Extended Mobile IP and support for Global Connectivity in Hybrid Networks

Christer Åhlund

Luleå University of Technology
Department of Computer Science and Electrical Engineering, Division of Media Technology
Extended Mobile IP and support for Global Connectivity in Hybrid Networks

Christer Åhlund

Division of Media Technology
Department of Computer Science and Electrical Engineering
Luleå University of Technology
SE-971 87 Luleå
Sweden

January 2005

Supervisor
Professor Arkady Zaslavsky

http://folk.uio.no/paalee/
# Table of Contents

Chapter 1. Thesis Introduction and Summary ...............................................................1  
  1.1. Introduction ..................................................................................................1  
  1.2. Summaries of included published papers .....................................................5  
  1.3 Mobile City ...................................................................................................8  
  1.3 Chapter Summary .......................................................................................9  

Chapter 2. Overview of current developments in IP network mobility and global connectivity .................................................................11  
  2.1 Mobility vs. Portability ..............................................................................11  
  2.2 Mobile IP .....................................................................................................11  
  2.3 Ad hoc Networks .........................................................................................15  
  2.4 Global Connectivity and Hybrid Networks ................................................16  
  2.5 Chapter Summary ......................................................................................18  

Chapter 3. State-of-the-art in Mobile IP and Global Connectivity literature ..............19  
  3.1 Mobile IP and handover related issues .......................................................19  
  3.2 Global connectivity and Wireless Hybrid Network issues ...........................28  
  3.3 Chapter summary ....................................................................................36  

Chapter 4. Software Solutions to Internet Connectivity in Mobile Ad Hoc Networks  ...........................................................................................................................39  
  4.1 Introduction ..................................................................................................41  
  4.2 Connecting IP Networks with Ad hoc Networks ........................................42  
  4.3 Related work ...............................................................................................49  

Chapter 5. Multihoming with Mobile IP ..................................................................51  
  5.1 Introduction ..................................................................................................53  
  5.3 A simulation study .......................................................................................59  
  5.4 Related work ...............................................................................................61  
  5.5 Chapter summary .......................................................................................62  

Chapter 6. Extending Global IP Connectivity for Ad Hoc Networks .........................63  
  6.1 Introduction ..................................................................................................65  
  6.2 Connecting IP Networks and Ad hoc Networks ........................................70  
  6.3 Related work ...............................................................................................82  

Chapter 7. Running Variance Metric for evaluating performance of Wireless IP Networks in the MobileCity Testbed .................................................................85  
  7.1 Introduction ..................................................................................................87  
  7.2 Running Variance Metric ..........................................................................89  
  7.3 RVM Simulations study .............................................................................91  
  7.4 Related work ..............................................................................................95  
  7.5 Chapter summary ......................................................................................97  

Chapter 8. M-MIP : extended Mobile IP to maintain multiple connections to overlapping wireless access networks .................................................................99  
  8.1 Introduction ................................................................................................101  
  8.2 M-MIP .........................................................................................................102  
  8.3 M-MIP analysis using RNL based simulation ............................................103  
  8.4 Related work .............................................................................................109  
  8.5 Chapter summary .....................................................................................110
Abstract

Mobility is an essential and necessary feature for roaming users who connect to wireless networks via access points. Access points may have different capabilities, be connected to different networks and be installed by different providers. A mobile host will discover multiple access points in this environment. A mobile host should be able to use the best available connection to communicate with a correspondent host and perhaps use multiple connections for different hosts. In areas with wireless local area network access, pockets with limited or no coverage could exist. Such restricted connectivity could be compensated by neighbor hosts who form an ad hoc network and relay packets until they reach an access point.

This thesis describes and discusses the proposed solutions towards enabling and supporting connectivity in wireless networks. In the proposed solutions the network layer software will evaluate and decide which wireless network connections to use.

This thesis proposes a Running Variance Metric (RVM) and a Relative Network Load (RNL) as performance metrics to classify the traffic load of access points in wireless access networks. RVM and RNL can be efficiently used for infrastructure networks and ad hoc networks.

This thesis proposes an extension to Mobile IP in order to enable mobile hosts to use multiple care-of addresses simultaneously. The extension enhances network connectivity by enabling the mobile host, the home agent and correspondent hosts to evaluate and select the best connection. The proposed extension to Mobile IP is called Multihomed Mobile IP to emphasize support for multiple connections for a mobile host at the same time.

This thesis describes a proposed gateway architecture that integrates wired IP networks with ad hoc networks. Routes between a mobile host and gateways are maintained continuously where multi hop ad hoc connections are supported. Communication between peers in ad hoc networks is based on reactive ad hoc routing. Mobile hosts moving between ad hoc networks are supported by Multihomed Mobile IP.

This thesis describes developed prototypes and simulation results to validate multihomed Mobile IP and the gateway architecture which integrates ad hoc networks and wired IP networks.
Publications

This thesis work has resulted in the following outcomes:


7. C. Åhlund, A. Zaslavsky, Integration of Ad Hoc Network and IP Network Capabilities for Mobile Hosts, "International Conference on Telecommunications”, ICT 2003, February 2003, Tahiti. This paper has been awarded best paper award.


Papers 1 to 4, 6 to 8 and 10 to 12 were peer-reviewed and accepted for publication at international conferences. Publication 5 is a licentiate thesis and paper 9 is published in a Journal. The contents of papers 6 and 8 to 12 are included in the thesis. The included publications have some modifications and have been reformatted for this thesis.

Outcomes of the project have been presented at the following seminars:

- Monash University, February 2000, Melbourne, Australia;
- The Path to 4G Mobility, September 2001, Helsinki, Finland;
- International Conference on Emerging Telecommunications Technologies and Applications, October 2001, Kosice, Slovak Republic;
- Monash University, December 2001, Melbourne, Australia;
- International Conference in Mobile Open Society through Telecommunication, October 2002, Warsaw, Poland;
- Norway University of Science and Technology, September 2002, Trondheim Norway;
- Luleå University of Technology, August 2002, Luleå, Sweden;
- Licentiate Seminar, Luleå University of Technology, December 2002, Skellefteå, Sweden;
- International Conference on Product Focused Software Process Improvement, December 2002, Rovaniemi, Finland;
- International Conference on Telecommunications, February 2003, Papeete, Tahiti;
- Monash University, February 2003, Melbourne, Australia;
- Telenor, May 2003, Oslo, Norway;
IEEE International Conference on High Speed Networks and Multimedia Communications, July 2003, Estoril, Portugal;
Umeå University, January 2004, Umeå, Sweden;
Luleå University of Technology, May 2004, Luleå, Sweden;
International Workshop on Service Assurance with Partial and Intermittent Resources, August 2004, Fortaleza, Brazil;
Cisco Systems, November 2004, San Jose, USA;
Sheerbrooke University, November 2004, Sheerbrook, Canada;
Monash University, November 2004, Melbourne, Australia.
Acknowledgements

First, I would like to thank my supervisor Arkady Zaslavsky for his support, encouragement and advice as well as for always being available to answer questions, all of which made this thesis possible.

I am very grateful to my co-supervisor Johan Nordlander for valuable comments and good advice. I express my gratitude to Robert Brännström who helped me with running simulations and was always available to discuss results.

I would also like to thank Kaustubh Phanse, Anders Lindgren and Johan Kristiansson for reading the thesis and commenting upon the work and Gary Webster for helping me to make the thesis readable.

Most of my research has been funded by the MobileCity project, an EU-funded project, and I would like to thank Eva Lindberg and Skeria Utveckling, coordinator of MobileCity, for this. My research has also been funded through the Centre of Distance-Spanning Technology, CDT, within the NUTEK projects PROMODIS and IRES as well as the EU-funded project VITAL.

I also would like to thank the STINT project for supporting collaboration between Monash University and Luleå University of Technology and the project manager Sven Molin for supporting Arkady and me in our joint work.

My beloved family deserves my greatest gratitude for always supporting me in this work and showing grate patients and sacrifice and understating. Thank you my wife Marie and my children Andreas, Therese and Viktor. I also would like to thank my parents Ola and Doris, and my parents-in-law Erik and Stina for always being there when needed.

Finally I would like to thank my colleagues and friends at campuses in Skellefteå and Luleå for their support.

Thank you all!
This chapter introduces the thesis, gives a roadmap of the work and summarizes the publications included. A project named MobileCity is presented.

1.1. Introduction

The rapid developments in wireless computer communication technologies have made wireless communication a vital element for connectivity to wired Internet Protocol (IP) networks and the Internet, as well as for communicating peer-to-peer. The wireless local area network (WLAN) standardly referred to as 802.11[1] is a widely deployed standard. 802.11b [1] theoretically supports a throughput up to 11Mbps, 802.11a [2] and 802.11g [3] supports up to 54Mbps. To manage quality of service (QoS) in 802.11 networks, 802.11e [4] is proposed. Another proposal for a wireless standard 802.20 [5] will support mobility and QoS, giving a bandwidth of 1Mbps per user. The 802.16 [6] standard supporting up to 70Mbps will incorporate a proposal for 802.16e that will enable mobility. Both proposed standards, 802.20 and 802.16, are for metropolitan area networks. The proposed standard 802.15 [7] is for shorter distances and will be used in personal area networks.

The advent of high bandwidth wireless networks requires support for extended network protocols. Today wireless network access is provided by connecting to one access point (AP) at a time.

New functionality needs to be added to mobile hosts (MH) and wireless access networks to enable networking software to fully utilize the features and opportunities that come with wireless network access. Only then will MHs truly benefit from the dynamic behaviour of wireless communications.

To manage network mobility for an MH connecting to IP networks, where applications and users are unaware of the support of network mobility, the Mobile IP (MIP) [8] has been proposed. In MIP environments handover is managed at the network layer when an MH is moving between networks. In wireless network setups handover takes place when an MH moves from one AP (the old AP) to another (the new AP).

If the new AP is connected to the same local area network (LAN) or the same virtual local area network (VLAN) as the old AP, the handover will only occur at the datalink layer. No network layer handover is required. If however, the APs are connected to different networks, both datalink layer handover and network layer handover will take place. These two handovers usually occur without synchronization between the datalink layer and the network layer since the IP stack aims at limiting the interaction between layers to sustain modularity.
In WLANs, the datalink layer handover is based on monitoring the signal-to-noise ratio (SNR) and related factors by an MH, which decides when to change the point of attachment. The MIP network layer handover decision is based on agent advertisements. MIP specifies that if three advertisements in a sequence are lost, the network connection is considered lost and a handover will take place.

The decision of where to associate with a new AP at the datalink layer may not be the optimal decision considering network layer performance. Even with a high SNR value the performance at the network layer may be poor. To overcome this problem network layer performance characteristics need to be discovered and considered when deciding which AP to associate with. Synchronization between the datalink layer and the network layer is therefore needed.

There is also a need for the network layer to be informed of an abruptly broken wireless connection to lower the handover delay at the network layer. As defined by MIP it will take three times the agent advertisement time to discover a lost network layer connection. By signaling a lost connection at the datalink layer to the network layer, network layer handover can quickly take place avoiding broken flows and minimizing the number of lost packets.

Another new and promising networking technology brought about by the wireless capabilities is ad hoc networking [9]. Ad hoc networking enables MHs to create a peer-to-peer network on their own without a backbone infrastructure. MHs in ad hoc networks usually operate as both end-user hosts and routers. Ad hoc networks have so far been regarded as stand-alone networks. However, there is an increasing interest in research toward solutions for connecting ad hoc networks to wired networks.

In areas with WLANs it is difficult to avoid dead spots where there is no radio coverage. In such places it may be beneficial for an MH to use intermediate MHs to relay its traffic to and from the AP. Two MHs should be able to communicate with a routing distance of a few hops in the ad hoc network. Traffic should not have to be sent through the AP if a more efficient route can be found in the ad hoc network.

A dynamic network structure can be created by enabling mobility support between networks and making ad hoc networks a part of the Internet. To be able to benefit from this, solutions for gateways connecting wired IP networks and ad hoc networks are needed. Such would enable efficient connectivity for ad hoc nodes (MHs) to the wired IP network and the Internet. Assuming today’s traffic patterns within the Internet for ad hoc networks as well, we can say that 20% is for peers in the same network and 80% for destinations outside the network. Efficient connectivity to gateways is important. Therefore methods are needed to select the best connection to a gateway if multiple gateways are available, or, in the case of multiple routes, to a single gateway.

The work presented in this thesis enables a dynamic wireless network infrastructure. MHs are able to communicate with peers in ad hoc mode as well as to use a wireless network infrastructure. Mobility between networks is supported. The environment is considered trusted, where computer security and network security can be defined and established. While security is very important for wireless networks it has been left out of the scope for this thesis. Such environments include corporations, Universities, etc. All MHs in this environment are assumed to relay other MHs traffic. The work does not look at scenarios where MHs may refuse to relay traffic, or at the cost efficiency of relaying traffic.
In this thesis, the terms \textit{access point (AP)} and \textit{gateway} are used interchangeably. They are considered to host access router (AR) functionality as well as a foreign agent (FA) and sometimes a home agent (HA). It will be clear from the context which functionality is included. The work described in this thesis makes the following contributions:

- Proposes a \textit{Running Variance Metric (RVM)} and a \textit{Relative Network Load (RNL)} that are used to analyze the performance of APs in wireless access networks. The RVM and RNL can be used both for infrastructure networks and for ad hoc networks;
- Extends and enhances the Mobile IP (MIP) to enable mobile hosts (MH) to register with multiple foreign agents (FA) simultaneously. Such extension enhances the network connectivity by enabling the MH, the home agent (HA) and correspondent hosts (CH) to evaluate and select the best connection to the network. This MIP extension is called \textit{Multihomed Mobile IP (M-MIP)};
- Proposes a gateway architecture enabling global connectivity, and integrating ad hoc networks with wired IP networks, where the ad hoc network uses a reactive ad hoc routing protocol. Ad hoc connections will be used for connectivity to a gateway. MHs that move between different ad hoc networks are managed by the MIP. By combining ad hoc networks and wired IP networks using MIP for mobility management, a dynamic and adaptable infrastructure is enabled, enhancing the network support for mobile users;
- A number of prototypes have been designed and implemented as proofs for the concepts of M-MIP and the gateway connecting ad hoc networks and wired IP networks. These prototypes validate the proposed algorithms and concepts.

This thesis is organized in the following way. The remainder of chapter 1 gives a road map of published articles and summarizes the work. Chapter 2 gives a background of mobility and ad hoc networks. Chapter 3 presents the state-of-the-art research in this area. Chapter 4 is based on the published paper “Software Solutions to Internet Connectivity in Mobile Ad Hoc Networks” and describes the architecture and a prototype connecting ad hoc networks and wired networks. Chapter 5 incorporates the published papers “Multihoming with Mobile IP” and “Agent selection strategies in wireless networks with multihomed Mobile IP”. These papers propose an extension to the MIP that enables multihoming with MIP. Chapter 6 is based on the published paper “Extending Global IP Connectivity for Ad Hoc Networks” and extends the work presented in chapters 4 and 5. In this chapter multihoming in MIP is extended to work in ad hoc networks. Chapter 7 is based on the published paper “Running Variance Metric for evaluating performance of Wireless IP Networks in the MobileCity Testbed” and describes a metric used to discover the performance of APs at the network layer in order to benchmark the APs. Chapter 8 is based on the published paper “M-MIP: extended Mobile IP to maintain multiple connections to overlapping wireless access networks” and describes a simulation study where the metric introduced in chapter 7 is used with multihomed MIP for selecting the AP to connect to. Chapter 9 discusses the contributions of the conducted research. Chapter 10 concludes the thesis.
Chapter 1. Thesis Introduction and Summary

The thesis work has resulted in 11 publications at peer-reviewed international conferences and one journal paper. Those articles cited in the green boxes with thick border have been included in the thesis. These publications are grouped into chapters which are discussed and summarized below.

Figure 1.1. A roadmap of the thesis work.
1.2. Summaries of included published papers

This thesis identifies research challenges and proposes solutions to enhance the functionality and performance of mobile hosts (MH) connecting to wireless access networks and the Internet.

A Running Variance Metric (RVM) is proposed to compare the utilization of access points (AP) based on the network layer performance. This information is then used to select one or more APs to connect to.

Mobile IP (MIP) is extended to manage multihoming (M-MIP). With multihoming functionality an MH is able to evaluate the network layer performance of APs including the wireless link using RVM as well as the round trip time (RTT) between the MH and its home agent (HA) and correspondent hosts (CH). The metric used for this evaluation is named the Relative Network Load (RNL).

A gateway architecture is proposed to integrate ad hoc networks with wired IP networks. The usage of RVM, RNL and M-MIP is extended to ad hoc networks thus enabling global connectivity and hybrid networks.

The following sections contain brief summaries of the articles included in this thesis

1.2.1. Articles describing implementations, conceptual architectures and theoretical contributions

**Software Solutions to Internet Connectivity in Mobile Ad Hoc Networks [10]:** Ad hoc networks have been regarded as standalone networks without necessary connectivity to wired IP networks.

This paper proposes a gateway architecture connecting an ad hoc network with wired IP networks and the Internet using the Ad-hoc On-demand Distance Vector (AODV) [11] protocol. MIP is used to manage mobility of MHs between ad hoc networks. To sustain the IP network architecture an ad hoc network is considered a network with its own network number and which incorporates APs.

A prototype is described which is based on the HUT-MIP [12] implementation and AODV-UU [13]. The paper describes how MIP messages are managed in the ad hoc network, how an FA is selected and how a peer is identified in the ad hoc network and in the Internet.

The selection of AP is based on the RTT between an MH and APs, where MHs send Internet Control Message Protocol (ICMP) echo requests to discover the best AP to use. These messages are sent periodically for the comparison of APs responses. MIP messages are modified and extended to handle multiple hops.

**Multihoming with Mobile IP [14]:** With MHs connecting to WLANs, connections may be abruptly lost and frequent handovers may take place based on datalink layer decisions. When using MIP, the MH may have to register a new care-of address for each datalink layer handover. This registration process will take a certain time and packets could be lost. One solution to overcome this problem is to use
Chapter 1. Thesis Introduction and Summary

multihoming at the network layer and to control the selection of which AP to associate with as well as to control the handover from the network layer.

The work presented in the paper extends MIP so that an HA as well as the CHs can maintain multiple bindings for an MH. The MH’s home address will then be bound to multiple care-of addresses. To decide which care-of addresses to register at the HA as well as the CHs, the MH monitors the deviation in arrival times of agent advertisements. A number of care-of addresses supplied by the APs having smaller deviations will be registered. To communicate with the MH, the HA as well as the CHs can then select among the registered care-of addresses. The selection of care-of address is based on the RTT between the HA and MH and between CHs and the MH. The RTT measurements are carried out on the MIP messages. In this way no explicit ICMP messages have to be sent.

With multihoming it is possible to overcome the problems where an MH may abruptly lose its connection or frequent handovers between nearby APs happen even if the MH remains in the same position.

Extending Global IP Connectivity for Ad Hoc Networks [15]: This paper describes work combining the proposals given in the papers “Software Solutions to Internet Connectivity in Mobile Ad Hoc Networks” [10] and “Multihoming with Mobile IP” [14]. Multihoming is used by MHs in the ad hoc network to evaluate connectivity to APs in order to select the most appropriate AP when communicating with its HA and CHs. Agent advertisements are forwarded in the ad hoc network and routes to the APs are worked out based on those messages. Routes are also created and maintained based on agent solicitations, registration requests and registration replies. In this way we develop a combination of a proactive and a reactive approach in the ad hoc network. The maintenance of routes to gateways is managed in a proactive approach while routes to CHs are managed in a reactive way. In this way a tree structure of MHs connecting to an AP will be maintained, where the AP is the root of the tree. If multiple APs are available an MH may belong to multiple trees.

As in [14] the MH decides the care-of addresses to register based on the deviation in arrival times of agent advertisements. We also use the same approach as in [14] for the HA as well as the CHs to decide which care-of address to use.

The paper also proposes and describes the decision making algorithms and how they are used in the prototype. A simulation study presented in the original paper has been omitted in this thesis since the later publications include a more comprehensive study of the same simulation-setup and offer extended results.

1.2.2. Articles describing simulations

Agent selection strategies in wireless networks with multihomed Mobile IP [16]: This paper presents a simulation study which is based on algorithms and a prototype described in the paper “Multihoming with Mobile IP” [14] (see section 1.2.1). The study shows results from two different simulations: the first describes how the network layer performance at an AP can be affected by nearby MHs using the same
AP, without noticing this at the datalink layer; the second simulation shows how Mobile IP extended with multihoming functionality responds to network load.

Running Variance Metric for evaluating performance of Wireless IP Networks in the MobileCity Testbed [17]: This paper includes the formal aspects of the RVM and its simulation study. The RVM is shown to discover the load of APs in infrastructure networks as well as in ad hoc networks.

The RVM metric is used in MHs and it is calculated based on the deviation in arrival times of agent advertisements broadcast by APs. RVM reflects both an AP’s ability to forward traffic as well as the load in the wireless link. If an AP is heavily loaded, advertisements will be delayed in buffers and they may also arrive in more dense succession. In the wireless media, collisions increase with added traffic and the time-period to access the media will increase rendering a larger deviation in arrival times of agent advertisements.

For ad hoc networks, the study looks at how the RVM responds to advertisements relayed via multiple hops and whether RVM can be used to evaluate the utilization of routes. The results show that the RVM can be used to discover an increased distance in hop count for small ad hoc networks. At the same time the RVM reflects the utilization of routes.

The RVM can therefore be use to compare the performance of APs and the wireless media to discover the least loaded AP. However, RVM can not be used to compute the exact utilization of an AP or the exact load of wireless links. With RVM it is possible to rank the utilization of different APs.

To enhance the reliability of the RVM, a predefined marginal variance should be used when sending agent advertisements. This is to avoid a flow with low throughput that may collide with advertisements repeatedly if packets are sent at exactly the same time as agent advertisements. Since agent advertisements are sent using broadcast this will not be discovered by the AP and advertisements will be lost. In ad hoc networks this can be resolved by listening for nearby nodes retransmitting advertisements. If a sender cannot hear a node retransmitting an agent advertisement it can resend the advertisement.

M-MIP: extended Mobile IP to maintain multiple connections to overlapping wireless access networks [18]: This paper describes a simulation study using the RVM and RNL to decide an MH’s default gateway and the care-of addresses that the HA and CHs will use.

During the simulation study it was discovered that the Jacobson/Karels formula [19] did not respond to the rapidly changing conditions in the wireless access network in a timely way. The RNL metric represents the modified Jacobson/Karels algorithm to select the care-of address as well as the gateway. The distinction of the modified algorithm is the weight when adding the RTT and the deviation as well as the calculation of the deviation. The deviation calculation using RVM was shown to respond quicker and is therefore used to calculate the RNL.

The simulation uses RVM to decide which care-of addresses to register at the HA and CHs. For the decision of which care-of address to use when starting to send packets, the RNL is applied. To calculate the RNL, the RVM is added with RTT measurements of MIP registration messages and binding updates. In this way the decision of which care-of address to use will be based on both the wireless access network and the wired network.
All CHs using the HA to reach the MH will use the same care-of address since this will be the care-of address decided by the HA. However, with route optimization different care-of addresses may be used by different CHs.

The paper compares benefits of selecting the AP based on the signal-to-noise ratio (SNR) and selecting the AP based on network layer performance using the RVM and RNL. In all cases the selection based on network layer measurements was more efficient. This is based on the fact that the SNR does not clearly reflect the throughput at the network layer. With a good SNR metric the network layer performance can still be poor.

The simulation also shows the effects of continuously monitoring the RVM and RNL. When traffic is sent to a selected AP it will impact the metrics. If the difference between two APs is lower than the increase of metrics caused by the added traffic, false handover will occur. With false handover an AP could be selected which may not have the lowest metrics minus the impact of the own traffic of an MH. The paper proposes a method to avoid false handovers.

1.3 MobileCity

This section discusses the MobileCity project where most of the prototypes and simulations have been validated and tested. MobileCity (www.mobilecity.nu) is an EU funded research and development project that develops both a wireless communication infrastructure and diverse mobile applications that can support various activities of individuals as well as communities. The project involves an extensive wireless infrastructure and a number of research initiatives in mobile communications and applications development. The real city in the focus of this project is Skellefteå, located in the northern part of Sweden in the Internet-Bay area. Skellefteå is a typical regional community in a sparsely populated area of the country (see figure 1.2), with well-developed infrastructure, mostly small and medium-sized enterprises and a broad range of applications that users might need. Within MobileCity a wireless access network based on 802.11 has been developed. This network spans different locations in the city of Skellefteå, including a university campus area, city centre, hotels and places visited by

Figure 1.2. Map
tourists. In the summer of 2004 the testbed was extended with an 802.16a wireless network covering sectors within and around the city.

A campus building hosting lecturers, researchers and students is being used for the prototypes developed in this thesis. The current wireless infrastructure has some areas with no coverage (dead spots). To fix this problem a prototype has been built that manages ad hoc networks as subnets and enables connectivity to multiple gateways to the Internet. Users will be able to communicate peer-to-peer and to use gateways for communication outside the ad hoc network. If an MH is in a dead spot area the adjacent reachable intermediate hosts will relay the traffic.

1.3 Chapter Summary

This chapter introduced the thesis objectives, presented the published work and a road map of this thesis. Articles describing conceptual architectures, theoretical contributions, prototype implementation and simulation studies have been briefly summarized. This chapter has also introduced the MobileCity project which is used as a testbed for validation of this thesis' results.

The next chapter will provide a background to mobility, Mobile IP, ad hoc networks and global connectivity.
Chapter 2. Overview of current developments in IP network mobility and global connectivity

This chapter presents background information on mobility in IP networks, Mobile IP, ad hoc networks and global connectivity.

2.1 Mobility vs. Portability

To manage movement of hosts connecting to the Internet, there are different possibilities depending on the requirements of a mobile host (MH). The expressions *portability* and *mobility* are used to describe MHs connecting to different networks. The portability concept refers to a device connecting to different networks and disconnecting when moving between them, while mobility describes MHs that maintain network connectivity also when moving.

For users connecting to foreign networks with the purpose of accessing the Internet, it will be enough to support MHs with the Dynamic Host Configuration Protocol (DHCP) [20]. When an MH connects to a network, the device will request an IP address as well as the network mask and the default gateway from a DHCP server. Usually, the address of the Domain Name Server (DNS) [21] is also provided. This information can then be used to access the network and the Internet.

For users requiring full access to the home network when connecting to foreign networks, Virtual Private Network (VPN) [22] can be used. A secure tunnel is created between the MH and the home network. VPN is mainly used to connect different geographical sites belonging to the same administrative domain, but separated by a public provider carrying other customer traffic (separate networks or a host connecting to the corporate network). A VPN can be created between a hot spot area and the home network of an MH. This kind of connectivity is usually considered to be long-term, where the mobile host connects to the foreign network and remains in the same place. To manage mobility in IP networks the Mobile IP (MIP) [8] is proposed.

2.2 Mobile IP

To manage mobility for an MH connecting to IP networks, where applications and users are unaware of the network mobility, the MIP is deployed.

MIP architecture incorporates a home agent (HA), a foreign agent (FA) in MIPv4 [23], and the MH. An MH connected to the home network will operate according to normal IP network operations, without using the MIP. When an MH connects to a
foreign network it will register its new location with the HA based on the address used by the MH in the foreign network. This address is called a care-of address or co-located care-of address depending on how addresses to visiting MHs are managed in the foreign network, and depending on the MIP version. In MIPv4 the address given a visiting MH can be supported by an FA (care-of address) or the DHCP server (co-located care-of address). In MIPv6 [24] there are no FAs so a co-located care-of address is always used and managed by IPv6. In IPv6 there are two options for an MH to receive an address: first, via stateless auto-configuration, based on the neighbor discovery protocol (NDP) [25], and second, via stateful process using a DHCP server. In this chapter, the care-of address is also used to describe a co-located care-of address. It will be clear from the context which is discussed below.

The registration sent to the HA by an MH connecting to a foreign network will create a binding in the HA, between the home address and the care-of address. If an FA is used, the MH sends the registration through the FA. The FA registers the MH in its visitor list and forwards the registration to the HA. In MIPv4 the registration is named a registration request and a registration reply is returned by the HA. In MIPv6 it is named a binding update and a binding acknowledgement is returned in response.

When packets for the MH are received at the home network, the HA will forward the packets to the care-of address using tunneling. A tunnel encapsulates the received packet for an MH as a payload in a new packet with an IP header having the care-of address as the destination and the HA as the source. When the packet arrives at the FA or the MH (if no FA is used) at the foreign network, the packet will be decapsulated by the networking software. If an FA is used the packets will be sent in a frame, the last hop with the MAC address stored in the visitor list for the MH. The MAC address is registered when the MH makes the registration by looking at the source address in the received frame. If the MH itself manages the encapsulation, the outer packet header is stripped off before being handed to upper layers.

The interception of packets by the HA for an MH, is based on the Address Resolution Protocol (ARP) [26] in IPv4 or the NDP in IPv6. A router connected to a network receiving a packet for a destination in the network, or a source in the same network as the destination, will translate the IP address to a Media Access Control (MAC) address. The MAC address is used to send a frame containing the packet the last hop to its destination. The ARP and the NDP are used to send a request for a MAC address containing the IP address. The host configured with the IP address will respond with its MAC address. When an MH has registered a care-of address at the HA, the HA will respond to the ARP and NDP requesting a MAC address for the MH’s IP address. The returned MAC address will be the MAC address of the HA.

All hosts as well as routers connected to a network maintain caches for the binding between IP and MAC addresses. To prevent caches from keeping obsolete bindings when an MH registers at a foreign network, the HA sends a gratuitous ARP (IPv4) or a gratuitous Neighbour Advertisement message (IPv6), to update these caches with the binding between the MH’s IP address and HA’s MAC address.

In MIPv4, an MH visiting a foreign network sends packets to a CH using the MH’s home address as the source and the CH’s address as the destination. However, ingress filtering may cause problems that require packets to be tunneled (reverse tunneling) to the HA first, and then sent from the HA to the CH.
In MIPv6 a packet sent to the CH will use the MH’s care-of address as the source, and the home address will be added in the home address destination option. Since the addresses are topology-correct, ingress filtering is avoided. The CH receiving the packet will put the address in the home address destination option as the source address before handing it to the transport layer.

The routing created by MIP is referred to as triangular routing (see figure 2.1). Here packets from a CH are sent to the HA.

**Figure 2.1.** Triangular routing in Mobile IP.

The HA tunnels packets to the MH, and the MH sends them directly from its current location to the CH, making a triangle (if reverse tunneling is also used it is named quadrilateral routing). To optimize routing between the MH and a CH, route optimization is used (see figure 2.2).

**Figure 2.2.** Route optimization in Mobile IP.
A CH sending packets to an MH is informed by the HA in MIPv4 or the MH in MIPv6 about the care-of address used by the MH. When a CH receives a binding update it will start to send packets directly to the MH using the care-of address as the destination address. In MIPv4, route optimization is seldom used because of security problems, and because the CH must be MIP-aware to manage tunneling to the MH. Currently there is no active proposal for route optimization in the Internet Engineering Task Force. In MIPv6, support for route optimization is built into the IPv6. The CH will use the routing header in IPv6, where the destination of the packet is the MH’s care-of address and the address in the routing header is the MH’s home address. When the MH receives such a packet, the destination field will be updated with MHs home address before handing the packet to the transport layer.

To secure route optimization in MIPv6 the return routability procedure is used (see figure 2.3). For this, four messages are sent; Home Test Init (HoTI), Care-of Test Init (CoTI), Home Test (HoT) and Care-of Test (CoT).

Before the MH send a binding update to the CH it sends a HoTI message to the CH through the HA. A CoTI message is also sent directly to the CH according to IP routing. When receiving these messages the CH responds with the HoT and CoT messages, where HoT is sent through the MH’s HA and the CoT message is sent directly to the MH’s care-of address. The MH derives a binding management key from the information in the HoT and CoT messages. After this, the binding update will be sent to the CH. The CH will derive the binding management key from the information in the binding update.

The return routability procedure verifies that the MH is reachable through both its home address and its care-of address. To secure the information exchanged in the return routability procedure, IPSec [27] can be used between the MH and it’s HA for the HoTI and HoT messages. A malicious node has to intercept both HoT and CoT messages to create the binding management key.

Return routability is required each time the MH changes care-of address. As long as the same care-of address is used the same binding management key is valid.

![Figure 2.3. Return routability in Mobile IP.](image-url)
Chapter 2. Overview of current developments in IP network mobility and global connectivity

The Mobile IP solution is attractive since it enables mobility with the most widely used protocols at the transport layer, the Transport Control Protocol (TCP) and the User Datagram Protocol (UDP). Protocols above the network layer are unaware of network mobility. One problem with MIP is the registration time when moving between networks, especially if there is a long distance between the HA and the foreign networks visited by an MH. MIP will probably be most used with MHs connecting wirelessly and this may cause problems because of rapidly changing conditions in the wireless network. An MH switching between APs connected to different networks will require a new registration each time.

The time it takes doing handover may cause UDP packets to be dropped and TCP flows to break.

2.3 Ad hoc Networks

A promising new networking technology brought about by the wireless capabilities is ad hoc networking. Ad hoc networking enables MHs to create a multi-hop network on their own without a backbone infrastructure.

Ad hoc networks consist of MHs which communicate wirelessly without support from a wired backbone infrastructure. Every MH within the network is both a host and a router for other MH flows. There are two main types of ad hoc routing protocols: proactive and reactive. In proactive protocols routing tables are maintained for the topology, continuously exchanging route information between the MHs, regardless of whether user data is sent or not. Reactive ad hoc routing protocols create routes on request by a source to send data.

Examples of proactive routing protocols include the Destination-Sequenced Distance Vector Protocol (DSDV) [28] and Cluster Switch Gateway Routing (CSGR) [29]. DSDV is a distance-vector routing protocol that transmits route updates on a periodic and event-driven basis. DSDV uses sequence numbers to avoid routing loops. MHs in a DSDV network are managed as a flat address space without routing hierarchy. CSGR is based on DSDV and creates a hierarchical topology using clusters, where the cluster-heads create a wireless backbone.

Dynamic Source Routing (DSR) [30] and Ad Hoc On-Demand Distance Vector Protocol (AODV) [11] are examples of reactive routing protocols. In these protocols, the source sends a route request and waits for the destination to respond with a route reply. At the reception of the reply, packets are forwarded by the source.

DSR uses a routing header for each packet that is added by the source, so intermediate MHs do not need to keep the routing information.

The AODV protocol is a development from DSDV based on symmetric links, managing both unicast and multicast routing. When a sender requests to send packets, a route request for the destination is broadcast in the network. When an MH rebroadcasts a route request it creates a reverse route to the sender of the route request. This route will later be used for the route reply. The destination receiving a route request, will reply with a route reply to the sender. The route reply is sent using the reverse route created by the route request. Upon receipt of the route reply at
intermediate MHs and at the source of the route request, a forward route is created and packets can be sent.

If an intermediate MH receiving a route request knows a route to the requested destination, it can respond with a route reply if the sender allows it. This is usually allowed for UDP flows where there might be no reverse traffic, but for TCP flows where at least acknowledgements will be returned, the route request should be sent all the way to the destination. If an intermediate MH responds, the destination may not have a route for its acknowledgements and have to invoke a new route request. Another option for this is to use gratuitous route reply.

When a route breaks because of a node movement or a bad communication link, the upstream MH noticing the break sends a route error to the source. The route error traversing MHs to the source will inform about the lost connection and all routes using the link will be erased. Created routes are soft-state and expire when not used for some time. To maintain neighbour relationships with nearby MHs, “hello” packets can be used.

2.4 Global Connectivity and Hybrid Networks

Global connectivity and hybrid networks are two terms used to describe connectivity of ad hoc networks to wired IP networks and the Internet. The ad hoc network can use both proactive routing and reactive routing. Research within the area mostly considers a reactive ad hoc routing protocol. Therefore this section gives an introduction to the area using a reactive routing scheme.

![Figure 2.4. Route discovery of a destination in the ad hoc network.](image-url)
Chapter 2. Overview of current developments in IP network mobility and global connectivity

This is a new research area that just recently gained interest in research communities and not yet been widely deployed. The background is given as it was proposed in [10].

An MH within the ad hoc network sends a route request to discover a peer regardless of the network address of the destination (see figure 2.4). In this scenario the MH is unaware of the connectivity to the wired IP network.

In figure 2.4, the gateway supports in the route discovery phase forwarding route requests. If the gateway knows that the destination is within the network it will operate as if it was an ad hoc node.

If the gateway discovers that the destination is in the wired network, it sends a route reply to the source (see figure 2.5). The source will not see the difference between a gateway replying and the destination itself.

Figure 2.5. Route discovery of a destination in the wired network.

When packets arrive at the gateway for a CH in the ad hoc network the gateway will search for the destination in its routing table. If a route is available the packets will be forwarded according to the routing entry. If there is no such route the gateway will request one by sending a route request. When receiving the route reply, packets will be forwarded (see figure 2.6). This architecture preserves the functionality of reactive ad hoc routing protocols.

Another solution is for the MH to maintain gateway information. If the MH believes that the CH is in the same ad hoc network it will send a route request for the destination. If a route reply is received packets will be sent using the ad hoc route. If there is no response from the CH a route to the gateway is identified and packets are sent there.
2.5 Chapter Summary

This chapter introduced background information of the areas where the major research described in this thesis has been carried out. Mobility is introduced and MIPv4 and MIPv6 are described. Ad hoc networks, global connectivity and hybrid networks are also presented.

The next chapter will describe state-of-the-art research in these areas using MIP and connecting ad hoc networks to wired IP networks and the Internet.
Chapter 3. State-of-the-art in Mobile IP and Global Connectivity literature

This chapter presents state-of-the-art research with Mobile IP and global connectivity. Research challenges are identified and discussed.

3.1 Mobile IP and handover related issues

This section describes proposals to enhance the functionality of MIP.

Hierarchical Mobile IP

Hierarchical Mobile IP (H-MIP) [31] is used to decrease the handover time in MIP. With an increasing distance between an MH and its HA the time for network-layer-handover will increase as well. To overcome this problem H-MIP is proposed. H-MIP uses a Mobility Anchor Points (MAP) in the access network, managing the mobility within the network (see figure 3.1). When an MH enters a wireless access network it must obtain two addresses: a regional care-of address (RCoA) and an on-link care-of address (LCoA).

The MH registers a binding between the RCoA and LCoA in the MAP. A binding update is also sent to the HA for a binding between the MHs home address and the RCoA. Based on this bindings a tunnel is created from the HA to the MAP, as defined by the MIP from the HA to the care-of address. Another tunnel is created from the MAP to the LCoA.

A packet sent to the MH’s home address is intercepted by the HA and forwarded to the MAP. The MAP de-capsulates the packet and tunnels it through a second tunnel to the LCoA. When an MH moves within the coverage of an MAP with changing LCoAs, the MH sends a binding update to the MAP for a binding between then RCoA and the new LCoA. In this way mobility can be managed locally within an access network keeping the time to register a change of location small, even with a long distance to the HA.

The same RCoA is maintained at the HA as long as the MH is within the network managed by the MAP. When moving into a new MAP’s coverage area rendering in a new RCoA, the MH has to obtain both a new RCoA and LCoA. This requires the MH to register with both the MAP and the HA. Route optimization is possible by sending the RCoA in a binding update to the CH.
Figure 3.1 Hierarchical Mobile IP.

Multiple MAPs can be active in an access network. Then an MH can register with more than one MAP and use different RCoAs with different CHs. MAPs are announced in agent advertisements possibly containing multiple MAPs. The reason using multiple MAPs is that if an MH communicates with a CH at the same layer in the network hierarchy, packets should not have to be sent to the top-most MAP. Instead, the closest MAP to the MH and the CH should be used.

An MH sending packets will tunnel them to the MAP, which de-tunnels arrived packets and handles them according to the proposed MIP standard.

H-MIP will isolate signaling to the access network and enable faster handover compared to when only using MIP. However, as with MIP, the same problem occurs when wireless links are used since there is no co-ordination between network layer and the datalink layer. The network layer performance is not considered when deciding AP to associate with. Also this requires big access networks in control of one
operator (one autonomous system) to enable this isolation of signaling. H-MIP limits the time for signaling, although the time for handover within the network is still the same.

**Fast handover**

Fast handover [32] proposes a solution to decrease the handover time with MIP. It is a proposal that uses signaling between the MH and the APs to keep the time for handover as low as possible. The proposal improves the address resolution time by pre-configuration. To enable this, seven messages are used for the signaling between the MH and APs. The messages are:

- Router Solicitation for Proxy Advertisement (RtSolPr);
- Proxy Router Advertisement (PrRtAdv);
- Fast Binding Update (FBU);
- Fast Binding Acknowledgement (Fback);
- Fast Neighbor Advertisement (FNA);
- Handover Initiation (HI);
- Handover Acknowledgement (Hack).

When an MH does a handover between APs, a tunnel is created between the old and the new AP for the time it takes for the MH to update its HA and CHs with the new care-of address. During this time packets from the HA and CHs (in the case of route optimization) are sent to the old AP according to MIP and then tunneled from the old AP to the new AP. Packets from the MH are tunneled from the new AP to the old AP and sent from there to the CHs.

To enable this functionality, the datalink layer signals to the network layer prior to handover at the datalink layer. Two messages are first used (PrRtAdv and RtSolPr) to resolve the datalink layer information of the new AP and network specific information. The MH sends an RtSolPr message, including the datalink layer identifiers that the MH has discovered at the new AP to the old AP. In response, a PrRtAdv is returned by the old AP informing the MH of the network specific information. Based on this information the MH creates a care-of address at the new AP and sends a FBU message to the old AP. The old AP then creates a tunnel between itself and the new AP. In response to the FBU message an Fback is returned. This message is sent by the old AP both in its own network and to the new AP. The new AP will forward the Fback message into its network. In this way the Fback message will be received at the MH from either of the APs. If the MH receives an Fback from the old AP the care-of address for the new AP has been sent as the proposed address in a HI message from the old AP to the new AP. The response was then returned in a Hack message. If the new AP rejects the proposed care-of address it will return another care-of address in the Hack message to the old AP. This address is then returned in the Fback message from the old AP to the MH. Packets are now tunneled until a FNA message is sent from the MH to the new AP. With this message the tunnel is erased and traffic is sent to the new care-of address according to MIP. Should the connection to the old AP be lost before the FBU-Fback message exchange is finalized, an FNA message which includes an FBU message can be sent by the MH.
to the new AP. The FBU-Fback messaging is then sent between the old and new AP and managed by the new AP.

A scenario showing the signaling of the fast handover when it operates normally according to the proposal is viewed in figure 3.2.

The fast handover proposal is for minimizing packet drops during handover by pre-configuring and forwarding packets from the old AP to the new AP. The solution proposes signaling between the datalink layer and the network layer. With signaling from the datalink layer the requirement to handover at the network layer will be discovered quickly, and in time to arrange for packets to be forwarded from the old to the new AP while MIP is being updated with the new care-of address.

While this is a good solution, it requires intelligent devices in the access networks and requires the wireless infrastructure to be modified.

![Fast handover diagram](image)

**Figure 3.2.** Fast handover.

**S-MIP**

S-MIP [33] presents a solution that builds on H-MIP and the fast handover mechanism. S-MIP is expressed to provide loss-less seamless handover with a similar handover delay as with 802.11. It was shown in [34] that by combining H-MIP and fast handover mechanism that the handover delays are between 300-400 milliseconds. To lower the delay S-MIP extends this approach. S-MIP is also expressed to handle the ping-pong effect that may happen with a combination of H-MIP and fast handover.
To keep packet losses small the MAP from H-MIP is put as close to the MH as possible. S-MIP uses a combination of Synchronized-Packet-Simulcast (SPS) and a hybrid handoff mechanism being an extension of the fast handover mechanism. With SPS the same packet is duplicated and sent to multiple APs: the AP currently attached to and the presumptive new AP(s) to associate with. The hybrid handoff lets the MH decide when to perform a handover and the access network decides where to perform a handover. The decision of where to handover is based on movement tracking in the network. Three conditions for movements are defined: linear movement, stochastic movement and stationary in overlapping cells. Stationary and not within overlapping cells need no consideration since no handover is possible. To enable movement tracing the MH monitors the signal strength from APs. This signal strength is reported to a Decision Engine (DE) that monitors the position of MHs and traces the mobility pattern. The DE is connected to the MAP used in H-MIP. When the MH signals that it wants to handover because of low signal strength at the AP associated with, the DE sends messages to the APs as well as the MH describing where to handover. During the handover process SPS is used to duplicate packets to the old AP as well as the AP(s) to associate with. Depending on the mobility pattern one or more APs may be available for the MH to associate with.

As with the fast handovers mechanism, packets (not being SPS packets) will be tunneled from the old AP to the new AP. The new AP buffers the SPS packets while older packets are tunneled from the old FA. These packets will be forwarded first by the new AP to the MH. Then the SPS packets will be sent to the MH. In addition, the old AP will send all packets on the wireless network. When all tunneled packets are delivered to the MH the new AP will notify the MAP to terminate the SPS and only forward packets to the new AP. To support the DE in making a good decision of what AP(s) to handover to, each AP sends information to the DE about how many MHs it is managing.

S-MIP enables fast handover and decreases the number of packets dropped. However it requires complex signaling and requires a complex access network. Also the MH cannot by itself decide which AP to use. In this way the ability for policy-based control is disabled. In an area with multiple operators and networks the user has to stay with the same provider. The number of associated MHs is used as the metric to decide the load of APs. This is not a good metric since MHs may be idle not sending any traffic while a few MHs communicates heavily. In this way an AP managing many MHs may be lightly loaded while an AP only serving one MH is heavily loaded.

**P-MIP: Paging Extension for Mobile IP**

P-MIP [35] is proposed to lower the MIP signaling for mobility in networks. To enable this paging areas are used, where a paging area consists of multiple APs. With MIP an MH registers with a new network (receiving a new subnet-prefix in an agent advertisement) despite being active (sending/receiving data) or idle. P-MIP proposes that idle MHs only should register when entering a new paging area. When moving within a paging-area with different networks no registration is required. An active node however needs to register accordingly to the MIP specification. The APs need to
Chapter 3. State-of-the-art in Mobile IP and Global Connectivity literature

keep information of MHs connected to it, if it is active or idle. The MH also keeps this information about itself. An MH sending or receiving data is in active state and after traffic stops it will remain in this state for an active timer period. If no further traffic is sent within this time period, the timer will expire and the MH will go into idle mode. If more packets are sent or received before the timer expires, the timer is reset to the active timer period.

An FA signals in the MIPv4 agent advertisement if it supports paging and the MH does the same in the registration request.

When packets for an MH arrive at its FA the FA controls if the MH supports paging. If paging is supported the FA will check the status of the MH. If it is in active mode packets will be sent to the MH according to MIP. If the MH is in idle state a paging request message will be sent by the currently used FA to all FAs within the paging area as well as broadcasted in its own network. All FAs receiving the message will also broadcast the paging request message in their network. If the idle MH is within the network where it is registered, it will respond with a paging reply to the FA. When receiving the reply the FA knows that the MH is in its network so packets can be sent there. If the idle MH has moved to the coverage of another FA, the MH will first register the new care-of address with its HA and then send a paging reply to the old FA through the new FA. Buffered packets at the old FA will be forwarded to the new FA and sent from there to the MH.

P-MIP addresses a scenario that frequently will be the case for MHs. As stated in the paper an MH will be in idle mode most of the time. The work saves energy and lowers the number of messages sent in the network. The drawback is the time it takes for the paging process, rendering a delay when delivering the first packets. An FA also needs to know other FAs within the same paging area. The benefit of the proposal also relates to how often MHs need to register with the FA and HA.

**Horizontal and Vertical handoffs**

The publications “Handoffs in Cellular Wireless Networks: The Daedalus Implementation and Experience” [36], “Vertical Handoffs in Wireless Overlay Networks” [37] and “Policy-enabled Handoffs Across Heterogeneous Wireless Networks” [38] propose various methods for horizontal and vertical handoffs.

The work [36] presents a method for horizontal handoff. Handoff latency between nearby APs of 8-15ms was achieved. The proposal works as MIP with the extension of using multicast to APs. With multicasting multiple nearby APs can be addressed.

The MH monitors the signal strength from APs and requests the APs where it may handover to join the multicast address assigned the MH. One AP is selected and it is instructed to forward packets while the other APs are told to buffer packets. The buffering APs drop packets in a FIFO order when the buffer becomes full (that is if they are not delivered to the MH). When doing handover the MH requests the old AP to buffer packets and the new AP to forward packets. At most one AP at a time is in forwarding state. In [37] a solution for seamless handovers in overlay networks is presented. An overlay network is defined as a combination of wireless networks, spanning in-room, in-building, campus, metropolitan and regional cell sizes.
Horizontal handoff is expressed as handover within the same network architecture and vertical handover is between different network technologies. Horizontal handover is managed by choosing the AP with the best signal strength. For vertical handover the signal strength cannot be used, because a low signal strength of an in-building WLAN may perform better than a wide-area network with high signal strength. Based on this the solution proposed for vertical handovers always uses the technology in the lowest available layer in the overlay network, usually providing the best throughput. A handover to a higher layer occurs when there are no signals from the lower layer, and vice versa. [37] extends the work presented in [36].

A solution to policy-enabled handoff is presented in [38]. The solution is based on three policies; power consumption using the network access device, cost and bandwidth that can be offered. The bandwidth usage is monitored and announced by APs so that MHs can calculate the utilization of a network. This is then used to determine which AP to use. This work is extended work to [36].

In [36,37] heterogeneous handover is managed by monitoring the signal strength and based on this information, to decide which the AP to use. The delay for MIP registrations required when changing between APs in different networks is avoided with the usage of multicast. However APs may need a big buffer-space with many MHs in the access network. Also APs must be extended with a new signaling mechanism and they must be multicast aware. In [38] APs monitor the bandwidth usage and announce it in beacons, enabling an MH to choose the least loaded AP.

**Intelligent Handoff for Mobile Wireless Internet**

This work [39] addresses the problems concerning lack of synchronization between datalink layer handover and network layer handover using MIP. A proposal is given called the “Intelligent handoff architecture”. The proposal consists of three extensions:

- Packet buffering;
- Neighbor list update;
- Datalink layer handoff notification.

Packet buffering enables the MH to tell the old FA to buffer packets for it prior to doing handover. After associating at the datalink layer with the new FA and registering the new care-of address at the HA, the HA sends a re-route message instructing the old FA to forward packets to the new FA.

A neighbor list update is used to inform an MH about neighbor agents to make handover faster and to avoid packet-drops. The neighbor list is announced in agent advertisements and contains information about the IP-address, datalink layer type and radio frequency information for each agent. With this information an MH can discover if handover is possible to an FA attached to the same network as the old FA. Selecting that FA avoids having to update its binding at the home network with a new care-of address.

Datalink layer handoff notification is used to avoid having to wait for three times the agent advertisement time as specified in MIP. With notification from the datalink
layer to the network layer the MH can register with a new FA immediately when a connection is lost with the old FA.

The work also addresses how to discover the foreign network closest to its home network. The proposals include:

- The MH can download the routing table from the access router;
- Traceroute can be used;
- Estimation can be made by increasing the distance to the HA by one each time handover occurs.

The proposal to use datalink layer signaling improves the network layer handover. Neighbor lists are required which means that this proposal requires that all APs are managed by the same autonomous system (AS). A solution is proposed to discover the closest care-of address with regard to the distance to the HA. This requires extra signaling and complexity especially if the AP has to interpret the routing tables downloaded from the access network.

### Multiple Care-of Address Registration on Mobile IPv6

The work [40] presented proposes multihoming in MIPv6. A single interface or multiple interfaces can be used. In IPv6 an interface can be configured with multiple IP addresses. Each interface is identified with a unique identification number (IFID) and an identification priority value (IFPRI). With multihoming using a single interface, multiple pseudo-interfaces are used. A pseudo-interface is created for each address and each one is given an IFID and IFPRI. With multiple interfaces each interface is configured with an IFID and a IFPR.

A primary interface is defined as the interface assigned the primary care-of address being the address registered at the HA. The primary interface is the interface given the lowest IFPRI value. Only the primary care-of address is registered at the HA. Interfaces with higher IFPRI values are called non-primary interfaces. A CH is updated with both the primary and non-primary interfaces. If the primary interface fails the non-primary interface with the lowest IFPRI value will be chosen. The IFID and IFPRI are sent in the binding updates. The return routability procedure takes place by sending one CoTI for each care-of address the MH uses.

It is not defined on what conditions the IFPRI value should be set. Only CHs are enabled to maintain multiple bindings, while the HA only keeps the binding for the primary care-of address. This requires the full registration process with the HA to take place when doing handover. If the HA is enabled to keep the non-primary interface(s) a more fault-tolerant system will be achieved. Since the MH uses the same IFID to all MHs there can be no balancing by using different care-of addresses for different CHs.

### Dynamic Network Interface Selection in Multihomed Mobile Hosts

The paper [41] describes a multihomed approach to MIPv6 using simultaneous multi-access to enable to use different routes for different flows through heterogeneous interfaces to the same CH (or to different CHs), where policies are used to select the
preferred interface. To be able to send flows through different interfaces for the same CH the routing table needs to be modified. The usual way a routing table is implemented is by the destination address identifying a route to the destination. In this case, packets for a destination have to use the same interface and the same next hop. To be able to choose different interfaces for the same CH the source address (care-of address) is used in the MH to identify where to send packets for each flow. A CH having multiple connections to an MH will use multiple home addresses for the MH. In the CH it looks like connections to different MHs since it is still a one-to-one binding between a home address and a care-of address. This avoids having to change the MIPv6 protocol.

The proposed method has the advantage of avoiding having to change the MIP specification. However the routing function in the MH must be modified. The work also requires the DNS service to maintain multiple addresses for the MH and respond to DNS requests with those. Without that no multihoming functionality will take place. There is no proposal given in how to selects which care-of addresses to use.

**HAWAII: A Domain-based Approach for Supporting Mobility in Wide-area Wireless Networks**

HAWAII [42] uses Mobile IP for macro-mobility and proposes a solution for micro-mobility due to the latency problems with registrations to the HA when an MH moves within a domain. The MH keeps its care-of address when moving within the domain avoiding time consuming registrations with the HA. To manage mobility of MHs within the domain host routes are used for each MH.

A mobile node is allocated a co-located care-of-address when entering the domain, where each domain uses a root domain router. While moving within the domain the MH retains its address and the connectivity is maintained through dynamically established paths.

The domain is created as a tree topology using Open Shortest Path First (OSPF) [43] or some other routing protocol. While moving, the MH sends “path setup update messages” to establish and update host route entries for it. The path state is “soft state” and expires if not updated within a time period. An MH therefore sends periodic path refresh messages to the AP. The AP and the intermediate routers (between the AP and the root router) send periodic “aggregate hop-by-hop refresh messages” higher up in the tree.

HAWAII uses four path setup schemes to manage host routes in the routers within the domain:

- Multiple Stream Forwarding (MSF);
- Single Stream Forwarding (SSF);
- Unicast Non Forwarding (UNF), for mobile nodes that can listen to two APs;
- Multicast Non Forwarding (MNF).

MSF and SSF use forwarding that may result in temporary routing loops and perhaps reordered packets (MSF) but no packets are lost. The SSF needs an extension to work, the extension includes information about the interface where the packet is received and where it is being sent (like virtual links).
UNF and MNF do not forward packets from the node being left to the new node. To see all packets they need to be duplicated to adjacent APs.

The proposal requires frequently changing routing tables because of the host route approach, routing tables may therefore be big. The temporary routing loops will affect the performance in the access network.

3.2 Global connectivity and Wireless Hybrid Network issues

This section describes work connecting wired networks and ad hoc networks.

Gateway Forwarding Strategies for Ad hoc Networks

This paper [44] addresses different approaches for MHs in ad hoc networks to maintain connectivity with gateways connecting to the Internet. It is pointed out that maintaining host routes to destinations in the Internet by using proxy replies from gateways has the drawback that many entries have to be kept in the routing table. To aggregate these entries multihop default routes can be used.

If an MH discovers that a CH is in the Internet it will forward the packets to its configured gateway. This means that intermediate MHs may also have to make this discovery before forwarding the packets to its configured gateway. This may render in cascading look-ups, meaning that intermediate MHs will try to request a route for the CH before forwarding the packets to the gateway.

To overcome this problem a solution is to use host routes that points to the default router. However this approach may result in that a believed default route by an upstream node is actually being sent to another gateway, because a downstream MH uses another gateway (see figure 3.3). To overcome the described problems, it is proposed that tunneling should be used between an MH and its gateway. The benefits given are:

- Protocol transparency, because intermediate nodes do not need to be aware of gateways and tunnels;
- Route aggregation, since all destinations within the Internet uses the tunnel;
- Stability and reduced overhead. The gateway selected at an MH will not be diverted by intermediate nodes;
- Multiple tunnels to multiple gateways enabling load balancing and fault tolerance;
- Efficient forwarding compared to the default gateway approach. Since only two lookups are needed in the source and one in intermediate nodes;
- Security compared to default routes since intermediate nodes cannot divert the traffic to another gateway.

The paper shows by simulations that tunneling performs better than using default routes. In the simulation study AODV is used in the ad hoc network and MIP is used for visiting MHs to register with a gateway (FA).
Chapter 3. State-of-the-art in Mobile IP and Global Connectivity literature

Figure 3.3. The usage of tunneling to gateways.

This is a good study that addresses important aspects in the usage default gateways in ad hoc networks. However with multiple gateways no proposal is given in how to select the gateway to use for a destination. Also, the MH itself first has to discover if a peer is within the same ad hoc network or not. This increases the time to discover a CH.

MIPMANET-Mobile IP for Mobile Ad Hoc Networks

The MIPMANET [45] proposal enables Internet connectivity for MHs in ad hoc networks. MIPMANET uses hop counts to decide which FA to register with. A new FA is selected if the hop count is two hops less than the FA currently used.

MIP messages as agent advertisement, registration request, etc., are modified to manage multiple hops. When an MH registers with an FA a tunnel is created from the MH to the FA. If a CH cannot be found in the ad hoc network packets for it are forwarded to the FA using this tunnel.
A simulation study is presented comparing two approaches to discover FAs. In the first approach FAs respond to broadcasted agent solicitations with an agent advertisements sent by unicast, and secondly FAs broadcast agent advertisements periodically. The results from the simulation show that when having one or more visiting MH(s) the broadcast approach performs better.

MIPMANET gives a good proposal on how to decide what FA to use if multiple FAs are available. It also shows that the message overhead when using a broadcast approach for agent advertisement may be lower than if agent solicitations are used. However, the solutions do not consider the load of wireless links when deciding which FA to use. With the emerging heterogeneous wireless networks, hop count will not be the optimal parameter for decision of gateway connectivity. It does not reflect the capacity or the load of wireless links.

**Supporting Hierarchy and Heterogeneous Interfaces in Multi-Hop Wireless Ad Hoc Networks**

The work [46] proposes a technique allowing a single ad hoc network to span heterogeneous link layers. Ad hoc networks are integrated with the Internet and MIP is used to support mobility between networks. The proposal uses a single gateway and Dynamic Source Routing (DSR) [30] as the ad hoc routing protocol. It is proposed that ad hoc networks should be identified with topology correct network numbers in the Internet. Packets sent from a CH in the Internet will be routed according to the routing in the wired network and when arriving at the gateway, the gateway will forward the packet to the MH using the ad hoc routing protocol. If the FA receiving the packets has a route to the MH, packets will be forwarded immediately. If not, a path discovery will first take place and then packets will be sent using the newly created path.

An MH having packets for a CH will request a path to the CH if it does not have a route installed. If the node is believed to be outside the ad hoc network, the gateway will respond with a proxy reply (using a big hop count) saying that it can reach the CH. If the CH is within the ad hoc network it will respond with a reply. The MH will then receive two replies and choose the route leading directly to the CH without using the gateway. An MH discovering that it has moved to a new network will send a router solicitation requesting the gateway to respond with a router advertisement. This solicitation is piggy backed on a route discovery message. To avoid route discovery messages traversing multiple overlapping networks a flag is used. With this flag only nodes within a network are allowed to forward the route discovery message. A node in a nearby network can respond if it is the destination or if it has a route to it. However it cannot rebroadcast the route discover message.

The proposal gives a solution that enables the ad hoc network to keep its structure and for MHs within the ad hoc network to operate according to the ad hoc networking paradigm. No proposal is given on how to manage multiple gateways.
Internet Connectivity for Ad hoc Mobile Networks

The work [47] presented in the paper proposes connectivity between ad hoc networks running AODV and the Internet. MHs in an ad hoc network use their home address in foreign ad hoc networks. An FA-list is maintained in the MH for FAs sending agent advertisements. Duplicate advertisements are used if they arrive through a shorter path considering hop count than the previous advertisement. The routing information in the MH will then be updated using the shorter route. If the advertisements are received for the first time it will be re-broadcasted within a randomized time to avoid collisions. When an MH proactively wants to discover agents it issues a route request for the all mobility agents multicast group.

An MH receiving a route request for a FA will respond with a reply, if it has an address to a FA. Priority is given to the FA registered with. The reply is sent by unicast to the originator. If the MH does not have an FA in its FA-list, it will re-broadcast the request. When receiving the reply, the MH sends a router solicitation using unicast to the FA which in response reply with an agent advertisement. A registration request is then sent to the FA that will forward it to the HA, and the reply is forwarded from the FA to the MH.

When searching for a CH the MH broadcasts a route request. If an FA has a route table entry for the CH it will respond with a route reply. If not, a FA-reply is sent to the MH expressing that the CH can be found in the Internet using the FA as a gateway. The FA-reply has a big hop count value. Before using the FA the MH must make sure that the CH is not within the ad hoc network. For this, multiple route requests are sent according to the maximum number of attempts configured. If no reply is received from the CH the FA will be used.

In the case of multiple FAs the decision for handover occurs if the following two criteria are fulfilled:

• The MH has not heard from its FA for at least one advertisement interval;
• The MHs route to the FA has become invalid (because of route expiration or node movement).

When only receiving a reply from a new FA, that FA should be registered with it to maintain Internet connectivity.

Since there is no support from the gateway, whether the CH is within the ad hoc network or not, the setup time may be long. Also, no consideration is given to how loaded a path to a gateway is.

Internet Connectivity for mobile ad hoc networks

The paper [48] proposes general solutions to connecting ad hoc networks to the Internet. It is noted that care must be taken to insure that intermediate hosts do not try to discover a route to the destination. Otherwise packets may be rerouted in an inappropriate way for the sender.

A gateway can be discovered during the initial address configuration. When an MH needs to create an address using automatic address configuration it sends an address request to check its uniqueness. If some MH has this address it responds with
an address reply. A gateway seeing an address request for a site local address responds to the MH by sending a gateway advertisement. The same happens if the gateway receives a route request for a site-local address. This enhances the efficiency to configure an MH with a gateway.

Another solution proposed for finding a gateway is to send a route request using the Internet-gateway multicast address. A flag called the Internet-Global Address Resolution flag is used in this request. In response, gateways will send a reply, including its IP address and the prefix information. If a route request is sent for a global address with the Internet-Global Address Resolution flag, a gateway will respond in the same way and if the CH is within the ad hoc network it will also respond. This will return two replies to the MH. The route directly to the CH should then be used.

Yet another way to discover a gateway is to modify the agent advertisement and agent solicitation messages. Here, a flag is added to signal that the message should be re-broadcasted by MHs. An agent solicitation is sent to the Internet gateways multicast address and in response the gateways respond with an agent advertisement including the IP address of the gateway and the prefix information.

If a routing protocol does not maintain next hop information for a gateway, two routes are required, one for gateway and the other for the next hop. These routes should be maintained and updated before timing out. In the case of a routing protocol maintaining next hop default router, the information may become inactive. Then the information required is the default route information without needing to be updated.

If the routing protocol supports next hop forwarding to a default router, the MH relies on the intermediate MHs to have a route and the shortest path to a gateway. This may not be the same gateway as the sender is configured with. If next-hop forwarding is not used, packets are sent with a routing header for the gateway. If the gateway has a route for the destination pointing to the network where the packet came from, it will send an Internet Control Message Protocol (ICMP) [49] redirect message to the sender.

When sending packets to a CH in the Internet (and if both a host route and a gateway is installed in the routing table), the shortest route should be used. If the MH only have a default route when requesting a route to a CH, it first have to request a route for the CH before using the default route, since the host may be within the network. If the CH responds to the request it is in the network and packets will be sent using that route. If not, the default gateway will be used.

If an intermediate MH forwarding a packet to the gateway sees that the destination is within the ad hoc network, it is allowed to redirect the packet without knowledge of the source. However, it should update the source with a gratuitous route reply. If a routing header is used, this is not possible for an intermediate MH. In this case the gateway has to support this. When the source is notified of this it should send a route request for the CH.

When sending packets to a CH in the Internet, and if both a host route and a gateway are installed in the routing table, it is proposed that the shortest route should be used. However the MH does not have that knowledge since the total distance to the CH is unknown. To enable this the MH must also consider the routing information in the wired network. Otherwise the paper gives several good proposals on how to manage gateways in ad hoc networks.
AODV and IPv6 Internet Access for Ad hoc networks

This paper [50] proposes IPv6 Internet connectivity for ad hoc networks using the AODV routing protocol. For communication with gateways an Internet gateway multicast group is proposed. MHS requesting gateway information send a route request to the multicast group. When arriving at the gateway a route reply is returned with the address of the gateway, the global routable prefix information and gateway related information. The reply uses a flag. This flag informs MHSs that the reply carries gateway information. Upon reception at the MH it can auto-configure its global routable IP address and the gateway. Intermediate nodes do the same.

When an MH has packets for a destination it first looks in its routing table to see if it has a host entry for the destination. If it has, it will forward the packets accordingly. If not, it has two option; the first is to send packets to the default gateway, and if the destination is in the ad hoc network a redirect message will be returned; the other option is to send a route request for the destination and if no route reply is returned packets are sent to the gateway.

When sending packets to the Internet the sender can rely in intermediate MHSs to find their way to the gateway. Another way is to use a routing header. By using a routing header the sender knows that packets are not redirected on their way to the gateway.

Care must be taken to avoid cascading effects created if every intermediate MH sends a route request for a destination when sending packets to the Internet. Simulations showed that it may take 2.2 seconds to discover if an MH is in the Internet if expanded ring search is used. The experiments were carried out with an average hop count of three. With expanded ring search being disabled it took 0.09 seconds.

The study shows the importance of being supported by the gateway when determining if nodes are located in the ad hoc network or not. There is not proposal in how to decide the “best” gateway to use.

An Architecture for connecting Ad hoc Networks with the IPv6 Backbone (6Bone) using a Wireless Gateway

The paper [51] describes an IPv6 testbed connecting ad hoc networks to the Internet. AODV is used as the ad hoc routing protocol and MIP is used for mobility between networks. Two approaches for gateway discovery are evaluated and compared. One uses a proactive approach broadcasting HELLO messages containing a PROADV option. This message is re-broadcasted by intermediate nodes. For the reactive approach a GWSOL message is broadcasted by MHs that need to discover a gateway. In reply a gateway sends a GWADV. In the reactive approach GWSOL is managed like the RREQ message and GWADV as the RREP message. A flag is used to distinguish the GWSOL and GWADV messages from RREQ and RREP messages. A performance evaluation for using the proactive or reactive approach to discover a gateway is presented. For a small ad hoc network with a size of maximum 3 hops, no difference in the throughput was found with the two approaches. The study looked at different packet sizes and different number of hops sending TCP traffic.
The contribution of the study is that it shows that the throughput is the same with both the proactive and reactive approach. The network size used in the study is limited to a size that is operational with a wireless technology like 802.11.

**Integrating Mobile IP with Ad Hoc Networking**

In this paper [52], ad hoc networks are considered subnetworks running the DSDV ad hoc routing protocol. Mobility between networks is managed by MIP. Each gateway has a service range and all MHs within this service range can receive agent advertisements broadcasted by the gateway. A service-area is defined as a number of hops from a gateway. If an MH can hear two gateways it will select the gateway closest to the MH considering the number of hops. If an MH does not receive agent advertisements it can send agent solicitation to request an advertisement. The hop count used in solicitations increases each time no advertisement is received. When a gateway sees a solicitation with a larger hop count than its service range it will extend the service range.

The gateway is used when a sender does not have a route for the destination. Routing information is exchanged between MHs within a configured hop distance. Broadcasts are managed by tunneling to the gateway, and then a broadcast is sent within the ad hoc network the number of hops defined by the service range.

The approach uses proactive routing that may not scale for ad hoc networks with a frequently changing topology. The solution to select a gateway may result in a ping-pong effect due to changing conditions in wireless links. Otherwise, the solutions with service areas are appealing.

**A Hybrid Approach to Internet Connectivity for Mobile Ad Hoc Networks**

A combination of a proactive and a reactive gateway discovery mechanism is proposed [53]. Proactive discovery of gateways is enabled by periodically broadcasting MIP agent advertisements. The reactive approach broadcasts an agent solicitation message when traffic needs to be sent through a gateway to the Internet. AODV is used as the routing protocol in the ad hoc network and MIP is used to manage mobility between gateways. It is proposed that to limit flooding of solicitations they can be piggybacked on a route request message using expanding ring search. Intermediate nodes receiving such a message will respond with its gateway information.

An MH within the ad hoc network sends a route request for a CH. If the CH is within the ad hoc network it will respond with a route reply. A gateway believing the CH to be within the Internet (i.e. not having a route to the CH), will reply with a FA route reply. If the MH only receives FA route reply the CH is believed to be in the Internet and it will route the packet(s) through the gateway.

Routes are created based on the agent advertisement and agent solicitation messages. The coverage of agent advertisements is limited by a hop count. MHs outside this area need to broadcast agent solicitations to discover a gateway. In response, an agent advertisement is sent in a unicast message. Nodes close-by can
eavesdrop the message and discover a gateway, avoiding having to send its own agent solicitation message.

The simulation studies network sizes of 50 nodes in an area of 1000m x 1000m and 100 nodes in an area of 2000m x 2000m. CBR flows are used with 512 byte packets and 10 packets/second is sent. The source and sink nodes are randomly picked in the network. In the scenario with 50 nodes a beacon interval of 10-15 seconds showed good performance. With a 10 second advertisement time a hop count of 2 was found to perform best. A hop count of 5 is equal to flooding. In the scenario with 100 nodes a hop count of 4-5 performed best.

The paper shows that even with a big number of MHs, the major number of MHs will only be a few hops from the gateway, assuming normal radio range with 802.11.

**Hybrid gateway advertisement scheme for connecting mobile ad hoc networks to the Internet**

This paper [54] proposes two types of advertisement schemes. These schemes avoid generating unnecessary traffic and are based on traffic and mobility patterns. Following these schemes, an MH will be able to find the shortest path to the Internet.

An algorithm for adaptive advertisements is proposed, to avoid flooding the network. An MH maintains information of all gateways it receives agent advertisements from. In the study the DSR protocol is used. If an MH has gateway information, it can respond to a request from an MH for a gateway. An MH receiving gateway information from more than one gateway will only rebroadcast the agent advertisement with the shortest hop count to the gateway. If two advertisements have the same hop count the first arrived will be re-broadcasted.

The adaptive advertisement can limit the flooding area to only those nodes needing gateway connectivity. To enable this, MHs need to be aware of Internet joining nodes. This can be achieved by looking in the DSR header for packets passing by looking at the first and last hop external bit. When an MH discovers an Internet joining node it goes into advertise forwarding mode. The gateway calculates a Regular Mobility Degree (RMD). The RMD is calculated by dividing the number of Internet joining nodes with the number of agent forwarding nodes. The calculation of RMD can be done by examining the source route path to the target nodes. The RMD is compared to a threshold between zero and one. With a threshold of zero adaptive advertisements will be sent as soon as there is mobility. With a threshold of one, adaptive advertisement will hardly be used. When an MH receives an adaptive advertisement it re-broadcasts the packet if it’s in adaptive forwarding mode or if a counter in the advertisement is bigger than zero. A node not in adaptive forwarding mode decrements the counter before re-broadcasting the advertisement.

The simulation presented shows that the adaptive advertisements give less overhead, lower delay and slightly higher fraction of packets delivered, than with a reactive approach.


3.3 Chapter summary

The related state-of-the-art research presented in this chapter describes two architectures that are proposed for future wireless networking solutions namely MIP and ad hoc networking. The related work presented proposes solutions to overcome the deficiency when MIP is used in wireless access networks as well as solutions to connect ad hoc networks to wired IP networks and the Internet.

With MIP, the related work can be classified into proposing extensions to MIP in the following areas:

- Forwarding between APs during handover to avoid losing packets;
- Shortening the signaling to minimize the delay for handover;
- Signaling from the datalink layer to the networks layer when the connection at the datalink layer is lost to lower the delay for handover at the network layer.

Except for one proposal the decision of AP in infrastructure networks is based on the SNR and related factors. These metrics are good for discovering the distance to an AP, as well as if there is interfering radio signaling. With the usage of datalink-layer protocols where an MH requests a link and only transmits if it is granted access. The number of collisions will be low and not generate a bad SNR metric. For an MH out of communication range from other MHs sending traffic and using the same AP but within communication range to the AP, the MH may still have a good SNR. Under these circumstances the performance at the network layer can still be poor. There is one proposal enabling the MH to discover the utilization of APs at the network layer. That solution requires the AP to send advertisements announcing the load it is carrying.

The solutions with forwarding between APs during handover require the MH to signal with both APs requiring a multihoming functionality with one interface at the datalink layer.

The proposals for multihoming with MIP do not address multihoming with respect to performance issues regarding the connection to use.

In the proposals for global connectivity and hybrid networks MIP is used to manage mobility between ad hoc networks and often the MIP messages agent advertisement and agent solicitation are used to discover gateways. Solutions for both IPv4 and IPv6 are given. The discussed proposals adapt ad hoc networks to the functionality of subnetworks in wired IP networks and the following major challenges are addressed:

- Should the gateway use advertisements or should MHs request advertisements by solicitation?
- How to discover a CH being in the ad hoc network or in the wired IP network?
- Ways to address a default gateway.

With the discussed proposals for gateway connectivity, the same gateway is used for all traffic to and from an MH. In networks with multiple gateways it is beneficial to use the gateway having the best performance and that gateway may differ depending on the CH’s location and the load of gateways and routes. None of the solutions addresses the dynamic behavior of wireless links. The route decision is
based on the hop count and with similar hop counts the first received is selected. This means that the time factor plays a role when setting up the route using the same hop count but not otherwise. Considering that some routes may become more congested than others after this decision there is a need to continuously monitor the performance. It may be that a 3 hop route performs better than a 2 hop route.

The next chapter describes a proposed architecture connecting an ad hoc network to wired IP networks and the Internet.
Chapter 3. State-of-the-art in Mobile IP and Global Connectivity literature
Chapter 4. Software Solutions to Internet Connectivity in Mobile Ad Hoc Networks

1 This chapter is based on the publication:


Minor changes have been made to the publication to improve the presentation.
Chapter 4. Software Solutions to Internet Connectivity in Mobile Ad Hoc Networks
Software Solutions to Internet Connectivity in Mobile Ad Hoc Networks

In recent years wireless Internet access and wireless communications between peers have become the focus of intensive research efforts in various areas of information and communication technologies. Mobility aspects, software development and support for mobile users are currently of major interest within this research area. The Mobile IP protocol is deployed for mobility management of hosts moving between networks. Ad hoc routing is also of major importance for connectivity between communicating mobile hosts without backbone infrastructure. In this chapter we propose and describe an integrated connectivity solution and its software implementation between an ad hoc network running the Ad Hoc On-Demand Distance-Vector Protocol and a wired IP network where Mobile IP is used for mobility management.

4.1 Introduction

Current network and distributed systems are unthinkable without sophisticated software solutions that require non-stop evolution and improvement [55,56]. The IP protocol is the major network protocol used in modern computer networks. This protocol was developed for a wired network topology with stationary routers and hosts. The IP address is used to reach a node in a network (like a router and a host) by determining a route to the destination and identifying the host within the destination network. IP addresses are hierarchically organized with a major network part identifying the network, an optional subnet part identifying a subnetwork within the major network and a host part identifying the host within the network or the subnet. The hierarchical IP address allows routers to only look at the network part when finding a route to a destination. The only router looking at the host part of an IP address is the router(s) connected to the network where the destination host resides. To overcome the static network topology problems and to support mobility within IP networks, Mobile IP (MIP) [8] is proposed and partly deployed. MIP can be used for hosts connecting to a foreign network and still function as if they were connected to the home network. Because of the mobility aspects, wireless connections are becoming very popular and are considered the preferred way for a mobile host’s (MH) connectivity. Wireless technologies like 802.11a [2] and 802.11b [1] among other technologies are used instead of wired connections and do create a one hop connection to a wired network like Ethernet. To be able to connect to networks equipped with 802.11 access points (AP), the mobile host must be within radio communication range to the AP.

Another type of network becoming popular because of the wireless capabilities in computer communications is ad hoc networking [9]. In ad hoc networks there is no such fixed infrastructure as a wired backbone with routers. Instead, all MHs, which
are ad hoc hosts, usually both work as end user hosts and routers within the ad hoc network.

The goal with the work described in this chapter is to enable mobile users to communicate ad hoc when connecting without a backbone, and at the same time have connectivity to a wired IP network infrastructure (the home network or a foreign network) accessing the Internet. When outside communication range to an AP, an intermediate adjacent ad hoc host should be used if possible to reach the AP.

The proposal describes an approach connecting wired IP networks with ad hoc networks running the Ad Hoc On-Demand Distance-Vector Protocol (AODV) [11] for routing in the ad hoc network. Ad hoc hosts will have connectivity with hosts in the ad hoc network as well as in the Internet. For mobility of hosts between networks we use the MIP protocol. Our approach integrates ad hoc networks with wired IP networks and the Internet.

The contribution of the work proposed and described in this chapter is an integrated connectivity solution and its prototype software implementation. We propose a new functionality for gateways between IP networks and ad hoc networks for “global connectivity”, so that MHs can move between wired IP networks and ad hoc networks as well as between ad hoc networks while maintaining network connectivity. We also propose protocol changes so that the AODV protocol and the MIP protocol can function together.

This chapter is organized as follows. Section 4.2 describes our design and implementation. Related work is described in section 4.3. Section 4.4 summarizes the chapter.

4.2 Connecting IP Networks with Ad hoc Networks

4.2.1 Design considerations

For ad hoc networks to be integrated with IP networks, ad hoc networks should adapt to the network functionality within IP networks. Often ad hoc networks are seen as self-contained and are of limited size. Ad hoc networks are considered a complementary to IP networks in this chapter, where Internet connectivity can be extended into the ad hoc network, making ad hoc networks a part of the Internet. We assume that an ad hoc network uses a flat address space.

The location of a host in IP networks is identified by the network bits within the IP address (the major network bits and the subnet bits). When a packet arrives at the router connecting the network hosting the destination, and if the destination is connected to a network being a Local Area Network (LAN), the hierarchical IP address is converted to a MAC address using the ARP protocol [26] in IPv4 or the Neighbour Discovery Protocol [25] in IPv6, before the packet is sent in a frame in the last hop using the MAC address as the identifier of the destination. The MAC address represents a flat address space without information about the host’s locality within the LAN.

Considering the ad hoc network has a flat address space, it can be seen as a network (major network or subnetwork) within Internet. This will identify an ad hoc
network connected to the Internet by its own network number. There is, however, a major difference between a LAN and an ad hoc network. Hosts connected to a LAN are within the same broadcast domain, and are managed as one hop connections by the IP protocol. A packet broadcasted in the LAN will reach all hosts connected to the network. In the ad hoc network a broadcast sent by one ad hoc host may not reach all other hosts. A broadcast needs to be retransmitted by hosts in the network so that it will reach all hosts. A broadcast in the ad hoc network running the IP protocol uses the time to live (TTL) value to limit the spreading of a packet. A packet’s TTL when arriving at a router connected to a LAN requires a value of 1 to reach a host connected to the network. In the ad hoc network a TTL of 1 when forwarded on the ad hoc network will be discarded after the first hop. This behaviour needs to be managed by gateways connecting ad hoc networks with IP networks.

Instead of using the ARP or the Neighbour Discovery Protocol, the ad hoc routing protocol needs to be used in what is defined as the last hop in IP networks. The functionality in reactive ad hoc routing protocols maps well to the functionality in the ARP protocol and the Neighbour Discovery Protocol, with a request for the host in the last hop and a soft state table. The route request (RREQ) in reactive ad hoc routing protocols can be compared to the ARP request, and the soft state ad hoc routing table to the ARP table.

We assume all MHs in the ad hoc network to have an IP address and to be able to connect to hosts within the Internet. MHs should be able to move between ad hoc networks and still function as if they were connected to the home network. Peers in the ad hoc network should be able to communicate one-hop or multi-hop regardless of their network part in the IP addresses. This means that two hosts homed in different networks will be able to communicate peer-to-peer. MHs homed in an ad hoc network will have the network number given the ad hoc network, and MHs visiting the network will either use their home address as an identifier, an address given by a DHCP [20] server, or if IPv6 is used, stateless auto configuration [25] can be used.

MHs connected to the home ad hoc network will be reached through the IP network by IP routing to the gateway connecting to the ad hoc network. Then the ad hoc routing protocol is used to locate the destination in the ad hoc network. If however the gateway already has an active path to the destination, the packet will be forwarded without a prior RREQ creating a route.

When an MH in the ad hoc network addresses a correspondent host (CH), we use the proxy ARP approach to manage where to send a packet for a destination in the same network or in another. In the proxy ARP approach a router looks at the IP address requested, and if the router sees that the destination is in another LAN it responds with its own MAC address. All packets to the IP address will be sent to the router and the router will forward the packet to the destination. The source does not need by itself to investigate if the destination is within the same LAN or not, the router will support this. If the destination is within the same network it will itself respond.

In ad hoc networks we must be able to manage a multi-hop distance between hosts and the gateway, compared to the mechanisms in IP networks with one hop. Instead of the link layer address used in LANs to reach a gateway or a CH in the same network, the IP addresses of the gateway and the CH have to be used, since they may be multiple hops away. If the ad hoc routing protocol uses link layer addresses for
routing within the ad hoc network, link layer addresses can be used. But in our work the IP address within the ad hoc network is used to identify MHs and to make route decisions.

To manage mobility of MHs connected to an IP network and trying to connect to different ad hoc networks, MIP is used. In MIP, messages for advertisement, solicitation and registration are sent link local. For ad hoc networks the MIP messages must be able to travel multiple hops in the ad hoc network. To enhance the performance of our implementation, routes should be installed to the source of the MIP messages.

4.2.2 Design

In our design and implementation described below we make use of MIPv4, and the AODV routing protocol within the ad hoc network.

The design principles described in this section underlie our design considerations discussed in the previous section. We will describe:

- How MIPv4 messages are managed in the ad hoc network;
- How an MH decides which foreign agent (FA) to register with, if it knows several FAs;
- How a destination address from the ad hoc network is found being in the ad hoc network or in the Internet.

For an MH to be able to communicate ad hoc and to visit foreign ad hoc networks, the MH needs to run both the MIP and the AODV protocols. The same stands for a gateway connecting an ad hoc network to a wired IP network hosting the AP functionality. For the MIP messages specified to be sent link-local in a LAN, with a TTL value of 1, we have changed the value to indicate the size of the ad hoc network or the length in hops to an MH.

We have made the AODV protocol MIP-aware to recognize MIP messages so that ad hoc hosts (MH and gateway) can install routes based on the messages. Hosts forwarding agent advertisements will install a route to the FA. When an MH then registers with an FA by sending a registration request to the agent, a route will be available, without the need to do an explicit RREQ for the FA. The registration request creates a route for the registration reply, and the agent solicitation message creates a route for the unicast agent advertisement.

An MH discovers a path to a destination by sending an RREQ for the destination. If the destination is known to be outside the ad hoc network, the gateway replies with a proxy RREP (ProxyRREP) to the source. For a destination within the ad hoc network the gateway will function as an ad hoc host forwarding the RREQ.

For a gateway to know which hosts are in the ad hoc network the AODV protocol requires the knowledge of the visitor list in the FA, and all hosts homed in the ad hoc network have to have the same network number as the gateway connecting to the ad hoc network. These requirements are the same as for a LAN in IP networks managing mobility. When a route is requested for a destination with a network number different from the ad hoc network, the visitor list is in the gateway searched by the AODV process to see if the destination is available in the ad hoc network. If the destination is
not within the network the gateway sends a ProxyRREP to the source. If the destination is a foreign MH visiting the ad hoc network, normal AODV operations are used to discover the destination within the network. Figure 4.1 illustrates this process. A ProxyRREP is also sent if a destination homed in the ad hoc network is connected to a foreign network. The home agent (HA) functionality used is as specified in MIPv4. Packets coming to the gateway will be processed as shown in figure 4.2.

![Diagram](image)

**Figure 4.1.** The process in the gateway to manage AODV RREQ messages.

To manage several FAs covering an ad hoc network there is a need to synchronize the visitor information between the FAs. Without synchronization, a gateway may conclude that a destination is within the wired IP network and send a ProxyRREP to the source, while the destination is, in fact, within the ad hoc network but registered with another FA. The FAs synchronize their visitor list using the wired IP network to offload the ad hoc network, the information is synchronized when an entry is added or deleted from the visitor list in a gateway. In this way, all gateways will be able to see if a visiting host is within the ad hoc network even though it is not registered with the gateway receiving the RREQ. The HA binding cache does not have to be synchronized between the gateways since the gateway responding with a ProxyRREP will be the gateway acting as a HA for the MH connected to a foreign network. Other gateways will believe that the MH is in the ad hoc network.

When several FAs serve an ad hoc network, an MH must be able to choose the best FA to register with. In our approach, the FA closest to the MH will be used based on the roundtrip time (RTT) between the MH and the FA. For each agent advertisement received at an MH, an Internet Control Message Protocol (ICMP) [49] echo request message is sent to the FA to measure the RTT.
Chapter 4. Software Solutions to Internet Connectivity in Mobile Ad Hoc Networks

4.2.3 Implementation

The gateway as well as the MHs run the MIP software and the AODV software (see figure 4.3). The MIP software used is the HUT distribution [12] and the AODV software is a distribution from Uppsala University [13]. The operating system used is Linux 2.4 [57].

The MIP and the AODV software operate by modifying the forwarding table within the Linux kernel. In the kernel, the forwarding process forwards the packets using the information in the forwarding table.

The software in the gateway looks as in figure 4.3 with the extension of a shared memory between the MIP process and the AODV process, so that the AODV process will be informed of visiting MHs. If the gateway hosts an HA as well, the shared memory will also contain the hosts homed in the ad hoc network connected to a foreign network.
To measure the RTT we use the ICMP echo request sent to the FA that the MH knows about. To calculate the metric used to select an FA we use the Jacobson/Karels algorithm [19] for the retransmission timer in the TCP protocol (see formula 4.1).

\[
\text{Difference} = \text{SampleRTT} - \text{EstimatedRTT}. \tag{4.1}
\]

\[
\text{EstimatedRTT} = \text{EstimatedRTT} + (\delta \times \text{Difference}).
\]

\[
\text{Deviation} = \text{Deviation} + \delta(|\text{Difference}| - \text{Deviation}).
\]

\[
\text{Metric} = \mu \times \text{EstimatedRTT} + \phi \times \text{Deviation}.
\]

By using the deviation in the metric calculation the variance will affect the FA selected. To switch between FAs the difference between the metrics for two FAs must be bigger than a threshold.

In MIP, an MH receiving an agent advertisement assumes the previous link layer sender to be the FA, and uses this address as the default gateway. This must be modified to function in ad hoc networks; otherwise the MH will believe an intermediate host in the ad hoc network forwarding the agent advertisement to be the FA. We reserve the first care-of address field in the agent advertisement to the FA address within the ad hoc network (see figure 4.4). The MH can then discover the address of an FA being multiple hops away.

![Figure 4.4](image-url) The modified agent advertisement extension for the ad hoc network.

The agent solicitation message is also modified to function in the ad hoc network for the same reason as the agent advertisement. We have also added a sequence number to limit the broadcast of the solicitation and for the creation of a route to the sender. The solicitation message used in the standard MIP is the same as the router solicitation message. The extension in figure 4.5 shows the agent solicitation extension added to the router solicitation.

Since we create a route based on the registration request as well to enhance the performance, we use an extension field in the message to add the source address.
4.2.4 Evaluation

We have thus far created some basic measurements to evaluate the performance of a TCP flow and the RTT from a source in the IP network to an MH connected to a foreign ad hoc network being between 1 and 4 hops from the gateway (see figure 4.6).

The links in the wired IP networks are 100Mbps, and for the wireless connectivity Orinoco 11Mbps Silver cards are used. The agent advertisement is sent every 5 seconds and the ad hoc route timeout is 10 second. The results are shown in figure 4.7. All MHs are within 10 meters radius. We filter frames by the MAC-address to control the connectivity between MHs, and to create the ad hoc connections shown in figure 4.6. The performance degradation in figure 4.7 is due to the link layer protocol and it shows how an overheard flow will affect the delay and the throughput.
4.3 Related work

In [58], authors modify RIP for ad hoc networks and to work with MIP. This approach uses a proactive ad hoc routing protocol, which is not very efficient in ad hoc networks. In [45, 47, 59, 60], a proposal for connections between MIP and AODV is made. In [60] MIPv4 and AODV are connected so that MIP messages will be managed in the ad hoc network. The question of how to select between multiple FAs is not addressed, and an MH in the ad hoc network has to discover by itself if a destination is within the ad hoc network or not. If the gateway thinks it can receive the destination it replies with an FA-RREP. But before an MH can use the gateway, it first needs to conclude that the destination is not within the ad hoc network and this will delay the connection setup time. MIPv6 management with AODV is proposed in [59] using the neighbour discovery protocol. The same approach for destinations in the IP network is taken as in [60] and it is proposed that router advertisements should not be sent without router solicitation. However, in [45] measurements show that it is more efficient to use the normal MIP behaviour where advertisements are sent without solicitations. In [45] and [47] an approach to selecting between multiple FAs is described, the selection in made on the hop count between the FA and the MH. Hop
count may not be the best way to measure what FA to register with since network load is not considered. In [46] MIP and the ad hoc routing protocol DSR are addressed.

4.4 Chapter summary

In this chapter we propose and describe an integrated connectivity solution and its implementation connecting IP networks and ad hoc networks running the reactive AODV routing protocol, where MIP is used to manage mobility. The software supports the creation of areas covered by gateways connecting wireless MHs to the Internet. MHs will be able to communicate with peers in the ad hoc network and with hosts in the IP network. Our approach proposes a new way to locate a destination inside the ad hoc network or in the IP network, and the selection of a FA based on the RTT between the MH and the FA.

The solution supports the creation of applications requiring global connectivity. An application using a software socket (TCP or UDP) can rely on the networking software within the MHs and the gateway to find a route to a destination, in the ad hoc network as well as in the Internet. This avoids complexity within the applications.

The next chapter describes a proposed architecture extending MIP with multihoming functionality.
Chapter 5. Multihoming with Mobile IP:

This chapter is based on the publications:

C. Åhlund, A. Zaslavsky, Multihoming in Mobile IP, "IEEE International Conference on High Speed Networks and Multimedia Communications HSNMC'03, July 2003, Estoril, Portugal. Lecture Notes in Computer Science (LNCS), Springer-Verlag.


The two publications are related and are therefore combined and included in this chapter. The publication “Multihoming in Mobile IP” describes a developed prototype and the publication “Agent Selection Strategies in Wireless Networks with Multihomed Mobile IP” describes a simulation study.

Minor changes have been made to the publications to improve the presentation.
Chapter 5. Multihoming with Mobile IP
Multihoming with Mobile IP

Mobile IP is the standard for mobility management in IP networks. With today’s emerging possibilities within wireless broadband communication, mobility within networks will increase. New applications and protocols will be created and Mobile IP is important to this development, since Mobile IP support is needed to allow mobile hosts to move between networks with maintained connectivity. This chapter describes multihomed MIP enabling mobile hosts to register multiple care-of addresses at the home agent as well as the correspondent hosts, to enhance the performance of wireless network connectivity. Correspondent hosts and the home agent can select different care-of addresses to achieve load-balancing and a more reliable connectivity. A prototype and a simulation study are described.

5.1 Introduction

In the future wireless local area networks (WLAN), connectivity to access points (AP) by different technologies and different providers will be a reality. Technologies like 802.11 [1], 802.16 [6] and HiperLAN [61] will support wireless network connectivity to wired network infrastructures for reaching the Internet and for other types of services. WLAN-technologies are becoming efficient enough to support network capabilities for applications running in desktop computers.

However, with the use of WLANs, new challenges arise and mobile hosts (MH) will face multiple APs with possibly different capabilities and utilization.

The work described in this chapter is based on the 802.11b technology. In 802.11b, there are two different Basic Service Sets (BSS) for connectivity: infrastructure BSS and independent BSS. In the infrastructure mode the association with an AP is based on link-layer mechanisms using the signal quality. The selection is invisible to upper layer protocols and one association at a time is possible. In independent BSS a network interface can communicate with all others within communication range without association. This is the mode used in ad hoc networks [9].

The selection of AP should also be available for higher level protocols, the applications and the users. For example the signal quality might be somewhat better to one AP but throughput better at another. Here it would be reasonable to use the AP with the best throughput. Since wireless connections are prone to errors and by using multiple simultaneous connections to APs, a more reliable connectivity is achieved.

In the largest study so far [62], a university campus equipped with WLANs where 476 APs are spread over 161 buildings divided into 81 subnets and 5,500 students and 1,215 professors are equipped with laptops. The study shows that 17% of the sessions involve roaming and that 40% of it was between different subnets, causing the IP traffic to fail. MHs sometimes performed frequent handovers between APs while being in the same place.

For mobility and maintained network connectivity for an MH that moves between APs, the Extended Service Set (ESS) can be used. The ESS manages handover in the datalink layer and so is restricted to the same Local Area Network (LAN) or Virtual
LAN (VLAN). An MH doing handover between APs in different networks (ESS) will break flows.

To manage handover between networks without disrupting flows, the Mobile IP (MIP) [8] has been proposed and partly deployed. For an MH connected to the home network, the IP will operate normally. If the MH disconnects from the home network and connects to a foreign network, the MIP will manage network mobility which will be transparent for the protocol layers above the network layer and to the user of the MH. There are two versions of MIP: MIPv4 [23] and MIPv6 [24].

The study [62] shows the MIP requirements and the potential to associate with multiple APs simultaneously to avoid breaking and disrupting sessions. Wireless connections are prone to errors and by using multiple simultaneous connections to APs, a more reliable connectivity can be achieved.

To enhance performance and reliability in network connectivity, host-based multihoming can be used. A host can be multihomed by using two interfaces configured with different IP addresses, or by using two IP addresses for the same network interface. Host-based multihoming is usually managed by a Domain Name Server (DNS) [21]. For a multihomed host the name-to-address binding binds multiple IP addresses to a single host name. Either the DNS can return the IP addresses in a round robin fashion upon name resolution requests or alternatively multiple IP addresses can be returned and the host selects which address to use.

The work in this chapter describes an approach to enhance the performance of network connectivity to MHs connecting to WLANs. The MIP is extended to support multihomed connectivity where multihoming is managed by MIP. This will enable the AP selection on other criteria than just the signal quality: Traffic to and from an MH can be sent using multiple APs. A prototype developed is also described.

In Section 5.2 multihoming with MIP is described. A simulation study is presented in section 5.3. Section 5.4 describes related work and section 5.5 summarizes the chapter.

5.2 Multihomed Mobile IP

Multihomed MIP enhances the performance and reliability for MHs connecting to WLANs. Wireless connections are prone to errors and changing conditions which must be considered to enable applications for desktop computers.

The multihoming is managed by the MIP and hidden from the IP routing, keeping IP routing protocols like the Routing Information Protocol (RIP) [63] and Open Shortest Path First (OSPF) [43] unaware. For a sender, multihomed MIP can be considered an any-cast approach [64] where a sender relies on the network protocol to use the best available destination for the packets. The destination will be one of many possible care-of addresses used by an MH. In IPv6, an any-cast address is used to reach the best available destination (server) among multiple destinations supporting the service required. The approach in this chapter for a sender to any-cast address an MH, is that the MH’s home address be used to locate the best care-of address. The difference from the any-cast approach in IPv6 is that it is address-based instead of server-based and the destination will be the same host.
Chapter 5. Multihoming with Mobile IP

The MH keeps a list of APs it receives agent advertisements from and registers the care-of addresses at the HA (and the CH if MIPv6 route optimization is used) for the networks supporting the best connectivity. To evaluate the connectivity, the MH monitors the deviation in arrival times between advertisements and makes calculations based on this information (see formula 5.1). The resulting metric is used to describe the MH’s connectivity to foreign networks. A small metric indicates that agent advertisements sent at discrete time intervals arrive without collisions and without being delayed by the FA. This also indicates available bandwidth as well as the FA’s capability to relay traffic for the MH. Among the care-of addresses registered at the HA, the FA with the smallest metric will be installed as the default gateway in the MH.

The selection of which care-of address to use for an MH is based on the delay between a CH or the HA and the MH, where the delay includes wireless links. In IP routing with protocols like RIP and OSPF a wireless last hop link is not considered in the route calculation. A hop count of 1 is used in the RIP protocol, and a static link cost in OSPF based on the link (usually Ethernet) connecting the APs. In multihomed MIP, IP routing is used to the care-of address selected but the selection of which care-of address to use is managed by MIP. The HA makes its own selection and the CH does the same if route optimization is used.

Before informing the HA and CHs about the current location of the MH, the MH must decide which foreign networks to register with. An MH receiving advertisements from foreign networks will monitor the available networks and calculate the deviation in arrival times. This is recorded for each sender and the networks with the smallest metric calculated from the deviation are registered at the HA. An MH is configured with the maximum number of care-of-addresses to register.

Since the MH may register multiple associations with foreign networks, the HA can have multiple bindings for an MH’s home address. Based on the round trip time (RTT) between the HA and the MH, one of the care-of addresses will be installed as the tunnel end-point to the MH. The measuring of RTTs is based on the registration messages sent between the MH and the HA as well as binding updates between the MH and its CHs using route optimization according to the MIPv6 proposal. The metrics for the selection of care-of address made by the HA and CH (if route optimization is managed by the MH) is based on the Jacobson/Karels algorithm [19] (see formula 4.1). A small value is preferred.

A CH sending packets to an MH without route optimization will send them to the MH’s home network, where the HA will make the selection of which care-of address to use to forward the packets. With route optimization the CH will send the packets to the MH without using the HA and based on the MIP version used, will either tunnel the packets (MIPv4) or use the routing header (MIPv6). Route optimization is managed by the HA (MIPv4) or the MH (MIPv6), sending the CH information of

\[
\text{Sample} = \text{CurrentArrivalTime} - \text{LastArrivalTime}. \quad (5.1)
\]

\[
\text{Mean} = \text{Sample} \times \delta + \text{Mean} \times (1 - \delta).
\]

\[
\text{Metric} = (\text{Sample} - \text{Mean})^2 \times \mu + \text{Metric} \times (1 - \mu).
\]
which care-of address to use. In multihomed MIP multiple care-of addresses may be sent.

The selection by the CH for which care-of address to use can be based on:
- Longest prefix match;
- The first address received;
- Delay.

Considering the proposal made to MIPv4 for route optimization without explicitly measuring the RTT (e.g. sending ICMP [49] echo requests) between the CH and the MH, longest prefix match and arrival order are the options available. With the proposal for route optimization made to MIPv6 the care-of address selection made by the CH is managed the same way as by the MH. However, instead of the MH sending the RTT time delayed by one registration request, the CH itself measures the RTT by monitoring the time between sent binding refresh update messages and binding update response.

The choice of care-of address is based on individual selections by the HA, the CH and the MH for packets sent by them. In a scenario where an MH has registered three care-of addresses and there are two CHs, one using the HA to communicate with the MH and the other using route optimization, three different APs may be used: one by the HA, another by the CH using route optimization, and the third by the MH to send packets (see figure 5.1).

![Figure 5.1. A multihomed connectivity scenario where the HA, CH and MH make their own selections of which care-of address to use.](image)

To avoid rapid changes resulting in flapping of the care-of addresses and the default gateway because of metrics close in value, a new care-of address or gateway is only chosen if they differ beyond a threshold.
5.2.1 The prototype

This section describes the prototype and the extensions made to MIP. MIPv4 is used as the framework and the use of an FA is assumed. Route optimization is handled by MIPv4 and MIPv6 in different ways: in MIPv4 route optimization messages are sent by the HA to CHs; in MIPv6 the MH informs CHs about its address. To add this behavior to the prototype, both types of route optimizations are considered.

To register a care-of address at the HA, a registration request is sent, and to enable the HA to distinguish between a non-multihomed and a multihomed registration, an N-flag is added to the registration request (see figure 5.2).

A HA receiving the registration request with an N-flag will keep the existing bindings for the MH. One of the registered care-of addresses will be used to forward packets to the MH. For the HA to be able to make the selection, the RTT to the MH through the different care-of addresses is measured. The MH monitors the time between registration requests and registration replies and calculates the RTT. The RTT is added as an extension in the next registration request. The HA will maintain all registrations for an MH and based on the metrics calculated by formula 4.1 will install a tunnel into the forwarding table for the care-of address having the smallest metric.

With a care-of address advertised by an FA, the MH is not allowed to use the Address Resolution Protocol (ARP) [26]. This will confuse other hosts connected to the network and may cause problems when the MH disconnects and moves to another network. To avoid this in MIP, the MH monitors the MAC address in the frame containing the agent advertisement, and installs the binding between the FA’s MAC address and the IP address in the ARP table, for the FA registered with. When a packet is sent using the default gateway, an entry in the ARP table will already be available and no ARP request is needed. In multihomed MIP, the MH will maintain multiple registrations with different FAs as well as keep control of available FAs not registered with. All IP addresses for the FAs registered with are installed in the forwarding table, and the bindings between the IP and the MAC addresses installed in the ARP table.

We propose two changes for route optimization, considering both the proposal for MIPv4 as well as how MIPv6 manages route optimization. For MIPv4 route optimization, the binding update sent from the HA to a CH is shown in figure 5.3. For MIP multihoming, multiple packets must be sent to inform the CH of multiple care-of addresses. The binding update is extended with an N-flag to signal multihomed...
binding updates. The first binding update clears the N-flag to erase the binding cache in the CH from possibly stale entries. The rest of the binding updates have the N-flag set. The CH will install a tunnel to the MH in its forwarding table based on the information in the binding updates. The decision of which care-of address to install as the tunnel end-point in the forwarding table is based on the selections described earlier. When an MH performs a handover between networks and changes from one care-of address to another, the binding update needs to express both the new and old care-of addresses, so that the HA and the CH knows which binding to update. In the case of a single-home binding, only the new care-of address is included.

Figure 5.3. The modified binding update message with the added N-flag and the optional care-of address.

When a binding is about to expire, CH sends a binding request message to the HA. The binding request must include the care-of address for the binding (see figure 5.4) so that the HA knows which binding to respond with. Without the care-of address included, binding updates are requested for all care-of addresses. An HA will respond with care-of addresses for all available bindings for an MH.

Figure 5.4. The modified binding request message with the optional care-of address that the request is sent for.

If a binding request is sent for one care-of address and multiple bindings are maintained, the binding update will have the N-flag set. This will inform the CH whether other bindings are maintained as well. If the care-of address requested in the binding request message has no binding in the HA, the HA will respond with a binding update with a lifetime set to 0.

To manage route optimization where an MN sends binding updates (instead of the HA) to the CHs as in MIPv6, the same messages is used as described above (see figures 5.3 and 5.4). The advantage of this method for route optimization is that RTT can be measured by the CH by monitoring the departure time of binding requests and the arrival of binding updates.
5.3 A simulation study

The evaluation uses the GlomoSim [65] network simulator. The topology shown in figure 5.5 is used in the simulations. The node MH2, equipped with two wireless interfaces, is connected via different channels to a network where it can reach FAs in different subnetworks. MH1 can only see FA1 and MH3 can only see FA2. MH1, MH2 and MH3 are out of communication range from each other.

Figure 5.5 MH2 reaches two FAs in different subnetworks and communicates with a CH in its home network. MH1 and MH3 are used to add additional flows to the FAs.

The graph in figure 5.6 plot results from a simulation where MH2 has selected FA1 as its gateway based on the signal-to-noise ratio (SNR). The SNR have a value of 22dB for all plotted curves in the graph.

![Graph showing throughput of 1 Mbps flow sent from CH to MH2 with different sizes of traffic between MH1 and FA1.](image-url)

Figure 5.6. Throughput of the 1 Mbps flow sent from the CH to MH2 with different sizes of traffic between MH1 and FA1.
Chapter 5. Multihoming with Mobile IP

All traffic generated in the simulations is constant bit rate traffic. With added and increasing throughput of the traffic flow between MH1 and FA1, the SNR value at MH2 remains the same. The throughput from the CH to MH2 for different sizes of added traffic between MH1 and FA1 is shown. A traffic flow of 1 Mbps is generated in both directions between the CH and MH2. The size of the inserted flow between MH1 and FA1 is used to name the curves in the graph. The yellow curve shows the throughput from the CH to MH2 with the MH2 and FA1 communicating 1.5 Mbps in both directions. The violet curve shows the same for 2 Mbps, the blue curve for 4 Mbps and the red curve for 6 Mbps.

As shown in figure 5.6 the throughput for traffic from the CH to MH2 varies depending on the traffic between MH1 and FA1 without being noticed by looking at the SNR.

Figure 5.7 plots the results from a simulation using different FA selection strategies. The plotted “SNR” curve (red dotted) shows the throughput when an FA is selected based on the SNR. The curve named “Dyn” (in blue) shows the throughput when the FA is selected based on the deviation of agent advertisements and RTT measurements.

Flows are added to congest the wireless link alternately between FA1 and FA2. Traffic between the MH1 and FA1 is generated the time periods 1 to 15 seconds and 30 to 45 seconds. The traffic between the MH3 and FA2 is generated the time periods 15 to 30 seconds and 45 to 60 seconds.

**Figure 5.7** Throughput between the CH and MH2 with different FA selection algorithms.
The plotted curve named SNR shows the throughput from the CH to MH2 when FA1 is selected based on the SNR. MH2 do not discover the traffic generated between FA1 and MH2 by looking at the SNR and continues to use FA1. In the time periods 1 to 15 seconds and 30 to 45 seconds the link used by MH2 to connect to FA1 is congested by traffic between MH1 and FA1. The throughput from the CH to MH2 is therefore low. The transients at times 15 and 45 seconds are explained by the buffering that take place in FA1 because of the congested link. When the link becomes free, buffered packets will be sent rendering in a temporary high throughput.

The solid blue plotted curve in figure 5.7 named “Dyn” shows the throughput when formula 5.1 is used to evaluate the connection to an FA for the selection of default gateway in MHs, based on agent advertisements sent by FA1 and FA2 once every second. Formula 5.2 calculates the metric used by the HA to select care-of address, based on the RTT measurements. The blue curve plots a higher throughput than the red dotted curve. The reason is that handovers take place when a wireless link is congested. The low throughput shown at times 15, 30 and 45 seconds are explained by the time it takes to react to bad metrics when a recently good connection becomes congested. As shown in figure 5.7, it is a correlation between the dept and the width of these drops in throughput. The reason is that if traffic starts congesting the link just after an advertisements is sent it will take longer time to react on a congested link, compared to if traffic starts to congest the link prior to the advertisement.

### 5.4 Related work

In MIPv4, an option for simultaneous bindings is proposed for sending packets to multiple care-of addresses for an MH. Packets will be duplicated at the HA and one copy sent to each registered care-of address, so that packets can be received through multiple APs. This option was proposed to decrease the number of dropouts of packets during handover, and for an MH with bad connections to APs to receive the same packet through several APs, with an increased probability of success. The solution does not enable the network layer to decide which connection to use and wastes resources in the WLAN.

In the current specification of MIPv6, all traffic uses the same care-of address. This prevents the dynamics of the MIP from fully utilizing the dynamics in WLANs and should be altered.

Dahlberg and Jung [66] have proposed an approach to multihoming for survivability managed at the datalink layer and based on radio signaling. However, this approach restricts the selection of APs to the datalink layer and is not available to higher levels.

Hsieh and Sivakumar [67] have proposed a transport layer protocol for striping data between multiple links to achieve bandwidth aggregation. The work presented in this paper instead aims to evaluate multiple connections and how to use the best available connection(s) to forward packets.

Another transport layer solution is presented by Stewarts and Metz [68] for multihomed hosts. Here the sender selects one of the host’s IP addresses as the destination address for the packets. If the IP address becomes unavailable due to
network failure, the protocol will switch to another IP address for the same destination host with maintained connectivity at the transport layer. The approach however does not address delays considering a wireless last-hop link.

Aida, et.al [88] show the correlation of signal-to-noise-ratio and throughput. However the approach does not cover multiple MHs communicating using e.g 802.11 (using clear-to-send, request-to-send and NAV times to avoid collisions). In this case the SNR will not be sufficient, since SNR is not affected. Traffic measurements are instead required.

Nagami et.al [89] have presented a proposal for multihoming with MIPv6. Our proposal is intended for both MIP versions. We also propose a care-of address selection algorithm and evaluate its performance. While no selection algorithm is presented by Nagami.

5.5 Chapter summary

The work described extends the MIP to manage multiple simultaneous connections with foreign networks. Based on the registered care-of addresses, multiple FAs can be used for packets to and from an MH. The approach also prevents MHs from flapping between foreign networks due to the fact that an MH has similar quality of connectivity to multiple networks.

Enhanced throughput and more reliable connections are achieved. The current prototype is based on MIPv4 but will in the next phase be deployed on MIPv6 as well.

A study will be performed on the impact of the delay between the measure of the RTT and the time the HA receives the information, since the RTT is sent one registration request later than when it was measured.

The association should be made available to higher protocol layers and provided through an Application Programmer’s Interface (API). Future work will look into possible solutions to achieve this.

The next chapter describes work combining the approaches given in the previous chapter and in this chapter.
Chapter 6. Extending Global IP Connectivity for Ad Hoc Networks

3 This chapter is based on the publication:


This chapter has overlapping parts with chapter 4 and 5 and is an extended publication based on the licentiate thesis:

C. Åhlund, Supporting Global Connectivity in Heterogeneous IP Networks, November 2002

A simulation study presented in the Journal publication is excluded from this chapter. More comprehensive simulation studies are instead included in chapter 7 and chapter 8.

Minor changes have been made to the publication to improve the presentation.
Chapter 6. Extending Global IP Connectivity for Ad Hoc Networks

Extending Global IP Connectivity for Ad Hoc Networks

Ad hoc networks have thus far been regarded as stand-alone networks without assumed connectivity to wired IP networks and the Internet. With wireless broadband communications and portable devices with appropriate CPU, memory and battery performance, ad hoc connectivity will become more feasible and demand for global connectivity through ad hoc networking is likely to rapidly grow. In this chapter we propose an algorithm and describe a developed prototype for connectivity between an ad hoc network running the Ad Hoc On-Demand Distance-Vector Protocol and a wired IP network where Mobile IP is used for mobility management.

6.1 Introduction

Rapid developments in wireless computer communication technology have made wireless communication appealing for connectivity to wired Internet Protocol (IP) networks and the Internet, as well as for communicating peer-to-peer. The wireless local area network (WLAN) standard referred to as 802.11b [1] is widely deployed and theoretically supports the throughput of up to 11Mbps. Another standard, 802.11a [2] supports up to 54Mbps. The bandwidth will scale for many applications used in desktop computers and will gradually increase.

Places with high concentrations of people and businesses (hotels, airports, shopping centers, companies, etc), are being covered by wireless high bandwidth network infrastructure that includes access points (AP). Areas supporting high bandwidth locally are called hot spot areas, and are based on technologies like 802.11. With the advent of high bandwidth wireless network access, new demands for network protocol support arise. Today, hot spot areas provide wireless network access by connecting to one AP and have somewhat limited support for seamless roaming while staying connected to one AP only.

To enable networking software to fully adopt the possibilities with wireless network access, new functionalities need to be added to mobile hosts (MH) and to the wireless access network.

Another promising new networking technology brought about by the wireless capabilities is ad hoc networking [9]. Ad hoc networking enables MHs to create a multi-hop network on their own without a backbone infrastructure.

Ad hoc networks consist of MHs communicating wirelessly without a wired backbone infrastructure of connected APs. Every MH within the network is both a host and a router for other MH flows. There are two main types of ad hoc routing protocols: proactive and reactive. The proactive protocols maintain routing tables describing the topology in the same way as routing protocols like Routing Information Protocols (RIP) [63] and Open Shortest Path First (OSPF) [43] in wired IP networks. Routing tables are maintained for the topology, continuously exchanging route information between the MHs, regardless of whether user data is sent or not. Reactive ad hoc routing protocols create routes on request by a source to send data.
Chapter 6. Extending Global IP Connectivity for Ad Hoc Networks

Examples of proactive routing protocols include the Destination-Sequenced Distance Vector Protocol (DSDV) [28] and Cluster Switch Gateway Routing (CSGR) [29]. DSDV is a distance-vector routing protocol that periodically and event-driven transmits route updates, and uses sequence numbers to avoid routing loops. MHs in a DSDV network are managed as a flat address space without routing hierarchy. CSGR is based on DSDV and creates a hierarchical topology using clusters, where the cluster-heads create a wireless backbone.

Dynamic Source Routing (DSR) [30], Location-Aided Routing (LAR) [69] and Ad Hoc On-Demand Distance Vector Protocol (AODV) [11] are examples of reactive routing protocols. The AODV protocol is a development from DSDV based on symmetric links, managing both unicast and multicast routing. When a sender requests to send packets, a route request (RREQ) for the destination is broadcast in the network. When an MH rebroadcasts an RREQ it creates a reverse route to the sender of the RREQ. The destination receiving an RREQ, will reply with a RREP to the sender. The RREP is sent using the reverse route created by the RREQ. Upon receipt of the RREP at intermediate MHs and the source of the RREQ, a forward route is created and packets can be sent as illustrated in figure 6.1.

![Figure 6.1. An ad hoc network running AODV.](image)

If an intermediate MH receiving an RREQ knows a route to the requested destination, it can respond with an RREP if the sender allows it. This is usually allowed for User Datagram Protocol (UDP) flows where there might be no reverse traffic, but for Transmission Control Protocol (TCP) flows where at least acknowledgements (ACK) will be returned, the RREQ should be sent all the way to the destination. If an intermediate MH responds, the destination may not have a route for its ACKs and have to invoke a new RREQ. Another option for this is to use gratuitous RREP.

When a route breaks because of a node movement or a bad communication link, the upstream MH noticing the break sends a route error (RERR) to the source. The RERR traversing MHs to the source will inform about the lost connection and all routes using the link will be erased. Created routes are soft-state and expire when not...
used for some time. To maintain neighbour relationships with nearby MHs, “hello” packets can be used.

In ad hoc networks where the topology may change frequently, reactive routing protocols have been shown to scale better than proactive routing protocols [70].

To manage mobility for an MH connecting to IP networks, where both applications as well as the user are unaware of the network mobility, the Mobile Internet Protocol (MIP) [8] is deployed.

MIP is based on a home agent (HA), a foreign agent (FA) in MIPv4 [23] and the MH. An MH connected to the home network will operate according to normal IP network operations, without using the MIP. When an MH connects to a foreign network it will register its new location with the HA based on the address used by the MH in the foreign network. This address is called a care-of address or co-located care-of address depending on how addresses to visiting MHs are managed in the foreign network, and depending on the MIP version. In MIPv4 the address given a visiting MH can be supported by an FA (care-of address) or the dynamic host configuration protocol (DHCP) [20] server (co-located care-of address). In MIPv6 [24] there are no FAs so a co-located care-of address is always used and managed by IPv6. In IPv6 there are two options for an MH to receive an address: first, via stateless auto-configuration based on the neighbor discovery protocol (NDP) [25], and second, via statefull process using a DHCP server. In this paper, the care-of address is also used to describe a co-located care-of address. It will be clear from the context which is being referred.

The registration sent to the HA by an MH connecting to a foreign network will create a binding in the HA, between the home address and the care-of address. If an FA is used, the MH sends the registration through the FA. The FA registers the MH in its visitor list and forwards the registration to the HA. In MIPv4 the registration is named a registration request and a registration reply is returned by the HA. In MIPv6 it is named a binding update and a binding acknowledgement is returned in response.

When packets for the MH are received at the home network, the HA will forward the packets to the care-of address using tunneling. A tunnel encapsulates the received packet for an MH as a payload in a new packet with an IP header having the care-of address as the destination and the HA as the source. When the packet arrives at the FA or the MH (if no FA is used) at the foreign network, the packet will be decapsulated by the networking software. If an FA is used the packets will be sent in a frame, the last hop with the MAC address stored in the visitor list for the MH. The MAC address is registered when the MH makes the registration by looking at the source address in the received frame. If the MH itself manages the decapsulation the outer packet header is stripped off before handed to upper layers.

The interception of packets by the HA for an MH, is based on the Address Resolution Protocol (ARP) [26] in IPv4 or the NDP in IPv6. A router connected to a network receiving a packet for a destination in the network or a source in the same network as the destination, will translate the IP address to a MAC address. The Media Access Control (MAC) address is used to send a frame containing the packet the last hop to the destination. The ARP and the NDP are used to send a request for a MAC address, containing the IP address. The host configured with the IP address will respond with its MAC address. When an MH has registered a care-of address at the
Chapter 6. Extending Global IP Connectivity for Ad Hoc Networks

HA, the HA will respond to the ARP and NDP requesting a MAC address for the MH’s IP address. The returned MAC address will be the MAC address of the HA.

All hosts as well as routers connected to a network maintain caches for the binding between IP and MAC addresses. To prevent caches from keeping obsolete bindings when an MH registers at a foreign network, the HA sends a gratuitous ARP (IPv4) or a gratuitous Neighbour Advertisement message (IPv6), to update these caches with the binding between the MH’s IP address and HA’s MAC address.

In MIPv4, an MH visiting a foreign network sends packets to a correspondent host (CH) using the MH’s home address as the source and the CH’s address as the destination. However, ingress filtering may cause problems that require packets to be tunneled (reverse tunneling) to the HA first, and then sent from the HA to the CH.

In MIPv6 a packet sent to the CH will use the MH’s care-of address as the source, and the home address will be added in the home address destination option. Since the addresses are topology-correct ingress filtering is avoided. The CH receiving the packet will put the address in the destination option as the source address before handing it to the transport layer.

The routing created by MIP is referred to as triangular routing. Here packets from a CH are sent to the HA. The HA tunnels them to the MH, and the MH sends packets directly from its current location to the CH, making a triangle (if reverse tunneling is also used it is named quadrilateral routing). To optimize routing between the MH and a CH, route optimization [71] is used. A CH sending packets to an MH is informed by the HA (MIPv4) or the MH (MIPv6) about the care-of address of the MH. When receiving a binding update with the care-of address, a CH can send packets directly to the MH using the care-of address as the destination address. In MIPv4, route optimization is seldom used because of security problems, and because the CH must be MIP-aware to manage the tunneling to the MH. Currently there is no active proposal for route optimization in the Internet Engineering Task Force. In MIPv6, support for route optimization is built into the IPv6. The CH will use the routing header in IPv6, where the destination of the packet is a care-of address and the address in the routing header is the MH’s home address.

The HA and FA can be implemented in hosts connected to the MH’s home network and the foreign network or in the routers connected to the networks.

In wireless access networks APs are used for connectivity to wired networks. The AP bridges wired and wireless media. In a wireless access network based on 802.11[1], the AP bridges wireless media and usually the Ethernet. Two types of Basic Service Sets (BSS) are defined in 802.11: the infrastructure BSS and the independent BSS. In the infrastructure BSS an MH associates with one AP, and sends all packets to other hosts through the AP even though peers are within communication range. The independent BSS connects MHs without an AP and MHs communicate without a wired backbone. This is the BSS used in ad hoc networks.

To support mobility between APs in infrastructure BSS, the wired backbone is used by APs to manage associations with MHs, enabling handover without connectivity disruption. The service is named the extended service set (ESS), see figure 6.2. The handover is managed at the data link layer, so it is limited to the same local area network (LAN) or virtual LAN (VLAN). MIP must be used to manage handover between different ESSs with maintained connectivity.
The size of a LAN/VLAN is restricted to an autonomous system (AS). Restrictions include the number of users connecting to the network and the traffic traversing it, to avoid congestion. An LAN/VLAN is a single broadcast domain, and broadcasts sent by one host are seen by all other hosts, affecting the performance. Traffic based on broadcast in IPv4 is for example the ARP protocol. In IPv6 there are no broadcasts. The functionality in IPv4 requiring broadcast uses multicast in IPv6 to enhance the performance. However, the all-nodes multicast address affects all hosts. Another protocol restricting the size of VLANs is called the spanning tree protocol because of its settling time, and is used to avoid loops in a switched infrastructure.

Global connectivity for Internet users is achieved by connectivity between heterogeneous networks. IP was created to connect different types of networks like Ethernet, Frame-Relay, and Token Ring etc. Future global connectivity will also include networks based on different wireless technologies like Bluetooth [72], 802.11 etc. Both 802.11 and Bluetooth are technologies used in ad hoc networks which have thus far been regarded as stand-alone networks without connectivity to wired IP networks and the Internet. With wireless broadband communications and portable devices with good CPU, memory, and battery performance, ad hoc connectivity will become more feasible and demand for global connectivity through ad hoc networking is likely to increase.

Multihoming is a term used for a host identified by two or more IP addresses, and a network identified by two or more network numbers. A multihomed host can use multiple network interfaces configured with different IP addresses, possibly connected to different physical networks. In IPv6 and most of today’s operating systems implementing IPv4, one network interface can have multiple IP addresses enabling multihoming with a single interface. In this case one physical network has multiple network numbers.

Network multihoming is for example used to connect an AS with multiple Internet service providers (ISP), managed by the Border Gateway Protocol (BGP) [73]. The reasons for using multihoming are improved performance and increased reliability.

Host-based multihoming is usually managed by a Domain Name Service (DNS) [74]. For a multihomed host the name-to-address binding binds multiple IP addresses
to a single host name. The DNS can return the IP addresses in a round robin fashion upon name resolution requests. Another option is that multiple IP addresses are returned and the host selects which address to use. The contributions of the work proposed and described in this paper include:

- A gateway is proposed, connecting wired IP networks and ad hoc networks, where the ad hoc network uses a reactive ad hoc routing protocol. Communication can be ad hoc with a peer or by associating with an AP to use the wired network. Multi-hop ad hoc connections can be used for connectivity to an AP. MHs moving between ad hoc networks are managed by the MIP. This will enable the (sub-) network architecture to be extended to manage multi hop ad hoc networks as (sub-) networks;
- The MIP is extended to enable MHs to register with multiple gateways simultaneously. This enhances the network connectivity by enabling the MH, the HA and the CH to evaluate and select the best connection for their communication;
- A prototype has been created to verify and evaluate the gateway. The prototype uses MIPv4 and AODV for ad hoc networks.

In section 6.2 we describe our design and prototype, and in section 6.3 related work. Section 6.4 concludes the paper and outlines future work.

6.2 Connecting IP Networks and Ad hoc Networks

The connectivity between wired IP networks and ad hoc networks integrates ad hoc networks into the global connectivity provided by the IP. Because of the capabilities of this connectivity, a more dynamic network infrastructure is created. Mobility between ad hoc networks is managed by MIP extended with multihoming functionality. This will enable the MIP to enhance its performance and reliability by using multiple network connections for MHs. We define this as multihomed MIP.

6.2.1 The Design Considerations

Global connectivity is achieved by the layering in the TCP/IP stack. In the physical layer, different physical equipment may be used, and in the datalink layer, different protocols can be used (e.g. Ethernet, Token Ring, Frame Relay). The network layer manages different datalink layer protocols and enables connectivity between them. The layers above the network layer (transport and application layer) are unaware of the differences in networking technologies, thus enabling global connectivity. When connecting ad hoc networks with wired IP networks (see figure 6.3), the differences between the two types of networks will be considered in the network layer.

MIP is used to manage MHs disconnecting from the home ad hoc network and connecting to foreign networks. MIP is extended to operate in ad hoc networks using a reactive routing protocol, where MIP messages are managed multiple hops instead of one hop as in the MIP specification. This enables MHs multiple hops in the ad hoc
network from a gateway, to register. The AODV protocol is modified to enable redistribution of MIP information, to create ad hoc routes based on MIP messages. This will enhance performance.

![Figure 6.3. Global connectivity with a multihomed MH.](image)

**Global Connectivity**

For ad hoc networks to be integrated with wired IP networks, ad hoc networks should adapt to the network functionality within wired IP networks. Often, ad hoc networks are seen as self-contained. Ad hoc networks are integrated with wired to IP networks in this article, where Internet connectivity can be extended into the ad hoc network, making ad hoc networks a part of the Internet. We assume that an ad hoc network uses a flat address space. Our work is based on the assumption that a wired IP network infrastructure is used to connect multiple distant ad hoc networks. The ad hoc networks considered as (sub-) networks in this work are of limited size with a network diameter of a few hops.

Routing and network management in wired IP networks are based on a relatively static topology, while routing and management in the ad hoc network need to manage a more dynamic and frequently changing topology.

In computer networking, routing protocols need to efficiently limit the routing information sent, in order to reserve network resources for user data. When two networks with different routing protocols are interconnected, the router connecting the two needs to run both protocols. To enable the information from one routing protocol to be available in the other, redistribution of the information is required. For example if a RIP network is connected to an OSPF network and the routes learned by RIP are to be available in OSPF, redistribution from RIP to OSPF is required.
Ad hoc networks consisting of MHs with scarce resources should not have to manage routing information from the wired IP network, since this would severely impair performance in the ad hoc network. Also, inserting frequently changing ad hoc routing information into the wired IP network would increase the administrative overhead.

To manage the differences in the two types of networks and to achieve an efficient connectivity, the ad hoc network is here a (sub-) network with its own network number. By identifying the ad hoc network with a network number, the only routing information sent in the wired IP network is this number, and there is no internal ad hoc routing information. In ad hoc networks, the only information needed by MHs to access the wired IP network and the Internet is which gateway is to be used.

The location of a host in IP networks is identified by the network bits within the IP address (the major network bits and the subnet bits). When a packet arrives at the router connecting the network hosting the destination, the hierarchical IP address is converted to a MAC address, before the packet is sent in a frame in the last hop using the MAC address as the identifier of the destination. The MAC address represents a flat address space without information about the host’s locality within the LAN.

Considering the ad hoc network has a flat address space, it can be seen as a network (major network or subnetwork) within the Internet. This will identify an ad hoc network connected to the Internet by its own network number. There is, however, a major difference between an LAN and an ad hoc network.

The ad hoc network can be seen as a Non-Broadcast Multiple Access (NBMA) [75] network and must be managed by the network protocol. Hosts connected to a LAN are within the same broadcast domain, and are managed as one hop connections by the IP protocol. A packet broadcasted in the LAN will reach all hosts connected to the network. In the ad hoc network a broadcast sent by one ad hoc host (MH) may not reach all other hosts. A broadcast needs to be retransmitted by hosts in the network so that it will reach all hosts. A broadcast in the ad hoc network running the IP protocol uses the time to live (TTL) value to limit the spreading of a packet. A packet’s TTL when arriving at a router connected to a LAN requires a value of 1 to reach a host connected to the network. In the ad hoc network a TTL of 1, when forwarded on the ad hoc network will be discarded after the first hop. This behaviour needs to be managed by gateways connecting ad hoc networks with IP networks.

Instead of using the ARP or the NDP, the ad hoc routing protocol needs to be used in what is defined as the last hop in IP networks. The functionality in reactive ad hoc routing protocols maps well to the functionality in the ARP protocol and the NDP, with a request for the host in the last hop and a soft state table. The RREQ in reactive ad hoc routing protocols can be compared to the ARP request, and the soft state ad hoc routing table to the ARP table.

We assume that all MHs in the ad hoc network have an IP address and are able to connect to hosts within the Internet. MHs should be able to move between ad hoc networks and still function as if they were connected to the home network. Peers in the ad hoc network will be able to communicate one-hop or multi-hop regardless of their network part in the IP addresses. This means that two hosts homed in different networks will be able to communicate within the ad hoc network. MHs homed in an ad hoc network will have the network number given the ad hoc network, and MHs
visiting the network will either use their home address as an identifier, an address given by a DHCP server, or if IPv6 is used, stateless auto configuration.

MHs connected to the home ad hoc network will be reached through the IP network by IP routing to the gateway connecting to the ad hoc network. Then the ad hoc routing protocol is used to locate the destination in the ad hoc network.

When an MH in the ad hoc network addresses a correspondent host, we use the way proxy ARP works to manage where to send a packet for a destination in the same network or in another. In the proxy ARP approach, a router looks at the IP address requested, and if the router sees that the destination is in another LAN it responds with its own MAC address. All packets to the IP address will be sent to the router and the router will forward the packet to the destination. The source does not by itself need to investigate if the destination is within the same LAN, the router will support this. If the destination is within the same network it will itself respond.

Instead of the link layer address used in LANs to reach a gateway or a CH in the same network, IP addresses have to be used, since they may be multiple hops away. If the ad hoc routing protocol uses link layer addresses for routing within the ad hoc network, link layer addresses can be used. But in our work the IP address within the ad hoc network is used to identify MHs and to make route decisions.

To manage mobility of MHs connecting to different ad hoc networks, MIP is used. In MIP, messages for advertisement, solicitation and registration are sent link-local. For ad hoc networks the MIP messages must be able to travel multiple hops in the ad hoc network. To enhance the performance of our prototype, routes will be installed to the source of the MIP messages.

**Multihomed Mobile IP**

Multihomed MIP enables an MH in ad hoc networks to register with multiple foreign networks. This enhances performance and reliability. Wireless connections are prone to errors and changing conditions. These must be considered to enable applications for desktop computers to be usable on MHs connecting wireless.

The multihoming is managed by the MIP and hidden from the IP routing, keeping IP routing protocols like the RIP and OSPF unaware. For a sender, multihomed MIP can be considered an any-cast approach [64] where a sender relies on the network protocol to use the best available destination for the packets. The available destination will be one of possibly multiple care-of addresses used by an MH. In IPv6, an any-cast address is used to reach the best available destination (server) among multiple destinations. The approach, taken in this chapter for a sender to any-cast address an MH, suggests that the MH’s home address be used to locate the best care-of address. The difference from the any-cast approach in IPv6 is that it is address-based instead of server-based and the destination will be the same host.

In our approach the MH keeps a list of all networks with valid advertisements and registers the care-of address at the HA (and the CH if MIPv6 route optimization is used) for the networks supporting the best connectivity. To evaluate the connectivity, the MH needs to monitor the performance of the paths in the ad hoc network and gateways connecting the ad hoc networks.
Chapter 6. Extending Global IP Connectivity for Ad Hoc Networks

The HA (and the CH in the case of route optimization) needs to make the selection of which care-of address to use for an MH. In multihomed MIP, IP routing is used to reach the selected care-of address selected but the selection of what care-of address to use is managed by MIP. The HA makes its own selection and the CH does the same if route optimization is used. With route optimization, multiple care-of addresses may be sent.

The choice of care-of address is based on individual selections by the HA, the CH and the MH for packets sent by them. In a scenario where an MH has registered three care-of addresses and there are two CHs, one using the HA to communicate with the MH and the other using route optimization, three different APs may be used: one by the HA, another by the CH using route optimization and the third by the MH to send packets (see figure 6.4).

Global Connectivity and Multihomed MIP

A network can be described by a graph $G = (V, E)$ where $V$ is the set of nodes and $E$ is a set of tuples connecting nodes from the set $V$. If $V = \{v_1, v_2, \ldots, v_n\}$ then the set of $E$ can be $E = \{\{v_1, v_2\}, \{v_3, v_1\}\}$. A path between two hosts $h_1$ and $h_2$ can be described by $P = <h_1, v_i, v_j, \ldots, h_2>$. In MIP a path from a CH to an MH connected to its home network can look like $P = <CH, v_i, v_j, \ldots, MH>$. When the MH connects to a foreign network the path changes to $P = <CH, v_i, v_j, \ldots, HA, v_k, \ldots, v_l, MH>$, where the subset $P' \subset P$, with $P' = <v_k, \ldots, v_l>$ is a tunnel. The node $v_l \in \{FA, gateway-without-FA\}$. If $v_l = FA$ it can be hosted in a host connected to the foreign network or in the router connecting the network (MIPv4). If $v_l = gateway-without-FA$, the HA uses a co-located care-of address.

In the converged network joining wired IP networks and ad hoc networks, no routing information is redistributed. We define two graphs: $G_w = (V_w, E_w)$ describes the wired IP network and $G_a = (V_a, E_a)$ the ad hoc network. With global connectivity a path may be in both graphs $P = <CH, v_{wi}, v_{wj}, \ldots, HA, v_{wk}, \ldots, v_{wl}, v_{wa}, v_{ai}, \ldots, MH>$ with $v_{wa} \in \{FA, gateway-without-FA\}$. $E_w$ is maintained continually while $E_a$ is created on demand (when a route is needed).

In multihomed MIP an MH keeps information on available foreign networks $N_{foreign}$, where $N_{foreign} = \{N_1, N_2, \ldots, N_n\}$ with $N_i = foreign-network$. A second set, $N_{registered}$ where $N_{registered} \subseteq N_{foreign}$ contains the foreign networks registered at the HA. In the connection with ad hoc networks no default gateway is required. The selection is based on the first gateway responding to a RREQ.

The HA (as well as the CH, if route optimization is used) has information about available care-of addresses for an MH in a set $B_{mh}$ where $B_{mh} = N_{registered}$. An IP tunnel is created where tunnel-endpoint $\in B_{mh}$ With route optimization, in MIPv4 tunnel-endpoint $\in B_{mh}$ and in MIPv6, routing-header $\in B_{mh}$.
6.2.2 The prototype

In the prototype described, we use MIPv4, and the AODV routing protocol. Even though route optimization in MIPv4 is questionable because of security reasons, we use the early proposal for route optimization [71]. We propose two changes for route optimization, considering both the early proposal for MIPv4 as well as how MIPv6 manages route optimization. An approach to route optimization in the ad hoc network is also presented. In this section we describe:

- How MIPv4 messages are managed in the ad hoc network;
- How an MH decides what FAs to register with, among several known;
- How a destination is found (located in the ad hoc network or in the Internet);
- How multihoming in MIP is managed;
- The changes made to MIP messages.

An MH uses the AODV protocol to communicate with peers and to access the wired IP network through the gateway (also running AODV). To manage mobility of MHs between ad hoc networks, MHs as well as the gateways run the MIP, where the software for the FA and HA runs in the gateway.

In ad hoc networks a TTL value of the network size in hops is required, since MHs need to rebroadcast these messages to reach all MHs. The MIP messages specified to be sent link-local in a LAN, with a TTL value of 1, are changed to indicate the size of

Figure 6.4. A multihomed connectivity scenario where the HA, CH and MH make their own selections of which care-of address to use.
the ad hoc network or the length in hops to an MH. In MIP, agent advertisement and agent solicitation are sent as link-local broadcast. The response of an agent solicitation is an agent advertisement sent by unicast.

Agent advertisements and solicitations are used to create routes in the ad hoc network to gateways (FAs) and MHS. This will enhance the performance in the ad hoc network. Without this, new broadcasts have to be sent. For example, when an MH receives an agent advertisement and decides to register a care-of address, the MH sends a registration request to the HA through the FA. Without a route created by the advertisement, a RREQ must first be sent by the MH to discover a path.

Our approach enables redistribution between MIP and AODV so that AODV can install routes based on information from MIP. This conforms to the approach used to redistribute information between routing protocols. Reverse routes are created by AODV for agent advertisements. These routes will be used by registration requests creating forward routes. Agent solicitation also creates a reverse route: the path is used by the advertisement sent by unicast to the MH. The responding agent advertisement creates the forward route. With redistribution between MIP and AODV and not requiring AODV to know MIP messages, modularity is sustained.

In MIPv4, the default gateway in an MH visiting a foreign network is modified to use the gateway announcing the agent advertisement. The default gateway is used for all packets sent. Because an MH connecting to a foreign network is not allowed to use the ARP, an entry is added in the ARP table using the IP address and the MAC address of the sender of the advertisement. In our prototype, a default gateway is not required for MHs connecting to ad hoc networks.

An MH discovers a path to a destination by sending a RREQ for the destination. If the destination is known to be within the ad hoc network, the gateway will function as an ad hoc host forwarding the RREQ and the destination will respond with a RREP. Both destinations homed in the ad hoc network and MHs visiting will respond to a RREQ if it is addressed. This can be regarded as route optimization for visiting MHs, without the source first having to send packets to the HA followed by binding updates from the HA to the MH.

For a destination outside the ad hoc network, the gateway replies with a RREP (proxy RREP). Packets will be sent to the gateway that forwards them on the wired network. Proxy RREP is responded by a gateway both for destinations homed in other networks (not visiting) and destinations homed in the network visiting a foreign network. This spares the source from having to find out if a CH is within the ad hoc network or not.

For a gateway to know which hosts are in the ad hoc network, the AODV protocol requires information from the visitor list in the FA. All hosts homed in the ad hoc network have to have the same network number as the gateway interface connecting to the ad hoc network (the visitor list is a part of the MIP information distributed into AODV). When a route is requested for a destination with a network number different from the ad hoc network, the visitor list is searched by the AODV process to see if the destination is available in the ad hoc network. Algorithm 6.1 illustrates this.

The proxy RREP for foreign MHs can be disabled if an MH does not want to reveal its location to a peer within the ad hoc network. Packets will then be sent to the HA and tunnelled back to the MH. To disable proxy RREP a visiting MH uses the P-
flag in the registration request (see figure 5.2). Packets coming to the gateway will be processed as shown in algorithm 6.2.

To manage multiple FAs covering an ad hoc network there is a need to synchronize the visitor information between the FAs. Without synchronization, a gateway may conclude that a destination is within the wired IP network and send a proxy RREP to the source, while the destination is in fact within the ad hoc network but registered with another FA.

The FAs synchronize their visitor lists using the wired IP network to relieve the ad hoc network. The information is synchronized when an entry is added or deleted from the visitor list in a gateway. All gateways will thus be able to see if a visiting host is within the ad hoc network even if it is not registered with the gateway receiving the RREQ. The HA binding cache does not have to be synchronized between the gateways since the gateway responding with a proxy RREP will be the gateway acting as an HA for the MH connected to a foreign network. Other gateways will “believe” that the MH is in the ad hoc network.

When multiple FAs serve an ad hoc network, an MH must be able to choose the best FA(s) to register with. Also, an MH may discover multiple ad hoc networks.

MHs keeps a list of all networks with valid advertisements and registers the care-of address at the HA for the networks supporting the best connectivity, see algorithm 6.3.

Algorithm 6.1. The process in the gateway to manage AODV RREQ messages.

To evaluate the connectivity, the MH monitors the deviation in arrival times of agent advertisements and calculates a metric based on this information (see formula 5.1). This metric is used to describe the MH’s connectivity to foreign networks. A
small metric indicates that agent advertisements sent at discrete time intervals arrive without collisions and without being delayed by the FA. This indicates available bandwidth in the ad hoc network as well as the FA’s capability to relay traffic to and from the MH. The usage of the deviation for the selection avoids inserting explicit control packets like Internet Control Message Protocol (ICMP) [49] echo requests.

\[
\begin{align*}
\text{var} & \quad A_{\text{connected}} : \text{set of connected ad hoc networks}; \\
& \quad W_{\text{connected}} : \text{set of connected wired networks}; \\
& \quad W_{\text{routes}} : \text{set of routes where, } W_{\text{routes}} = \text{routes in the wired network } \cup A_{\text{connected}} \cup W_{\text{connected}}; \\
& \quad A_{\text{routes}} : \text{set of AODV host routes}; \\
& \quad R : \text{set of routes where } R = W_{\text{route}} \cup A_{\text{routes}}.
\end{align*}
\]

Processing a \textit{<data, dest, source>} message:

\begin{verbatim}
begin
receive \textit{<data, dest, source>} from x;
if {..,dest} \in R (* host route *) then
  y \in \{n : \{n, z\} \in R \land z = \text{dest}\}
else if \{.., \text{network(dest)}\} \in R then
  if \text{network(dest)} \in A_{\text{connected}} then begin (* host route required *)
    A_{\text{routes}} := A_{\text{routes}} \cup \{\text{route : discovered route with AODV on network(dest)}\};
    y \in \{n : \{n, z\} \in A_{\text{routes}} \land z = \text{dest}\}
  end
else if \text{network(dest)} \in W_{\text{connected}} then
  y := \text{dest}
else
  y := \text{defaultRoute};
send \textit{<packet, source, dest>} to y
end
\end{verbatim}

Algorithm 6.2. The processing in the gateway of an incoming data packet.

The HA may have multiple bindings for an MH’s home address and needs to select which one to use. Since no routing information is redistributed between the wired IP network and the ad hoc network, the total distance to an MH can not be managed in any of the routing protocols.

The selection of which care-of address to use for an MH is made based on the delay (see algorithm 6.4), where the delay will include both the wired network and the ad hoc network. IP routing is used between the HA or the CH (in the wired network) and the care-of address selected, but the selection of which care-of address to use is managed by multihomed MIP and based on the delay including both networks.

To measure the delay, the round trip time (RTT) based on the registration messages sent between the MH and the HA used. We use the Jacobson/Karels algorithm [19] (see formula 4.1) to calculate the metrics used for comparison. The algorithm accounts for both the RTT and the deviation in the measurements.

To avoid rapid changes resulting in flapping of the care-of addresses registered at the HA and the CH, a new FA is only chosen if the value differs beyond a certain threshold.
Chapter 6. Extending Global IP Connectivity for Ad Hoc Networks

The usage of RTT to classify connectivity is used in protocols like e.g Berkeley Internet Name Domain (BIND) [21].

\[
\begin{align*}
\text{var} & \quad \mathcal{N}_{\text{foreign}} : \text{set of available fa and announced care-of address;} \\
& \quad \mathcal{N}_{\text{reg}} : \text{set of registered care-of addresses;} \\
& \quad \mathcal{M}_{\text{adv}} : \text{array of calculated metrics;} \\
& \quad \mathcal{T}_{\text{regReq}} : \text{array of clock times for RTT measurements;} \\
& \quad \mathcal{A}_{\text{adv}} : \text{set of agent advertisements received;} \\
\end{align*}
\]

Processing a \(<\text{agent advertisement, fa, care-of-address, seqno}>\) message: \textbf{begin}
receive \(<\text{agent advertisement, fa, care-of-address, seqno}>\):
if \(fa \notin \mathcal{N}_{\text{foreign}}\) then \textbf{begin}
\[
\begin{align*}
\mathcal{N}_{\text{foreign}} & := \mathcal{N}_{\text{foreign}} \cup \{fa, \text{care-of-address}\}; \\
\mathcal{M}_{\text{adv}}[fa] & := \text{initialize}; \\
\text{if} \quad |\mathcal{N}_{\text{reg}}| < \text{max care-of addresses to register} \quad \text{then begin} \\
\mathcal{N}_{\text{reg}} & := \mathcal{N}_{\text{reg}} \cup \{fa, \text{care-of-address}\}; \\
\text{if} \quad |\mathcal{N}_{\text{reg}}| > 1 \quad \text{then} \\
\text{set}(n\text{-flag}) \\
\text{else} \\
\text{clear}(n\text{-flag}) \\
\text{send} \quad <\text{registration request, home-address, ha, care-of-address, n\text{-flag}, 0}> \text{to ha} \\
\text{via fa}; \\
\mathcal{T}_{\text{regReq}}[fa] & := \text{clock} \\
\end{align*}
\]
end
\] else if \(<\text{agent advertisement, fa, care-of-address, seqno}> \notin \mathcal{A}_{\text{adv}}, \text{then}\)
\[
\begin{align*}
\mathcal{M}_{\text{adv}}[fa] & := \text{calculated metric according to formula 5.1}; \\
\mathcal{A}_{\text{adv}} & := \mathcal{A}_{\text{adv}} \cup <\text{agent advertisement, fa, care-of-address, seqno}> \\
\end{align*}
\]
end

Processing a \(<\text{registration reply, home-address, ha}>\) message: \textbf{begin}
receive \(<\text{registration reply, home-address, ha}>\) from fa:
\[
\mathcal{T}_{\text{regReq}}[fa] := \text{clock} - \mathcal{T}_{\text{regReq}}[fa] \\
\end{align*}
\]
end

Time expires for a binding to a fa: \textbf{begin}
if \(|\mathcal{N}_{\text{reg}}| > 1\) then
\[
\text{set}(n\text{-flag}) \\
\] else
\[
\text{clear}(n\text{-flag}) \\
\text{send} \quad <\text{registration request, home-address, ha, care-of-address, n\text{-flag}, T}_{\text{regReq}}[fa]> \text{to ha} \\
\text{via fa}; \\
\mathcal{T}_{\text{regReq}}[fa] & := \text{clock} \\
\end{align*}
\]
end

Time expires, compare \(\mathcal{N}_{\text{reg}}\) and \(\mathcal{N}_{\text{foreign}}\): \textbf{begin}
if \(\min(\mathcal{M}_{\text{adv}}[w] : w \in \mathcal{N}_{\text{foreign}} \land w \notin \mathcal{N}_{\text{reg}}) < \max(\mathcal{M}_{\text{adv}}[w] : w \in \mathcal{N}_{\text{reg}}) - \text{threshold}\) \textbf{then}
\[
\begin{align*}
\text{fa} & \in \{w : \min(\mathcal{M}_{\text{adv}}[w]) \land \mathcal{N}_{\text{foreign}} \land w \notin \mathcal{N}_{\text{reg}} \}; \\
\{\text{fa}_{\text{min}}, \text{care-of-address}_{\text{min}}\} & \in \{\{x, y\} : \{x, y\} \in \mathcal{N}_{\text{foreign}} \land x = \text{fa}\}; \\
\text{fa} & \in \{w : \max(\mathcal{M}_{\text{adv}}[w]) \land \mathcal{N}_{\text{reg}}\};
\end{align*}
\]
end
Chapter 6. Extending Global IP Connectivity for Ad Hoc Networks

\[ \{ f_{max}, \text{care-of-address}_{max} \} \in \{ x, y \} : \{ x, y \} \in N_{reg} \land x = fa \}; \\
N_{reg} := N_{reg} \setminus \{ f_{max}, \text{care-of-address}_{max} \}; \\
N_{reg} := N_{reg} \cup \{ f_{min}, \text{care-of-address}_{min} \}; \\
\text{if } |N_{reg}| > 1 \text{ then} \\
\quad \text{set(n-flag)} \\
\text{else} \\
\quad \text{clear(n-flag);} \\
\quad \text{send } <\text{registration request, home-address, ha, care-of-address}_{min}, \text{n-flag, 0}> \text{ to ha via } fa_{min}; \\
\quad T_{regReq[fa_{min}]} := \text{clock} \\
\end{algorithm}

Algorithm 6.3. The processing of MIP messages in the MH.

\begin{algorithm}
\begin{var}
B_{mh} & : \text{set of bindings;}
T_{mh} & : \text{array of tunnels;}
M_{rtt} & : \text{array of calculated metrics;}
\end{var}
\end{algorithm}

Processing a \(<\text{registration request, home-address, ha, care-of-address, n-flag, rtt}>\) message:
begin
receive \(<\text{registration request, home-address, ha, care-of-address, n-flag, rtt}>\) from mh via fa;
\text{if } \{\text{home-address, care-of-address}\} \notin B_{mh} \text{ then begin}
\quad B_{mh} := B_{mh} \cup \{\text{home-address, care-of-address}\};
\quad M_{rtt}[\text{home-address, care-of-address}] := \text{initialize}
\text{end;}
\text{if } \neg \text{n-flag then}
\quad \forall \text{ binding } \in \{ x, y \} : \{ x, y \} \in B_{mh} \land x=\text{home-address} \land y \neq \text{care-of-address} \text{ do}
\quad B_{mh} := B_{mh} \setminus \text{binding;}
\quad M_{rtt}[\text{home-address, care-of-address}] := \text{calculated metric according to formula 4.1 (rtt);}
\quad \text{tunnel} := T_{mh}[\text{home-address}];
\text{if } M_{rtt}[\text{home-address, tunnel}] - \text{threshold} > \min(\{M_{rtt}[\text{home-address, x}] : x \neq \text{tunnel}\})
\quad T_{mh}[\text{home-address}] \in \{ x : \min(M_{rtt}[\text{home-address, x}]) \land x \neq \text{tunnel} \};
\text{send } <\text{registration reply, home-address, ha}> \text{ to mh via fa}
\end{algorithm}

Algorithm 6.4. The processing of registration requests in the HA.

A CH sending packets to an MH without route optimization will send them to the MH’s home network. Here the HA will make the selection of which care-of address to use to forward the packets. With route optimization, the CH will send the packets to the MH without using the HA and will itself tunnel the packets to the selected care-of address.

In multihomed MIP, multiple care-of addresses may be sent. The choice of care-of address is based on individual selections by the HA and the CH for packets sent by them. The gateway selected to send packets from an MH to the IP network is based on the “closest” gateway. The “closest” gateway will be the first gateway responding to a RREQ.
In MIP, an MH receiving an agent advertisement assumes the previous link layer sender to be the FA, and uses this address for its registration request. This must be modified to function in ad hoc networks, otherwise the MH will “believe” an intermediate host in the ad hoc network forwarding the agent advertisement to be the FA. We reserve the first care-of address field in the agent advertisement for the FA address within the ad hoc network (see figure 4.4). The MH can then discover the address of an FA multiple hops away.

The agent solicitation message is also modified to function in the ad hoc network for the same reason as the agent advertisement. The solicitation message used in the standard MIP is the same as the router solicitation message. Figure 4.5 shows the agent solicitation extension added to the router solicitation.

To register a care-of address at the HA, a registration request is sent. And to enable the HA to distinguish between a non-multihomed and a multihomed registration, an N-flag is added to the registration request (see figure 5.2). The first binding update clears the N-flag to erase the binding cache in the HA from possibly stale entries. The rest of the binding updates have the N-flag set. An HA receiving the registration request with an N-flag will keep the existing bindings for the MH. The MH monitors the time between registration requests and registration replies and calculates the RTT. The RTT is added as an extension in the next registration request. The HA will maintain all registrations for an MH and based on the metrics it will install a tunnel into the forwarding table with the care-of address with the smallest metrics.

We propose two changes for route optimization, considering both the proposal for MIPv4 as well as how MIPv6 manages route optimization. For MIPv4 route optimization, the binding update sent from the HA to a CH is shown in figure 5.3. For MIP multihoming, multiple packets must be sent to inform the CH of multiple care-of addresses. The binding update is extended with an N-flag to signal multihomed binding updates. The first binding update clears the N-flag to erase the binding cache in the CH from possibly stale entries. The rest of the binding updates have the N-flag set. The CH will install a tunnel to the MH in its forwarding table based on the information in the binding updates. When an MH performs a handover between networks and changes from one care-of address to another, the binding update needs to express both the new and old care-of address, so that the HA and the CH knows which binding to update. In the case of a single-home binding, only the new care-of address is included.

When a binding is about to expire, CH sends a binding request message to the HA. The binding request must include the care-of address for the binding (see figure 5.4) so that the HA knows which binding to respond with. Without the care-of address included, binding updates are requested for all care-of addresses. An HA will respond with care-of addresses for all available bindings for an MH.

If a binding request is sent for one care-of address and multiple bindings are maintained, the binding update will have the N-flag set. This will inform the CH whether other bindings are maintained as well. If the care-of address requested in the binding request message has no binding in the HA, the HA will respond with a binding update with a lifetime set to 0.

The selection at the CH for which care-of address to use is based on the longest prefix match. In the case of the same prefix match, the first care-of address received is chosen. This avoids sending explicit control packets to monitor the delay.
To manage route optimization where an MN sends binding updates (instead of the HA) to the CHs, as in MIPv6, the same messages is used as described above (see figures 5.3 and 5.4). The advantage of this method for route optimization is that RTT can be measured by the CH by monitoring the departure time of binding requests and the arrival of binding updates. The selection of which care-of address to install as the tunnel end-point is based on formula 4.1.

### 6.3 Related work

Lei and Perkins describe work in [58] to modify RIP for ad hoc networks and with MIP. This approach uses a proactive ad hoc routing protocol less efficient in ad hoc networks. Work described in [45,47,59,60] proposes connectivity between MIP and AODV. Belding-Royer, Sun and Perkins [60] propose MIPv4 and AODV be connected so that MIP messages will be managed in the ad hoc network. The question of how to choose between multiple FAs however is not addressed. Moreover an MH in the ad hoc network has to discover by itself if a destination is within the ad hoc network or not. If the gateway ‘thinks’ it can reach the destination, it replies with an FA RREP (like the proxy RREP). But before an MH can use the gateway, it first needs to conclude that the destination is not within the ad hoc network and this will delay the connection setup time. MIPv6 management with AODV is proposed by Wakikawa et al [59]. Belding-Royer, Sun and Perkins [60] take the same approach for destinations in the IP network and propose that router advertisements should not be sent without router solicitation. However, Jonsson et al. [45] give measurements that show it is more efficient to use the normal MIP behaviour where advertisements are sent without solicitations. Jonsson et al. [45] and Sun, Belding-Royer and Perkins [47], describe an approach to choosing between multiple FAs. Here the selection is made based on the hop count between the FA and the MH. Hop count may however not be the best way to measure which FA to register with since network load is not considered. Broch, Maltz and Johnson [46] address MIP and the ad hoc routing protocol DSR.

In MIPv4, an option for simultaneous bindings is proposed for sending packets to multiple care-of addresses for an MH. Packets will be duplicated at the HA and one copy sent to each registered care-of address, so that packets can be received through multiple APs. This option is being proposed to decrease the number of dropouts of packets during handover, and for an MH with bad connections to APs to receive the same packet through several APs, with an increased probability of a good connection. The solution does not however enable network layer decisions about the best connection to use, and it wastes resources in the ad hoc network.

In the current specification of MIPv6, all traffic uses the same care-of address. This prevents the MIP from fully utilizing the dynamics within global connectivity including ad hoc networking, and should therefore be altered.
6.4 Conclusion and further work

In this chapter, we propose and describe an integrated connectivity solution and its prototype connecting IP networks and ad hoc networks running the reactive AODV routing protocol, where multihomed MIP is used for mobility as well as multihoming of MHs. The work presented combines our previous two proposals; Åhlund and Zaslavsky [10] and Åhlund and Zaslavsky [14] into a connectivity solution for MHs. This paper extends previous work from Åhlund and Zaslavsky [76].

The proposal supports interconnection of ad hoc networks by wired IP networks and the creation of areas covered by gateways connecting wireless MHs to the Internet. MHs will be able to communicate peer-to-peer or with hosts in the IP network. Our approach proposes a new way to locate a destination inside the ad hoc network or in the IP network, and to select FA(s) based on the RTT between the MH and the HA. A new solution of how to select between multiple FAs based on the deviation in agent advertisements is also presented.

MIP is extended to manage multiple simultaneous connections with foreign networks. Based on the registered care-of addresses, multiple paths can be used for packets to and from an MH. Enhanced throughput and a more reliable connection are achieved. With multihomed MIP we propose a new way to manage multihoming in MIP that differs from the proposal in MIPv4. In MIPv6 there is currently no multihoming functionality managed by the HA.

The current prototype is based on MIPv4, but can be applied in MIPv6 as well. It will also be useful to study the impact of the delay between the measurement of the RTT and the time the HA receives the information, since the RTT is sent one registration request later than when it was measured.

The next chapter presents a simulation study for evaluating the performance at the network layer of APs and gateways as well as the wireless channel.
Chapter 6. Extending Global IP Connectivity for Ad Hoc Networks
Chapter 7. Running Variance Metric for evaluating performance of Wireless IP Networks in the MobileCity Testbed

4 This chapter is based on the publication:


Minor changes have been made to the publication to improve the presentation.
Chapter 7. Running Variance Metric for evaluating performance of Wireless Networks in the MobileCity Testbed
This chapter proposes and analyzes a Running Variance Metric performance measurement of wireless local area networks and its formal aspects. Our approach evaluates the performance of wireless local area networks in infrastructure mode as well as in ad hoc mode. The Running Variance Metric is used to discover relative traffic loads of available access-points/gateways at the network layer in order to provide connectivity to the wired network. The chapter discusses a simulation study. The simulation results demonstrate the usefulness and efficiency of the Running Variance Metric to evaluate the utilization of available access-points/gateways. It is also shown that this metric can be used for hop-analysis in multi-hop ad hoc wireless networks.

7.1 Introduction

This chapter proposes and discusses an approach to evaluate the relative traffic load at the network layer when connecting to access points (AP) used in infrastructure networks and gateways connecting between wired IP networks and ad hoc networks. This is useful for a mobile host (MH) using Mobile IP (MIP) [8] and for Global Connectivity [15] during handover or when being multihomed and selecting the AP to use. When using MIP with infrastructure networks, the MH has to rely on the datalink layer to make a good decision on which AP to use if multiple APs are available. After associating with the AP, the network layer is able to discover the network connecting the AP and register according to MIP. The decision made at the datalink layer may not be optimal considering the performance based on throughput. To enable this there is a need to discover the network layer performance when deciding which AP to use. With ad hoc networks connectivity to gateways connecting to wired IP networks also needs a way to decide which gateway to use. Proposals given for this are usually based on the hop-count as described in [45]. Another solution is presented in [47]. However, a dynamic metric reflecting the utilization will be beneficial for this decision.

7.1.1 Infrastructure Networks

MHs when connecting to an AP make decisions based on the signal-to-noise ratio (SNR) and related factors. This information originates at the physical layer and is analyzed at the datalink layer in the IP-stack. However, SNR does not reflect the performance of the AP at the network layer. This means that calculating the SNR values will not be enough to decide the best AP to associate with considering the throughput. In some situations a better throughput can be achieved by using APs with
lower SNR values. With the same SNR the throughput may also differ. According to the 802.11 [1] standard, MH3 in figure 7.1 may associate with AP1 even though more traffic is sent by MH1 and MH2 than by MH4 and MH5. Or, in other words, AP2 is carrying less traffic than AP1. As illustrated by the left circle, MH3 is out of communication range from MH1 and MH2, and cannot detect collisions generated by these nodes in the SNR calculation. In 802.11 there is also a Network Allocation Vector (NAV) that is used by a sender to signal the time needed to send a frame. With the usage of NAV fewer collisions will occur. So it is clear that the SNR is not appropriate to use as the only metric when deciding which AP to use.

Figure 7.1. A sample topology.

For infrastructure Wireless Local Area Networks (WLANs), calculations based on measurements at the network layer can be used to decide which AP to use if multiple APs are available.

7.1.2 Ad Hoc Networks

For ad hoc networks where gateway connectivity to the wired network is required, the network layer performance should be used both when multiple gateways are available as well as when an MH has multiple paths to a gateway.

In existing networks with today’s traffic pattern, most network traffic is to destinations outside a LAN. The 20/80 ratio used to classify today’s network traffic indicates that 20% of the network traffic is within that LAN, and 80% of the traffic is to destinations outside the LAN. This means that 80% of the traffic has to go through the gateway.

In our model, we consider ad hoc networks as subnetworks [10], and that services like the Domain Name Service (DNS), Dynamic Host Configuration Protocol (DHCP) remain external to the ad hoc networks. This is due to the fact that MHs are mobile with a high probability of moving to other networks. However, there is
ongoing research on how to support these services inside ad hoc networks, for example, DNS services in ad hoc networks [77].

Based on these observations, maintaining connectivity to gateways is important, and choosing the one with the best performance will improve the throughput. The routing protocols proposed for ad hoc networks (e.g. DSR [30], AODV [11]) usually assume the same capacity for all links across the network, and use the hop count as the routing metric. Therefore a 2 hop route will be preferred over a 3 hop route despite the utilization of links. Even though the 2 hop route carries more traffic than the 3 hop route it will be selected. Ad hoc routing protocols that are considering only the hop count will face the same problems as RIP version 1 does in wired IP networks. Dynamic metrics need to be proposed and applied to ad hoc networks to overcome these problems. In this chapter we limit the scope of dynamic metrics to gateway connectivity only.

We propose a complementary metric that will enable an MH to evaluate the performance of a wireless link at the network layer and to choose the AP/gateway which provides the best throughput.

This chapter is structured in the following way. Section 7.2 describes the formal reasoning used to calculate the Running Variance Metric. Section 7.3 describes a simulation model and the results of that simulation and section. Related work is presented in section 7.4 and section 7.5 summarizes the chapter.

### 7.2 Running Variance Metric

To evaluate the relative traffic load of available APs/gateways we use periodical advertisements sent by them. These advertisements can be router advertisements [78](available in IP version 4 (IPv4) and IP version 6 (IPv6)) or agent advertisements in MIP version 4 (MIPv4) [23]. In MIP version 6 (MIPv6) [24], the router advertisement in IPv6 is used. With increased traffic the AP/gateway may not cope with incoming and out-going traffic. This will lead to buffering of advertisements and collisions between advertisements and traffic. If the send buffer at an AP/gateway is full, some advertisements will be dropped. When the link becomes less congested two or more advertisements could be sent in more dense succession. This, in turn, means that with increased traffic the arrival times of advertisements at MHs will vary. Collisions of advertisements also affects the arrival times, since these advertisements are destroyed and do not arrive at MHs. We make use of the variance in arrival times of advertisements to evaluate the degree of links load. The following formulas introduce the variance metric.

Formula 7.1 calculates the mean value of the time between arrivals of advertisements and is based on the formula for weighted mean (\( \bar{x}_n \)) values [79].

\[
\mu_n = \frac{\sum_{i=1}^{n} x_i}{n}
\]

Formula 7.2 then calculates the variance (\( V_n \)) of the arrived advertisements and this is used for the evaluation of wireless links. The variable \( t_n \) is the arrival time of the last advertisement, \( t_{n-1} \) is the arrival time of the previous advertisement. The variable \( n \) symbolizes the number of advertisements received since the MH started to receive advertisements from an AP/gateway. With the variable \( h \) we select a history window
Chapter 7. Running Variance Metric for evaluating performance of Wireless Networks in the MobileCity Testbed

expressing how long history to consider when calculating the mean value and variance.

\[
\bar{x}_n = \frac{1}{h} x_n + \frac{h-1}{h} \bar{x}_{n-1} \tag{7.1}
\]

\[
V_n = \frac{1}{h} (x_n - \bar{x}_n)^2 + \frac{h-1}{h} V_{n-1} \tag{7.2}
\]

The variables \( h \), \( \bar{x}_0 \) and \( V_0 \) are initialized with the following values:

\[
\frac{1}{h} \in (0,1] \text{ where } (0,1] \text{ is the half open interval } \{x : 0 < x \leq 1\}
\]

\( V_0 = 0 \)

\( \bar{x}_0 \) = Defined advertisement time

The variable \( x_n \) is calculated as:

\( x_n = t_n - t_{n-1} \) where \( n \) is a integer > 0

Formula 7.2 is an approximation of the mathematically defined variance and is shown by:

\[
V_n = \frac{1}{n} \sum_{i=1}^{n} (x_i - \bar{x}_n)^2 = \frac{1}{n} \left[ (x_n - \bar{x}_n)^2 + \sum_{i=1}^{n-1} (x_i - \bar{x}_n)^2 \right] = \frac{1}{n} \left[ (x_n - \bar{x}_n)^2 + (n-1) \frac{1}{n-1} \sum_{i=1}^{n-1} (x_i - \bar{x}_n)^2 \right]
\]

We put \( V_{n-1} = \frac{1}{(n-1)} \sum_{i=1}^{n-1} (x_i - \bar{x}_n)^2 \) \( \Rightarrow V_n = \frac{1}{n} \left[ (x_n - \bar{x}_n)^2 + (n-1)V_{n-1} \right] \)

\[
\therefore V_n = \frac{1}{n} (x_n - \bar{x}_n)^2 + \frac{n-1}{n} V_{n-1}
\]

The approximation is created by \( V_{n-1} = \frac{1}{(n-1)} \sum_{i=1}^{n-1} (x_i - \bar{x}_n)^2 \) where \( \bar{x}_n \) includes \( x_n \).

The previous variance metric would not include \( x_0 \), only \( x_0 \) to \( x_{n-1} \) is included for the “true” variance in the mean value. We will refer to our calculation of the variance as the “Running Variance Metric” (RVM) in the rest of the chapter. The next section will discuss the simulation study based on the RVM.
Chapter 7. Running Variance Metric for evaluating performance of Wireless Networks in the MobileCity Testbed

7.3 RVM Simulations study

This section evaluates the RVM calculation and how RVM is applied in the analysis of wireless links in infrastructure mode and in independent BSS mode (ad hoc mode). Our simulation study uses the GlomoSim simulation model version 2.4 [65].

Simulation study results are presented in figures 7.2, 7.4, 7.5, 7.7 and 7.9. The graphs with error bars represent the mean value of multiple simulations (different seeds) using a confidence interval of 95%. Our simulation study has selected two packet sizes based on the publications [80,81]. In [80] it is stated that the major parts (50%) of the packets have the size of the Maximum Transmission Unit (MTU). We choose an MTU of 1500 bytes in the simulation, being the MTU of Ethernet. The second most widely used MTU is 576 bytes [81]. Packets about this size are, except for TCP traffic, used for UDP traffic, for example for Voice over IP (VoIP). The advertisements used in the simulations have a size of 32 bytes.

Figure 7.2. The correlation between the RVM and the “true” variance.

Our simulation first analyzes the difference between the RVM and the “true” variance in the following way. Advertisements are sent every second from an AP with varying load. This load is based on different numbers of MHs communicating through the AP with varying throughput. Figure 7.2 shows the correlation between the RVM and the “true” variance. The solid green curve plots the “true” variance and the red dotted curve plots the RVM. The figure shows 105 calculations of the variance with 40 to 60 values \(x_n\) in each calculation. The range of values generated by the simulation is between 0.96-5.0 seconds. The graph shows a good correlation between the RVM and “true” variance. In this simulation we used \(h=n\) for this comparison.
7.3.1 Infrastructure Networks

To demonstrate the RVM’s capability for discovering the relative traffic load in wireless infrastructure networks we use the topology shown in figure 7.3. From one to five MHs send wireless traffic ranging between 0.5 Mbps and 1.5 Mbps with an MTU of 1500 bytes through the AP. The monitoring MH does the RVM calculation. The bandwidth used is defined by 802.11b and is 11 Mbps.

![Figure 7.3. The infrastructure mode topology used in the simulation.](image)

The results from the simulation are shown in figure 7.4. The RVM increases with the number of nodes as well as with the amount of traffic sent. Advertisements are sent once every second by the AP. The node monitoring the variance is only within communication range from the AP and not the other nodes.

![Figure 7.4. RVM calculations in infrastructure mode, with a packet size of 1500 byte.](image)

The plotted lower curve shows the RVM when up to five nodes send 0.5 Mbps each to the AP. The middle curve shows the same for 1 Mbps/MH and the upper
curve shows the RVM for 1.5 Mbps/MH. The RVM demonstrated for one node sending 0.5 Mbps, 1 Mbps and 1.5 Mbps is too small to be shown in the graph presented here. However the RVM is doubled for each increased step of the traffic. The big jump of the RVM in the upper and middle curves is explained by the fact that the link is congested, resulting in more collisions.

The same simulation was tried using a transmission unit of 576 bytes. The results are shown in figure 7.5. With smaller packets the RVM for 1.0 Mbps and 1.5 Mbps tend to converge near saturation in the wireless link. This is due to small differences in the deviation of advertisements between the two flows when the link is nearly congested. However, for each added node the RVM increases.

In figure 7.5 the big jump appears before congestion. This is explained by the increased number of packets sent with a packet size of 576 bytes compared to a packet size of 1500 bytes. The number of collisions therefore increases, rendering in big contention windows.

![Figure 7.5. RVM calculations in infrastructure mode, with a packet size of 576 byte.](image)

### 7.3.2 Ad Hoc Networks

Figure 7.6 and 7.7 depicts the wireless multi-hop networks used for the simulations of ad hoc networks. These topologies have been used to evaluate the RVM calculation in ad hoc networks when all wireless links use the same channel.
Chapter 7. Running Variance Metric for evaluating performance of Wireless Networks in the MobileCity Testbed

Figure 7.6. The ad hoc mode topology used in the simulation for calculating RVM at each hop.

The simulation study looks at RVM from the view point of differentiation in the number of hops an advertisements travels as well as the utilization of multi-hop routes. Every node only sees one or two neighbors. A wireless link capacity of 2Mbps is used in the simulation.

The first simulation uses the topology shown in figure 7.6. Advertisements are sent by the gateway and forwarded from MH1 to MH10. The RVM calculated at each hop is presented in figure 7.7. As shown, the RVM increases for each hop.

Figure 7.7. The RVM calculated at each hop.

To see how added traffic flows affect the RVM, we use the topology shown in figure 7.8. We monitor the RVM after 5 hops (in MH5) and insert up to four additional 0.5 Mbps flows between MH11 and MH18. The radio ranges of these will only affect MH1 and MH2.
Chapter 7. Running Variance Metric for evaluating performance of Wireless Networks in the MobileCity Testbed

Figure 7.8. The ad hoc mode topology for RVM monitoring in MH5.

Figure 7.9 shows results of the simulation where the monitored RVM at MH5 increases for each inserted 0.5 Mbps flow.

Figure 7.9. The RVM in MH5.

7.4 Related work

With mobility of MHs between WLANs handover is managed both at the datalink layer and the network layer, when MIP is used. The network layer handover can take place only after handover at the datalink layer has been completed. The related work presented here focuses on enhancing the performance of MIP in wireless access networks and to minimize the number of packets lost due to handover. The research described in [31-33,36,37,39,82]. All uses the SNR to decide the AP to associate with. A solution enabling an MH to select the AP based on the load at the network layer is presented in [38].
Methods for horizontal and vertical handovers are discussed in [36,37]. These approaches use multicast to reach multiple nearby APs. MHs instruct APs to forward or buffer data packets for it. If not delivered to the MH, these packets are dropped after some time.

In [31] a proposal is presented to lower the delay with MIP messages and thereby manage handover at the network layer more efficiently, considering the time for handovers. The proposal uses two care-of addresses; link local care-of address (LCoA) and regional care-of address (RCoA). A Mobility Anchor Points (MAP) is used. A MAP manages multiple networks and can be hosted in a gateway connecting an autonomous system (AS). When an MH enters an AS it requires two addresses, LCoA and RCoA. The RCoA address is registered at the home agent (HA) and the LCoA is used for registrations with the MAP. A binding between the RCoA and LCoA is maintained in the MAP. As long at the MH remains within the networks controlled by the MAP the only binding update needed when moving between different networks is the LCoA address sent to the MAP. The registration at the HA remains unaffected.

A solution for fast handover [32] in MIP uses signaling between the MH, the old AP and the new AP entered to avoid losing packets. Packets will be forwarded from the old to the new AP to avoid packet losses.

Another solution [33] combines the proposals [31] [32] and extend it to lessen the handover time even further. The handover time in this work is the same as the handover times for datalink layer handovers. In this proposal the MH decides when to handover and the network decides where to handover. The network monitors the MHs movement and based on this makes the decision of which AP to use next.

A solution to policy-enabled handoffs is proposed in [38]. This solution is based on three factors: power consumption, cost and bandwidth. The bandwidth usage is monitored and announced by APs so that MHs can calculate the utilization of an AP. This information is used to decide which AP to use.

In [39] a proposal using MIP is given to decrease the time for handover, and to lessen the packet drops. An MH doing handover at the datalink layer tells the old foreign agent (FA) to buffer packets for it. After the MH associates with a new FA, the HA tells the old FA to forward buffered packets to the new FA. In the proposal an FA sending agent advertisements includes a neighbor list in the message. The neighbor list includes the IP address, link layer type and channel information of the neighbors. The information enables the MH to select which FA to handover to. To avoid having to wait for three times the advertisement time (as specified in the MIP specification) to discover loss of connection to a FA, a signal from the datalink level is used to inform the network layer. All agents need to know the position of all neighbor FAs.

In [82] support for fast handover is managed at the datalink level. This proposal is based on the usage of a MAC bridge assisting in bridging packets to a roaming MH’s new location, while MIP registration is in process. This avoids losing packets during network layer handover. The delay for handover where packets can be lost only includes the datalink-layer-handover time. This method only works as long as all MHs do handover to APs connected to the MAC bridge. In a real system this is hardly the case, but for micro mobility it can be used.
[45,47] discusses connectivity between wired IP networks and ad hoc networks where MIP is used for mobility between networks. In [45] the hop-count is used for the decision of which FA to use. Handover is triggered when the hop-count to a new FA is two hops less than to the FA currently used. The proposal for gateway selection in [47] uses the following criteria: the MH has not heard from its FA for at least one advertisement interval, and the MHs route to the FA has become invalid. When this happens handover occurs.

The related work presented addresses how to achieve a more effective handover at the network layer and the decision where to handover both in wireless infrastructure networks and in ad hoc networks. Except for [38] none of the related work addresses network layer performance for this decision. In [38] the traffic carried by an AP/gateway is advertised but it does not give a metric describing how congested the channel is.

7.5 Chapter summary

This chapter addresses performance measurement in WLANs. We have proposed and shown how to discover the relative traffic load at MHs in the network layer when connecting wirelessly to APs/gateways. Our methodology uses passive measurements based on advertisements like MIP agent advertisements and router advertisements. The RVM can be used in infrastructure mode as well as in ad hoc mode. With increased traffic on a wireless link, collisions will increase and packets will be delayed in buffers. The simulation study reported in this chapter demonstrates that RVM is a complementary metric. It can be used in combination with SNR in infrastructure networks and hop count in ad hoc networks to improve efficiency and throughput of wireless communications between MHs and APs/gateways. In ad hoc networking This simulation study also supports the theoretical contribution presented in [14,15].

RVM will be used with MIP and Global Connectivity solutions to manage handover and multihoming. We use RVM with Multihomed MIP in [14] to associate with multiple APs. With the proposed approach it is possible to select the least loaded AP(s) when two or more APs are used. No double casting or multicasting is needed because the MH is connected to multiple APs receiving unique packets. Multiple associations are maintained in order to evaluate the performance of APs. In [15] the RVM is used to evaluate multihop connectivity to gateways in ad hoc networks.

A small “ground” variance should be used for sending advertisements [78], so that a flow (possibly with low utilization of the wireless link) with the same timing as the advertisements does not put out advertisements by colliding with them.

The next chapter describes a simulation study of multihomed MIP where RVM is used to decide the AP.
Chapter 8. M-MIP : extended Mobile IP to maintain multiple connections to overlapping wireless access networks

5 This chapter is based on the publication:


Minor changes have been made to the publication to improve the presentation.
Chapter 8. M-MIP: extended Mobile IP to maintain multiple connections to overlapping wireless access networks
M-MIP: extended Mobile IP to maintain multiple connections to overlapping wireless access networks

In future wireless access networks, connectivity to wired infrastructure will be provided through multiple access points with possibly different capabilities and utilization. The demand for increased network performance requires the ability to predict the best overall performance of those access points and to switch access point when the performance changes. Then there is the demand for mobility between networks with maintained connectivity which requires the ability to switch the point of attachment. Mobile IP support is needed to allow mobile hosts to move between networks with maintained connectivity. Multihomed Mobile IP enables mobile hosts to register multiple care-of addresses at the home agent, to enhance the performance of wireless network connectivity. Multihomed Mobile IP enables performance discovery at the network layer and the capability to decide what AP to use. This chapter describes a simulator evaluation of multihomed Mobile IP.

8.1 Introduction

With increasing demands for wireless connectivity and mobility support, new solutions are required to maintain the wireless network connection and to optimize the performance. This is important for mobile hosts (MHs) both moving and when stationary for a period of time. The major access technology used today in wireless local area networks (WLAN) is 802.11 [1]. The support of mobility and handover at the datalink layer enables flows to be maintained within the same network. However mobility between networks is no supported since this requires handover at the network layer. For this, Mobile IP (MIP) [8] is proposed.

When combining wireless access (802.11) and network mobility (MIP) association with an access point (AP) is managed at the datalink level without interference from the network layer. An MH decides which AP to associate with based on the signal to noise ratio (SNR) and related factors. The MH needs to associate to receive MIP agent advertisements used to discover available networks. If the MH discovers a foreign network (or if the MH arrives back to the home network), it requires a registration with the home agent (HA). Since the performance at the network layer may not be reflected in the SNR, the association may be with an AP having bad performance. With a high SNR metric the actual performance can still be low since an MH may not sense collisions from other MHs using the same AP if they are out of communication range. Also, since the Network Allocation Vector (NAV) is used in 802.11, hosts will defer their communication and thereby avoid collisions. Therefore MIP cannot entirely rely on the datalink level to make the right decision about the selection of an AP. Instead, network layer characteristics need to be considered.

To enable this, performance discovery at the networks layer is required and the capability to decide what AP to use. This can be achieved with multihoming.
Multihoming is enabled using a single wireless network card switching between APs [83] or by using multiple network cards. By maintaining multiple network connections, network layer performances can be compared and the best one selected.

Handover can be classified into soft and hard handover. With soft handover the association with the old AP is sustained while associating with a new AP. In this way two connections will be maintained for some time. With hard handover the connection to the old AP is ended before associating with a new AP.

In this chapter we present an approach to multihoming with MIP, called M-MIP. With M-MIP, passive network layer measurements are enabled by maintaining multiple registrations at the HA. In this way we can maintain connectivity and handle handovers without generating delays due to MIP registrations.

The chapter is structured in the following way. Section 8.2 describes the architecture of M-MIP. Section 8.3 describes a simulation study and the results of the study. In section 8.4 related works are presented and section 8.5 provides a summary.

8.2 M-MIP

This section briefly describes the changes made to MIP to enable multihoming functionality (M-MIP). For a more detailed description see [14]. M-MIP enhances the performance and reliability of MHs connections to WLANs. The multihoming is managed by the M-MIP and hidden from the IP routing process.

To register a care-of address at the HA, a registration request is sent by the MH. To enable the HA to distinguish between a non-multihomed and a multihomed registration, an N-flag is added to the registration request (see figure 5.2).

An HA receiving the registration request with an N-flag will keep the existing bindings for the MH. If a registration is received without the N-flag, the HA will clear the existing bindings for the MH which makes M-MIP compatible with standard MIP. One of the registered care-of addresses will be used to forward packets to the MH. To enable the selection at the HA, a metric is added as an extension in the registration request. The HA will maintain all registrations for an MH and based on the metrics it will install a tunnel to the selected care-of address into the forwarding table.

To enable an MH to select the “best” AP to use, we evaluate the performance of an AP at the network layer. In M-MIP the MH keeps a list of all networks it receives agent advertisements from and registers the care-of address of the network(s) supporting the best connectivity, with respect to the throughput, at the HA. To evaluate the connectivity, the MH monitors the deviation in arrival times between MIP agent advertisements and makes a Running Variance Metric (RVM) calculation based on this information (see formula 7.1 and 7.2).

The RVM is used to evaluate MHs wireless connectivity to foreign networks. A small RVM indicates that agent advertisements received at regular time intervals arrive without collisions and without being delayed by the (foreign agent) FA. This indicates available bandwidth as well as the FA’s capability to relay traffic for the MH.

The RVM is then added to the round trip time (RTT) between the MH and it’s HA shown in formula 8.1 and 8.2.
Chapter 8. M-MIP: extended Mobile IP to maintain multiple connections to overlapping wireless access networks

\[ \bar{x}_n = \frac{1}{h} x_n + \frac{h-1}{h} \bar{x}_{n-1} \]  

where \( x_n \) symbolizes the \( n \):th RTT measurement and \( \bar{x}_n \) is the weighted mean value

\[ RNL_n = \bar{x}_n + V_n \]  

where \( V_n \) symbolizes the RVM value

\[ \frac{1}{h} \in (0,1] \text{ where } (0,1] \text{ is the half open interval } \{ x : 0 < x \leq 1 \} \]

\( \bar{x}_0 = \) is set to the first RTT measurement

This formula is defined as the Relative Network Load (RNL). The calculation is carried out at the MH and the metric is attached to the next registration request sent to the HA. The RTT measure is based on the registration messages sent between the MH and the HA.

The measurements and metric calculations are made prior to registration and maintained while being registered at foreign networks. Since the MH may register multiple foreign networks, the HA can have multiple bindings for an MH. Among the registered care-of addresses, the FA with the smallest RNL metric will be installed as the default gateway in the MH and as the selected care-of address at the HA.

With route optimization it is possible to choose a different FA (to communicate with the correspondent host) than the FA used to communicate through the HA. An MH (as in MIPv6) sends binding updates to the CH with available care-of addresses. By requesting the CH to respond to binding updates with an acknowledgement, RTT can be measured by the MH. We then have the same functionality between CHs and the MH with route optimization as the registrations between the MH and it’s HA.

8.3 M-MIP analysis using RNL based simulation

In this section we present our work simulating M-MIP with the network simulator GlomoSim, version 2.4 [65]. The topology used is shown in figure 8.2.

The simulation evaluates how well M-MIP discovers the utilization of APs and, based on this, selects the AP with the best network layer performance, considering the throughput.

Agent advertisements are sent every second and the MH registers every 3 seconds with the HA. This is based on the MIP specification, where the timeout for a binding is three times the agent advertisement time. At each received advertisement the MH calculates the RNL metric and based on this decides which FA to use. The MH then attaches the RNL metric to the next registration request message.

The MH registers with two foreign agents (FA1 and FA2) using different channels and maintain multiple bindings with the HA. Hereby the HA as well as the MH maintain the RNL metric for each connection.
To add load to the wireless links we use the hosts LoadMH1 to LoadMH10 communicating with FA1 and FA2. We will use the phrase load traffic in the text below to name this traffic between the LoadMHs and the FAs. Based on the load traffic, we investigate how M-MIP responds to this load. The throughput presented in the graphs is the traffic sent by the peerMH and received at the MH, with and without using M-MIP. We name this traffic the monitored traffic.

Load traffic between peers is sent in both directions: the hosts LoadMH1 to LoadMH5 communicate with FA1 and LoadMH6 to LoadMH10 with FA2. The monitored traffic is also sent in both direction between the MH and the peerMH. Since the throughput presented looks similar in both the MH and the peerMH, we only present the monitored traffic for the MH.

Without using M-MIP, we evaluate the monitored traffic when the MH associates with an FA based on the SNR, without considering the performance at the network layer.

We use different combinations of traffic types (TCP and UDP) for the evaluation. For UDP traffic we use Constant Bit Rate (CBR) traffic and for TCP we use the generic File Transfer Protocol (FTP) provided by GlomoSim.

In our scenarios, the combination of traffic types for the load traffic and the monitored traffic is as follows:
- FTP is used as the load traffic and CBR as the monitored traffic;
- CBR is used as the load traffic and FTP as the monitored traffic;
- All hosts use FTP traffic;
- All hosts use CBR traffic.

We run each scenario with the two major packet sizes used in the Internet: 1500 bytes and 576 bytes [80,81]. Although another frequently used packet size is 40 bytes (ACK packets in TCP), we do not look into this size.

In the graphs the solid line plots the throughput with M-MIP and the dashed line with a SNR-selected AP. In figures 8.3 to 8.6 the x-axis shows the number of LoadMHs generating load traffic. The y-axis shows the throughput of the monitored traffic received at the MH. The load traffic pattern is as follows: the first 10 seconds
up to five LoadMHs add traffic to FA1; then 10 seconds to FA2. This is then repeated with a 20 second interval as well as a 30 second interval. The time to discover a loaded FA using the RNL calculation is about 2 seconds in all simulations.

The results are presented as mean values of multiple simulations (different seeds) and the error-bars express a 95% confidence-interval.

In all graphs below with an MTU of 576 bytes: less data in sent in each packet resulting in a lower throughput. This occurs because the time it takes to sense the channel and there is a settling time for the interface each time a frame is sent. Overflow in buffers also take place.

Figure 8.3 plots the result from the scenario where FTP is used as load traffic. Here traffic between the MH and the peerMH uses CBR traffic. The plotted solid green line is the throughput with a packet size of 1500 bytes using M-MIP. Behind the green line is a dotted blue line plotted showing the throughput with the SNR selected AP. The red lines show the throughput with a packet size of 576 bytes. Both the MH and the peerMH send 2.5Mbps CBR traffic.

As expected, there is no difference between M-MIP and choosing the AP based on the SNR. The reason for this is that FTP (the TCP mechanism) degrades throughput caused by collisions, while CBR (UDP) continues sending at the same rate, forcing FTP to continue degrading its throughput.

Figure 8.3. CBR traffic received at MH with FTP traffic as load.

In figure 8.4a we show the results where all hosts use CBR traffic with an MTU of 1500 bytes. The blue lines plot the monitored traffic when up to five LoadMHs generate load traffic of 0.25 Mbps. The green curves plot the same for load traffic of 0.5 Mbps and the red line for 0.75 Mbps. In figure 8.4b this is repeated for an MTU of 576 bytes.
Chapter 8. M-MIP: extended Mobile IP to maintain multiple connections to overlapping wireless access networks

Figure 8.4. CBR traffic received at MH with CBR traffic as load with an MTU of 1500 bytes and 576 bytes.

The results from the scenario where all hosts use FTP traffic is plotted in figure 8.5. The throughput with a MTU of 1500 bytes and a MTU of 576 bytes shows the...
same results. FTP using an MTU of 1500 bytes is plotted by the blue line and the
green line plots throughput with the MTU of 576 bytes.

![Figure 8.5 FTP traffic received at MH with FTP traffic as load.](image)

The results from the last scenario are shown in figure 8.6, where CBR is used as
the load traffic, and where monitored traffic uses FTP communication. In figure 8.6a,
load traffic with a MTU of 1500 bytes are shown. The blue line plots the FTP traffic
received at the MH with each LoadMH sending and receiving 0.25 Mbps. The green
line plots the same with load traffic of 0.5 Mbps and the red line with load of 0.75
Mbps. In figure 8.6b this is repeated for an MTU of 576 bytes.

In all scenarios M-MIP (plotted by solid lines) perform better than when only the
SNR (dashed lines) is considered. An interesting observation from the last scenario
(plotted in figure 8.6) is that the throughput increases with increased load as plotted in
some of the curves.

The reason for this is that we do not consider how traffic communicated by the MH
affects the RNL. Before communication takes place the MH monitors the RVM and
RTT and calculates the RNL metric. The RNL metric is sent to the HA in a
registration request. Based on the metric a FA is selected. When communication takes
place we continue to monitor the RVM and RTT and calculate the RNL metric. Since
MHs own traffic affects the metric a new selection of FA may take place, selecting
the FA being more loaded (not considering the own traffic). This will happen for both
CBR and FTP traffic. With CBR traffic this happens if the MHs traffic increases
beyond the difference between the least loaded FA and the next least loaded FA. With
FTP, since TCP is used, the MH will take as much of the available link as possible,
rendering a handover. This is most visible in the red curve in figure 8.6a and 8.6b.
With a small difference in RNL, handover to the more loaded FA happens more often,
keeping the sending window smaller. The same happens in all scenarios, but it is most
visible in the last simulation. It also means that the performance of M-MIP will
increase if we can avoid “false” handovers.
Chapter 8. M-MIP: extended Mobile IP to maintain multiple connections to overlapping wireless access networks

Figure 8.6. FTP traffic received at MH with CBR load traffic using a MTU of 1500 bytes and 576 bytes.

One solution to handle “false” handovers is for the MH to predict how much the own-added flow increases the metric. However this is difficult. We are not able to say that X kbps effects the RNL metric with a value of Y. This depends on the utilization of the link, e.g. whether it is near congestion or not. Another option for the MH is to calculate the difference between the RNL metric after starting to send the own flow with the RNL metric before doing so. However the resulting metric may be in error. Let us say that another host begin communicating at the same time, the calculated difference will be too big. Or that a host that communicated stops, the calculated difference will be too small. A more straight forward solution is to make a decision
when selecting the FA and starting to communicate. After that the FA cannot change for that flow. As soon as communication stops, new selections become possible. If all MHs behave in the same way we will have a distribution of MHs between APs.

In the case where route optimization is not used all traffic will use the FA selected for the HA. With route optimization multiple FAs may be used. This is possible since a unique binding update is sent to each CH.

8.4 Related work

In MIPv4 [23] a proposal to multihoming is presented, sending one copy of a packet to each AP an MH is associated to. This means sending duplicated packets in the wireless media wasting scarce resources. In MIPv6 [24] there is no proposal for multiple bindings enabling multihoming with MIP.

MIP similar methods for handovers using IP multicasting are discussed in [36-38]. A multicast address is used to reach nearby APs in WLANs where the MH is located. An MH instructs one of the APs listening at the multicast address to forward packets to it, and the other APs to buffer packets. When doing handover the MH first tells the previous AP to stop forwarding packets and the new AP to start doing so. In [36,37] the MH decides which AP to use based on the SNR. The AP having the best SNR is ranked as the best one to use. However, this may not be true in the topology shown in figure 8.2 when the LoadMHs is out of radio range from the MH.

In [38] the bandwidth usage is monitored by APs. This calculated bandwidth utilization is announced in beacons sent by the AP. Our approach decides which AP to use based on network layer characteristics and does not require any modification of existing WLAN infrastructure compared to [38]. [39] suggests a proposal using MIP to decrease the time for handover and to reduce the number of dropped packets. An MH doing handover at the datalink layer tells the old FA to buffer packets for it. After the MH associates with a new FA, the HA tells the old FA to forward buffered packets to the new FA. In the proposal, an agent advertisement includes a neighbour list in the message. The neighbour list includes the IP address, link-layer type and channel information. The information is used to enable the MH to select which FA to handover to. To avoid having to wait for three times the advertisement time (as specified in the MIP specification) to discover loss of connection to a FA, a signal from the datalink level is used to inform the network layer. Here all agents need to know the position of all neighbour FAs. This is not required in our proposal.

In [82] support for fast handover is managed at the datalink level. This proposal is based on the usage of a MAC bridge assisting in bridging packets to a roaming MH's new location, while MIP registration is in process. This avoids losing packets during network layer handover. The delay for handover where packets can be lost only includes the datalink layer handover time. This method only works as long as all MHs do handover to APs connected to the MAC bridge. In a real system this is hardly the case, but for micro mobility it can be used.

More related work is presented in [32,33]. Compared to our proposal a high message complexity is required.
Chapter 8. M-MIP: extended Mobile IP to maintain multiple connections to overlapping wireless access networks

8.5 Chapter summary

This chapter addresses performance measurements in WLANs. We have proposed and shown how to discover the relative load at MHs in the network layer when connecting wirelessly to APs. Our methodology uses passive measurements based on advertisements like MIP agent advertisements and router advertisements. With increased traffic on a wireless link, collisions will increase and packets will be delayed in buffers. The simulation study reported in this chapter demonstrates that RVM is a complement metric that can be used in combination with SNR to improve efficiency and throughput of wireless communications between MHs and APs. This simulation study also supports the theoretical contribution presented in [14].

We have presented a proposed and validated solution to Multihoming in MIP named M-MIP. M-MIP enables an MH to discover multiple networks and to register them at the HA. We have also presented a new metric called RNL that is used to select a care-of address. A simulation study describing the performance of our approach is presented and discussed.

The work presented in this chapter has focused on improving performance of MHs using MIP and connecting to 802.11 access networks by enabling MHs to associate with multiple FAs and to evaluate the performance at the network layer. M-MIP gives a higher throughput using RNL than if the selection is based only on the SNR. With multiple FAs, one FA will be used for traffic sent through the HA and other FAs can be used for CHs using route optimization. With M-MIP soft handover is enabled, allowing an MH to use multiple FAs. A roaming MH will receive unique packets through both FAs. When the MH decides to handover, it will register with the new FA at the same time as it uses the old FA. With registration completed, packets will be sent using the new FA. With this approach loss of packets because of handover can be avoided. M-MIP does not require any new types of MIP-messages.

Compared to other proposals to enable soft handover with MIP, we present a solution that do not require extended message complexity or modified APs. We use the messages proposed by MIP and analyses the network performance based on this messages.

The next chapter summarizes and analyzes the results of the thesis work presented in chapters 4 to 8.
Chapter 9. Analysis of results and contribution

This chapter evaluates the work presented in this thesis, its results and outcomes. In addition, most recent research results are also included.

Section 9.1 explains the results from the simulation model introduced in chapter 7. In section 9.2, an M-MIP scenario is presented that shows how M-MIP operates based on the results from chapters 5 and 8. Section 9.3 titled “Global Connectivity” includes extended and most recent work that continues the research described in chapters 4 and 6. Section 9.4 presents algorithms expressed in pseudo code to illustrate the logical flow and functionality of the prototypes and simulation models.

9.1 RVM

This section describes how the Running Variance Metric (RVM) presented in chapter 7 reflects the load of APs and wireless channels. RVM is shown to be useful for comparing the performance of different access points (AP). RVM can be used for selecting the least loaded AP. The mobile host (MH) monitors the deviation in arrival times of agent advertisements and based on this calculates the RVM. The deviation is caused by agent advertisements delayed in buffers due to queuing and the time taken to access the wireless channel.

When an agent advertisement is broadcasted by an AP, the channel is first detected to be free. Once the channel is free the AP selects a random time-slot in the contention window and it waits for a distributed inter-frame space (DIFS) time. When the DIFS time expires, the MH starts a back-off timer equal to the time of the randomly selected time-slot. If the channel remains idle for this time, the FA will transmit its frame. If the channel instead becomes busy, the back-off timer is suspended until the channel is free again. When this happens, the FA will have to wait for another DIFS time and the back-off timer can be resumed. This is repeated each time the radio senses a busy channel while waiting for the back-off timer to expire.

The smallest size of the contention window in this thesis is 32 time-slots and it may extend to 1024 time-slots in the case of multiple lost request-to-send (RTS). Each time-slot is 20 microseconds. With a random number of 10 when picking a time-slot in the contention window, the frame will be sent after 200 microseconds plus the DIFS time. Each time the radio senses a busy channel within this period, the back-off timer is suspended waiting for the channel to become free. When free the AP has to wait for a new DIFS time and then the back-off timer can be resumed. In the simulation study (in chapter 7) the DIFS time is equal to 45 microseconds.
When sending unicast traffic, the contention window is managed in the same way as with broadcast frames, if the unicast frame size is below the threshold for request to send (RTS). For each sent unicast frame an acknowledgement (ACK) frame is returned by the receiver. If the frame size is beyond this threshold, a RTS will first be sent by the sender of the unicast frame and in response a clear to send (CTS) is returned by the receiver. When the CTS is received, the frame can be sent. In the simulation studies presented in chapter 7 and 8 RTS-CTS is always used for unicast frames.

If a unicast frame is sent without lost RTS, CTS and ACK, the same contention window is used as with broadcast (i.e 32). If the RTS, CTS or ACK is lost, the contention window is increased exponentially for each loss up to a maximum of 1024 time-slots.

With a suspended back-off timer and increasing size of a contention window, the variance in sending agent advertisements will increase. With competing data frames of 1500 bytes, an agent advertisement will be delayed with approximately 2.2 milliseconds (for 11 Mbps) each time the channel becomes busy because of such a data frame sent by another MH.

An FA which sends only agent advertisements will send the frames selecting a random time-slot in a contention window with the size with 32 time-slots. Without contention or buffering, the agent advertisements will be sent with a low deviation. The deviation in times will be in the range of 0 to 620 microseconds. Figure 9.1 shows agent advertisements sent in a timely order without delays because of a busy media. The contention window to the right shows the size of the contention window for each agent advertisement sent.

![Figure 9.1. Agent advertisements broadcasted in a timely manner.](image)

With an increasing number of MHs competing for the channel, the probability that the radio interface will detect other MH’s traffic increases. When this happens, the back-off timer is suspended and after the channel becomes available, is resumed. Figure 9.2 shows a scenario where an FA sends only agent advertisements at the same time as nearby nodes transmit data frames. The data packets have a size of 1500 bytes suspending the back-off timer approximately 2.2 milliseconds for each such frame.

If buffering in output queues takes place and the unicast traffic is sent, the contention window is increased if a RTS, CTS or ACK is lost. The increase is exponential with a maximum window size of 1024, giving a time span between 0 and 20.48 milliseconds, excluding the suspension of the back-off timer due to a busy
channel. So, with a combination of unicast and agent advertisement frames in the output buffer, deviation will increase even more.

Figure 9.2. Broadcast of agent advertisements delayed by competing traffic from other MHs.

With other wireless MAC protocols using a distributed coordination function where a contention window is used and where the window size increases when sensing a busy channel and collisions, the deviation will increase with increased window size and when the back-off timer is suspended and resumed.

With increasing contention windows and suspended back-off timers the RVM will have a high value even if the load is constantly high.

9.2 M-MIP

M-MIP enables an MH to register multiple care-of addresses, where different care-of addresses can be used by the HA and CHs capable of managing route optimization.

In M-MIP, the decision of which care-of address to select is made independently, however the same metric is used at both ends of a registration process and binding updates. Hence, the traffic between the MH and it’s HA will use the same care-of address and the same comes to the traffic between the MH and each of its CH when route optimization is used. CHs using route optimization may use a different care-of address compared to the one used by the HA.

The care-of address selected by the MH and HA will be installed as the default gateway in the MH. All CHs using the HA to reach an MH will use this care-of address. For CHs using route optimization, a host route entry will be installed in the MH. Figure 9.4 shows a routing table in the MH for the topology given in figure 9.3.

In this scenario the MH receives a topology correct address from each subnetwork and registers these addresses with its HA and the CHs using route optimization. The address 0.0.0.0 is the entry for the default gateway and it points to the gateway selected to be used for traffic between the HA and MH as well as for traffic to CHs not using route optimization. In figure 9.3, 130.240.10.100 is used as the default gateway. CH1 and CH2 in figure 9.3 use route optimization and the MH can therefore update the CHs with its care-of addresses. In this scenario the gateway 130.241.100.10 is selected for communication with CH1 and the gateway 130.240.10.10 is selected for communication with CH2.
In M-MIP, RVM is used to evaluate the load of different APs. Another option would be for APs to announce their load in agent advertisements. An advantage of using RVM is that it will better reflect the load in the channel since RVM (by increasing contention windows, suspended and resumed back-off timers) reflects the load in the channel compared to receiving bandwidth information from the AP. In the case of a congested channel it may be that only a fraction of the packets make it through the AP. In this case the utilization of the AP reflected in agent advertisements will be low, rendering a possibly wrong decision in the MH of which AP to use. However, with RVM there is no way to tell how much traffic is carried by an AP. RVM can only be used to compare the performance of different APs.

When using RVM for making a decision of which AP to use, the own flow will affect the metric after starting to send packets, since RVM is continuously calculated. The reason for doing so is that a flow to a new CH will select the least loaded AP considering the current load of APs.

When selecting an AP and starting to send packets, the gateway selection for the peer may not change until any of the following occurs:

---

### Figure 9.3.
A network topology where M-MIP is used.

---

### Figure 9.4.
The routing table in the MH shown in figure 9.3.

<table>
<thead>
<tr>
<th>Address</th>
<th>Mask</th>
<th>Next hop</th>
<th>Interface</th>
<th>Metric</th>
</tr>
</thead>
<tbody>
<tr>
<td>130.10.100.10</td>
<td>255.255.255.255</td>
<td>130.240.10.100</td>
<td>130.240.10.200</td>
<td>X</td>
</tr>
<tr>
<td>130.100.100.30</td>
<td>255.255.255.255</td>
<td>130.241.100.10</td>
<td>130.241.100.100</td>
<td>Y</td>
</tr>
<tr>
<td>130.241.100.10</td>
<td>255.255.255.255</td>
<td>130.241.100.10</td>
<td>130.241.100.100</td>
<td>Z</td>
</tr>
<tr>
<td>130.240.10.100</td>
<td>255.255.255.255</td>
<td>130.240.10.100</td>
<td>130.240.10.200</td>
<td>U</td>
</tr>
<tr>
<td>0.0.0.0</td>
<td>0.0.0.0</td>
<td>130.240.10.100</td>
<td>130.240.10.200</td>
<td>V</td>
</tr>
</tbody>
</table>
• The datalink connection to the associated AP is lost;
• The MH stops communicating with the peer for a specified period of time;
• The network layer connection is considered lost due to three successive lost agent advertisements as defined by MIP.

The RVM in the simulations presented in chapter 7 gave a metric between 0 and 0.2. The Relative Network Load (RNL) used to decide the care-of address for a peer adds the round trip time (RTT) to the RVM.

Since the RVM varies between 0 and 0.2 and the measured RTT is usually below 100 milliseconds, these two metrics are added without weight. By adding the RTT, a value between 0 and 100 milliseconds is added to the RVM.

9.3 Global Connectivity

The approach to global connectivity is a combination of proactive and reactive approaches. Connectivity to gateways is continuously maintained by agent advertisements and is based on the assumption of small ad hoc networks with the same traffic characteristics as in wired IP subnets. Here the major part of the traffic is to CHs outside the ad hoc network. Connectivity between peers within the ad hoc network is reactive.

According to the MIP specification, agent advertisements are to be sent “link local”. Since we consider ad hoc networks as subnetworks, the advertisements are modified to be sent via multiple hops. The same agent advertisement may then arrive through multiple paths to an MH. The decision of which gateway to use is based on the RVM and RNL.

When using the RVM to select a gateway, each MH keeps an array consisting of \{gateway-address, last-hop, RVM\}. The reason for maintaining the last hop is explained by the scenario drawn in figure 9.5.

In a simulation study using \{gateway-address, RVM\} as the information to select the gateway (GW1 or GW2), it was discovered that GW1 was selected instead of GW2, even though paths to GW1 were more congested by other traffic. The computed RVM is based on advertisements from GW1 and gives a lower RVM than the one computed from GW2. The reason is that there are four nodes (MH1 to MH4) that are able to relay the advertisements and the MH relaying differ from advertisement to advertisement. With a route between GW1 and MH6 only one of those nodes will be used. So the RVM does not reflect the load of a single path from GW1 to MH6. By adding the last hop address to the information maintained for a gateway, the RVM can be monitored for each path between GW1 and MH6.

The selection of which agent advertisements to rebroadcast is based on the RVM. The agent advertisement from a previous hop giving the lowest metric for a gateway is rebroadcasted. Figure 9.6 shows a scenario where two gateways (GW1 and GW2) send agent advertisements. MH1 and MH2 receive advertisements directly from GW1 and GW2 via MH3. MH3 receives agent advertisements from GW1 via MH1 and MH2 and directly from GW2. MH3 then selects the advertisements with the lowest RVM for GW1 and rebroadcasts these advertisements. In figure 9.6 this will
be the advertisements from MH1. The advertisements from GW2 are also rebroadcasted.

Figure 9.5. A simulation topology where MHs calculate the RVM.

The reason for rebroadcasting advertisements from both gateways is to enable an MH to register multiple care-of addresses at the HA as well as the CHs using route optimization. Since our proposal only considers small ad hoc networks this is feasible.

Figure 9.7 shows a scenario with a node (MH4) visiting foreign networks. MH4 receives agent advertisements from both gateways. The gateway used for the HA will be set as the default gateway. If MH4 in figure 9.7 discovers that the route to GW2 is the best route, this care-of address is used to communicate with the HA and hence is selected as the default gateway. The functionality of default routes in currently implemented routing tables assumes the default gateway to be of one hop distance. This means that if MH4 decides to use GW2 in figure 9.7, MH4 will have MH3’s IP address (130.240.10.110) configured as the default gateway.

At the time MH4 makes its decision, MH3 will also have the lowest RVM value to GW2. When MH4 starts to send traffic through GW2 the RVM value in MH3 for GW2 may increase to a value higher than the RVM value calculated for GW1. As defined earlier, a gateway should not be changed while traffic is sent through it in order to avoid flapping between gateways. This means that MH4 should not change gateway until it stops communicating with the peer for a specified period of time or in case the connection to the gateway is lost.
Figure 9.6. A scenario showing the propagation of gateway information.

If MH3 is not sending or receiving any traffic it is free to select a new gateway. If the RVM value for GW2 increases beyond the RVM for GW1, MH3 selects GW1 as its default gateway and the traffic sent by MH4 will be rerouted to GW1. To avoid this and to make an MH aware of which gateway it uses, tunneling to the selected gateway is required. This approach differs from the one given in chapter 6 in that the MH uses the default gateway registered with it’s HA when sending packets to a peer (if route optimization is not used). In chapter 6, the functionality of the reactive ad hoc routing protocol was sustained by the MH sending a route request for all destinations regardless of the destination’s IP address. However, with this approach a gateway not associated to the MH may respond. In the case of tunneling between the FA and the HA to avoid ingress filtering it is required that the MH uses any of the gateways registered at the HA. Also, since RVM is used to decide the path to a gateway it should be used both for packets sent and received by the MH.

The routing table created in MH4 for the scenario in figure 9.7 is shown in figure 9.8. MH4 uses GW2 as its default gateway. GW1 is selected for communication to CH1 and GW2 is used to communicate with CH2.

To enable tunneling, virtual interfaces are used. In figure 9.8, the virtual interface 0 is the interface managing tunneling to GW2 and virtual interface 1 manages the tunnel to GW1. When a packet is sent to a virtual interface, an outer IP header is added to the packet. If MH4 sends packets to CH1 in figure 9.7 there will be two iterations in the routing table. In the first iteration, the forwarding process identifies the destination address 130.100.100.30 and sends the packets to the virtual interface 1. This interface is a process that adds an outer header to the packet. The IP address in the outer header will be the address of GW1, i.e., 130.241.100.10. Now the packet is returned to the forwarding process for a second iteration. This time the entry
130.241.100.10 is selected. The packet will then be sent to interface 130.100.10.210 with 130.240.10.100 as the next destination.

Figure 9.7. A network topology creating the routing table in figure 9.8

The registration request message carries the RNL metric as described in chapter 8 and the decision of which care-of address to use is based on this metric.

An MH communicating with a CH that has the same network number as the MH is connected to, uses AODV to discover the route. If the CH has moved to another network the HA will respond to the route request with a route reply. The packets will be sent to the HA that tunnels them to the CH’s current location.

<table>
<thead>
<tr>
<th>Address</th>
<th>Mask</th>
<th>Next hop</th>
<th>Interface</th>
<th>Metric</th>
</tr>
</thead>
<tbody>
<tr>
<td>130.10.100.10</td>
<td>255.255.255.255</td>
<td>130.240.10.100</td>
<td>Virtual int. 0</td>
<td>-</td>
</tr>
<tr>
<td>130.100.100.30</td>
<td>255.255.255.255</td>
<td>130.241.100.10</td>
<td>Virtual int. 1</td>
<td>-</td>
</tr>
<tr>
<td>130.241.100.10</td>
<td>255.255.255.255</td>
<td>130.240.10.110</td>
<td>130.100.10.210</td>
<td>3</td>
</tr>
<tr>
<td>130.240.10.100</td>
<td>255.255.255.255</td>
<td>130.240.10.110</td>
<td>130.100.10.210</td>
<td>2</td>
</tr>
<tr>
<td>0.0.0.0</td>
<td>0.0.0.0</td>
<td>130.240.10.100</td>
<td>Virtual int. 0</td>
<td>-</td>
</tr>
</tbody>
</table>

Figure 9.8. The routing table created in MH4 in figure 9.7.

If the CH has a network number that differs from the network where the MH is connected, the packets will be sent to the default gateway using the maintained route
based on agent advertisements. If the default gateway running the FA has the CH registered as a visitor in the network, an Internet Control Message Protocol (ICMP) [49] redirect is returned to the MH. The MH will then request a route to the CH using AODV. If the CH is outside the network, the gateway will forward the packets according to the IP routing protocol in the wired IP network.

When selecting a gateway and starting to send packets, the gateway selection for CH’s may not change until any of the following occurs:

- An agent advertisement is lost from the selected gateway, and the RVM computed for some other gateway has becomes lower than the RVM of the selected gateway at the time the selection was made;
- The MH stops sending and receiving packets from the CH for a specified period of time;
- The network layer connection is considered lost due to three successive lost agent advertisements as defined by MIP.

To maintain routes to gateways and to be able to manage MIP messages without enforcing new broadcasts, the active time out time in AODV is set to the registration timeout in MIP. The period of time a route remains active without being used is in AODV called the active route timeout. A route not used within this time is erased. Agent advertisements are sent once a second and the timeout time for MIP registrations is three times the agent advertisement time (as defined by MIP). This gives a timeout time of MIP registrations of three seconds. This is the same time as the active route timeout proposed in AODV.

With these timeout settings a route from an MH to a gateway is maintained by agent advertisements, registration and binding replies. And a route from a gateway to an MH is maintained by registration requests and binding updates.

### 9.4 Algorithms

This section presents several algorithms that highlight the modifications and extensions made to algorithms proposed and described in chapter 6.

The algorithms are expressed in pseudo-code form and are intended to demonstrate the logical flow and the key functionality. The following variables are used in the algorithms:

<table>
<thead>
<tr>
<th>var</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>$N_{\text{foreign}}$</td>
<td>set of discovered FAs;</td>
</tr>
<tr>
<td>$N_{\text{reg}}$</td>
<td>set of registered FAs;</td>
</tr>
<tr>
<td>$N_{\text{ch}}$</td>
<td>set of CHs using route optimization;</td>
</tr>
<tr>
<td>$M_{\text{rvm}}$</td>
<td>array of calculated RVM;</td>
</tr>
<tr>
<td>$M_{\text{rnl}}$</td>
<td>array of calculated RNL;</td>
</tr>
<tr>
<td>$M_{\text{rtt}}$</td>
<td>array of RTT measurements;</td>
</tr>
<tr>
<td>$M_{\text{coa}}$</td>
<td>array of care-of addresses for FAs;</td>
</tr>
<tr>
<td>$M_{\text{gw}}$</td>
<td>array of gateways selected for the HA and CHs;</td>
</tr>
<tr>
<td>$M_{\text{gwRNL}}$</td>
<td>RNL metrics to the HA and CHs for the selected gateways;</td>
</tr>
<tr>
<td>const</td>
<td>MAX_REG_FA : max FAs to register at HA and CHs;</td>
</tr>
</tbody>
</table>


119
Chapter 9. Analysis of results and contribution

9.4.1 M-MIP

The algorithms described below assume the FAs to advertise one care-of address. The processing of agent advertisement in an MH is shown in the algorithm 9.1. When an advertisement arrives from a new FA a registration request is sent if it is possible to register another FA. If the FA is known by previous agent advertisements, the RVM is calculated including the arrived agent advertisement. When registering a new care-of address at the HA, all CHs are updated with the new address. The timestamp when the registration request is sent along with the binding update(s) are stored to calculate the RTT when the registration reply and ACK(s) are received. If there is a previously registered care-of address, the N-flag is set by the operation $|N_{reg}| > 1$.

Processing *agent advertisement*: begin
receive $\langle$agent advertisement, care-of-address$\rangle$ from fa;
if $fa \notin N_{foreign}$ then begin
$N_{foreign} := N_{foreign} \cup \{fa\}$;
$M_{rvm}[fa] := \text{initializeRVM}$;
$M_{coa}[fa] := \text{care-of-address}$;
if $|N_{reg}| < \text{MAX\_REG\_FA}$ then begin
$N_{reg} := N_{reg} \cup \{fa\}$;
send $\langle$registration request, home-address, care-of-address, 0, $|N_{reg}| > 1$, 0$\rangle$ to ha via fa;
$M_{rtt}[ha, fa] := \text{clock}$;
forall $ch \in N_{ch}$ then begin
send $\langle$binding update, home-address, care-of-address, 0, $|N_{reg}| > 1$, true, 0$\rangle$
to ch via fa;
$M_{rtt}[ch, fa] := \text{clock}$
end
end
else $M_{rvm}[fa] := \text{updateRVM}(fa, \text{clock})$
end

Algorithm 9.1. The processing in an MH when an agent advertisement is received.

When a registration needs to be updated, a registration request is sent to the HA including the RNL. The timestamp when the registration request is sent is stored (see algorithm 9.2).

Processing a time-trigged registration request to a ha: begin
send $\langle$registration request, home-address, care-of-address, 0, $|N_{reg}| > 1$, $M_{rtt}[ha, fa]$ $\rangle$ to ha via fa;
$M_{rtt}[ha, fa] := \text{clock}$
end

Algorithm 9.2. The processing in an MH for periodical updates.

When receiving a registration reply, the RNL is calculated and based on this metric the default gateway is selected. This gateway is used for all CHs not using...
Chapter 9. Analysis of results and contribution

route optimization (see algorithm 9.3).

Processing registration reply: begin
receive <registration reply>, home-address, ha> from fa;
Mr[ha, fa] := clock - Min[ha, fa];
Mr[ha, fa] := updateRNL(Mr[ha, fa], Mrvm[fa]);
if Mrvm[ha] > min(Mrvm[ha, fa]) + threshold ∨ Mrvm[ha] = 0 then begin
Mr[gh, mh] := argmin{Mrvm[ha, y] : y ∈ Nreg};
Mrvm[ha] := Mrvm[ha, Mr[gh, mh]]
end
end

Algorithm 9.3. The processing in an MH to decide the default gateway.

The FAs registered with is included in Nreg are compared with the FAs not registered with and included in Nforeign. If an FA in Nforeign not registered with has a lower RVM than one in Nreg a switch of FAs will happen (see algorithm 9.4).

Processing comparison of Nforeign and Nreg: begin
let rvmnotreg := min {Mrvm[w] : w ∈ Nforeign \ Nreg};
let rvmreg := max {Mrvm[w] : w ∈ Nreg};
if rvmnotreg < rvmreg - threshold then begin
let famin = argmin{Mrvm[w] : w ∈ Nforeign \ Nreg};
let famax = argmax{Mrvm[w] : w ∈ Nreg};
Nreg := Nreg \ {famax} ∪ {famin};
send <registration request, home-address, ha, Mcoa[famin], Mcoa[famax], |Nreg| > 1, 0 > to ha via famin;
Mrvm[ha, fa] := clock;
forall ch ∈ Nch then begin
send <binding update, home-address, Mcoa[famin], Mcoa[famax], |Nreg| > 1, true, 0> to ch via fa;
Mrvm[fa] := clock
end
end

Algorithm 9.4. The processing in an MH deciding if a change of registered FAs is needed.

When the MH receives a binding request a binding update is sent in response. For each binding update an ACK is requested. If the optional field for the care-of address is set in the binding request only this care-of address is used for the binding update. Without this field added all care-of addresses are sent to the CH through the FAs providing the address. This functionality can also be managed by the HA, the only difference from the algorithm is that no RTT is measured (see algorithm 9.5).

Processing binding request: begin
receive <binding request, home-address, care-of-address> from ch via fa;
if care-of-address ≠ ε then begin
send <binding update, home-address, care-of-address, 0, |Nreg| > 1, true, Mrvm[ch, fa]> to ch via fa;
Chapter 9. Analysis of results and contribution

Algorithm 9.5. The processing in an MH when a binding request is received.

The binding ACK received via different FAs is used to calculate the RNL metrics. The FA giving the lowest RNL is selected for communication with the CH (see algorithm 9.6).

Processing binding ack : begin
receive <binding ack> from ch via fa;
M_in[ch, fa] := clock - M_in[ch, fa];
M_in[ch, fa] := updateRNL(M_in[ch, fa], M_rvm[fa]);
if M_gwRNL[ch] > min(M_in[ha, fa]) + threshold ∨ gatewayRNL[ch] = 0 then begin
M_gw[ch] := argmin_y {M_in[ch, y] : y ∈ Nreg };
M_gwRNL[ch] := M_in[ch, M_gw[ch]]
end
end

Algorithm 9.6. The processing in an MH when a binding acknowledge is received.

9.4.2 Global Connectivity

The algorithms proposed and discussed in the previous section are also used for the proposed global connectivity solution. It should be noted that for the global connectivity solution MIP messages are sent multiple hops.

When data is received at the gateway it may operate as an ad hoc node forwarding the data in the ad hoc network or act as a gateway forwarding the packets outside the network. Packets received via a tunnel with the gateway address will be decapsulated and forwarded according to the inner IP header destination field.

If the destination is visiting the ad hoc network, the gateway will send an ICMP redirect message to the source. If there is a route for the destination in the gateway, the packets will be sent that route.

If the packets are destined for an MH that has a binding to a foreign network, they will be tunneled to the care-of address.

In the case of a packet received without tunneling for a destination homed outside the network and not visiting, an ICMP redirect message is returned to the source and the packets are dropped.
Chapter 9. Analysis of results and contribution

9.5 Chapter summary

This chapter discusses the results from the simulations using the proposed Running Variance Metric (RVM), the operation of proposed multihomed Mobile IP (M-MIP) and extended architecture for global connectivity.

The settings of the wireless protocol in the simulations were described, how they operate and how the RVM calculation detects a loaded channel and AP.

A scenario using M-MIP describes how the MH operates and manages routes to peers.

Extended work with global connectivity was presented, describing how gateways are maintained and routes created.

Several algorithms have also been presented. These algorithms demonstrate the logical flow and key functionality of the prototypes and simulation models.

The next chapter concludes the thesis, identifies the place and impact of the thesis work in state-of-the-art research and outlines future work.
Chapter 9. Analysis of results and contribution
Chapter 10. Conclusion

The challenges and objectives of the research described in this thesis have been to enable a dynamic wireless infrastructure where mobile hosts (MHs) communicate with and without support of a network infrastructure; enhance the reliability and availability of wireless connections for ad hoc networks connecting to wired IP networks; and enable ad hoc networks to dynamically become a part of the Internet.

10.1 Achievements of this thesis work

The work described in this thesis presents and discusses the following three contributions:

- Two types of dynamic metrics that can be used to analyze and compare the load of access points (AP) at the network layer, in infrastructure networks and for connectivity to gateways in ad hoc networks. The first metric is called the Running Variance Metric (RVM) and has been discussed in chapter 7. The second metric is called the Relative Network Load (RNL) and has been discussed in chapter 8.
- An extension to Mobile IP (MIP) which enables multihoming for MHs. This extension is called Multihomed MIP (M-MIP) and has been described in chapters 5 and 8.
- A proposed gateway architecture which integrates ad hoc networks with wired IP networks and the Internet. This architecture has been described in chapters 4 and 6.

The outcomes of this thesis work are reflected in 10 published peer-refereed papers that have been presented at major international conferences. One paper has been published in a major telecommunications journal. These publications are a combination of theoretical ideas, discussions, simulation studies and prototyping.

The first publication [84] presents a project called MobileCity that was initiated by this work at the very beginning. The publication also presents an early proposal for a solution connecting wired IP networks and ad hoc networks. The idea behind MobileCity was to bring ICT companies and Universities together to do research within the area of this thesis.

The publications [85,86] present early ideas of multihoming and global connectivity in wireless networks and were intended to properly position this work and to present those ideas at research forums. In [87] these have been given a more business oriented approach.
Chapter 10. Conclusion

The first developed prototype is presented in [10], this publication proposes and describes a software solution to connect ad hoc networks to wired IP networks. The next prototype extending MIP with multihoming functionality is described in [14]. A combination of these two prototypes with new features has been presented in [76] and further extended in [15].

Validation of prototypes and theoretical discussions as well as reporting about experimental evaluation with simulations have been presented in [16,17,18]. The publication [16] presents a simulation with multihoming extension for MIP. In [17] the proposed RVM is described and results from simulations calculating RVM are presented. Based on results presented in [16,17] previous ideas with multihomed extension to MIP has evolved and been extended in [18].

RVM uses advertisements sent by APs to analyze and compare the load at the network layer. These advertisements could be router advertisements or MIP agent advertisements. By evaluating the load of an AP at the network layer a better metric of the load in the wireless media with respect to throughput is achieved than when considering only signal related factors.

As described in chapter 7, signal related factors can not be credibly used to discover the load at the network layer. Even though RVM can not tell the exact load of an AP, it is possible to compare the load of APs at the network layer. RVM can also be used in ad hoc networks and reflects the number of hops as well as the utilization of paths.

M-MIP enables the home agent (HA) and correspondent hosts (CH) to maintain multiple bindings to an MH. M-MIP supports multiple registrations with foreign agents (FA) for an MH. The MH decides the FAs to register with and hence the care-of addresses to register with its HA as well as updating the CHs capable of managing route optimization.

The RVM is used to decide which care-of addresses to register and the relative network load (RNL) is used to decide which of the registered care-of addresses to use when starting to communicate. The RNL metric is formed by adding RVM and the round trip time (RTT). The RNL metric between the MH and its HA uses the RTT between these two, and the RTT between MH and a CH is used for the RNL metric between them.

The RTT metric is used to reflect the distance to an HA and CHs as well as the load along the path. With RNL, the wireless channel plays a more important role than the wired network since it affects both the RVM and RNL metrics. It is assumed that the wireless link will be the connection with the scarcest resources available.

Even though the work presented in this thesis uses RVM and RNL, the design of M-MIP can use other metrics as well. Another option would be to use APs advertising the available bandwidth in router or agent advertisements.

The existing MIP uses the datalink layer for handover decision. M-MIP, on the other hand extends MIP with the extra benefit of making a handover decision based on network layer characteristics. In this case the network layer should be able to control associations at the datalink layer. M-MIP is also useful in overlapping networks selecting the least loaded AP and to balance the load generated by MHs between the APs.

The proposed RVM, RNL and M-MIP can support connectivity of ad hoc networks to wired IP networks. The selection of which gateway to use by an MH is
based on RVM calculations of agent advertisements relayed via multiple hops to reach all MHs in the network. Registration messages and binding updates are also managed through multiple hops. The measured RTT is used to calculate the RNL. Routes to gateways are maintained by the advertisements sent continuously.

All traffic to destinations within the same network as the MH resides in, is sent according to ad hoc routes created by AODV. For destinations outside the network, packets are sent to the gateway. If a packet arrives at the gateway for a destination inside the network, an Internet Control Message Protocol (ICMP) [49] redirect message is sent to the MH. This enables a proactive approach for connectivity to gateways and a reactive approach for connectivity to CHs within the network. The AODV protocol identifies a peer within the same network. By using a tunnel from an MH to its gateways, the default gateway functionality in current implementations of routing tables can be sustained.

10.2 Contribution to the state-of-the-art research

In comparison to related work presented in chapter 3 this thesis makes the following three contributions:

- The decision of which AP/gateway to use based on network layer performance;
- An extension to MIP to manage multihoming where network layer characteristics are used to decide the care-of address;
- A proposal for a gateway architecture that integrates ad hoc networks to wired IP networks and therefore with the Internet.

The selection of AP is usually based on the signal strength and related factors as reflected in several publications surveyed in chapter 3. Some proposals use APs advertising the bandwidth it relays, so that an MH can select the least loaded AP to associate with. The approach using RVM differs in that it uses advertisements sent at the network layer and the deviation of these messages is used as a metric for how loaded an AP is and how congested the wireless media is. There are a number of proposals addressing the lack of synchronization between the datalink layer handover and MIP network layer handover. The proposals still base the selection of AP on the signal-to-noise ratio and related factors. These approaches use rather complex signaling between APs, the MH and sometimes entities in the access network. Some approaches propose advertising the bandwidth carried by APs in agent advertisements in order to support MHs running MIP.

MIPv4 propose duplicating packets to two care-of addresses to avoid packet losses during handovers. MIPv6 has no such functionality. However, some research proposals within MIPv6 look at multihoming functionality where only CHs can have multiple bindings for an MH. To achieve that, the MH has to have multiple home addresses.

The approach to multihoming in M-MIP differs from related work in that multihoming is used to discover the care-of address supporting the least loaded AP
based on passive performance measurements and that there is a load balancing of the
HA and CHs between care-of addresses.

Related work with global connectivity spans both IPv4 and IPv6. The major
contribution made by this thesis when compared to related work is the proposal to
use multiple gateways and offers a new approach to gateway selection. In most of the
related work an MH is single homed. Even when multiple gateways are available
only one gateway is used. This gateway is configured to be the default gateway.

The proposed solutions address functionality that is applicable to both IPv4 and
IPv6. Even though most of the work has been carried out on IPv4, IPv6 have been in
mind.

10.3 Future work and challenges

Future work extends to analyzing the developed prototypes using RVM, RNL and M-
MIP in ad hoc subnets and their performance on a daily basis.

Only network layer characteristics have been considered when deciding which AP
to use. Future research will see if an additional evaluation of the wireless link can be
achieved using a combination of RVM and the signal-to-noise ratio. The combination
of RVM and hop count in ad hoc networks will be addressed.

The project has so far considered the 802.11 technology. Research work using
RVM with other wireless local area network (WLAN) technologies is needed. It is
hoped that this will enable usage of heterogeneous WLANs in the near future.
Further research into how RVM can be used in these networks will be carried out.

The proposed M-MIP has been validated in WLANs. However, M-MIP could
prove efficient in heterogeneous networks including GPRS and UMTS as well.
Further research might address comparison of different technologies and the
development of policies enabling users to make an informed decision.

Even though both IPv4 and IPv6 have been considered in this project, the
prototyping is based on IPv4. Mobility in IP networks has support in IPv6 and will
probably not be widely deployed until IPv6 is used. The prototypes, therefore, will
have to be ported to IPv6.

A solution to gradually manage the transition from IPv4 to IPv6 is to use dual
protocol stacks (i.e both IPv4 and IPv6 protocol stacks) in network devices. Network
components can then be modified one by one giving a smooth changeover from IPv4
to IPv6. This should be considered for mobility solutions as well for enabling M-
MIPv4 and M-MIPv6 to coexist, since MHs may connect to access networks running
both IPv4 and IPv6.
References


References


References


References


References


References


References


### Appendix A: Abbreviations

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>AODV</td>
<td>Ad hoc On-Demand Distance Vector Protocol</td>
</tr>
<tr>
<td>ARP</td>
<td>Address Resolution Protocol</td>
</tr>
<tr>
<td>AP</td>
<td>Access Point</td>
</tr>
<tr>
<td>AR</td>
<td>Access Router</td>
</tr>
<tr>
<td>AS</td>
<td>Autonomous System</td>
</tr>
<tr>
<td>BGP</td>
<td>Border Gateway Protocol</td>
</tr>
<tr>
<td>BSS</td>
<td>Basic Service Set</td>
</tr>
<tr>
<td>CH</td>
<td>Correspondent Host</td>
</tr>
<tr>
<td>CSGR</td>
<td>Cluster Switch Gateway Routing</td>
</tr>
<tr>
<td>DHCP</td>
<td>Dynamic Host Configuration Protocol</td>
</tr>
<tr>
<td>DNS</td>
<td>Domain Name Service</td>
</tr>
<tr>
<td>DSDV</td>
<td>Destination-Sequenced Distance Vector protocol</td>
</tr>
<tr>
<td>DSR</td>
<td>Dynamic Source Routing</td>
</tr>
<tr>
<td>FA</td>
<td>Foreign Agent</td>
</tr>
<tr>
<td>HA</td>
<td>Home Agent</td>
</tr>
<tr>
<td>GPRS</td>
<td>General Packet Radio Service</td>
</tr>
<tr>
<td>ICMP</td>
<td>Internet Control Message Protocol</td>
</tr>
<tr>
<td>IGRP</td>
<td>Interior Gateway Routing Protocol</td>
</tr>
<tr>
<td>IP</td>
<td>Internet Protocol</td>
</tr>
<tr>
<td>IPv4</td>
<td>Internet Protocol version 4</td>
</tr>
<tr>
<td>IPv6</td>
<td>Internet Protocol version 6</td>
</tr>
<tr>
<td>IPSec</td>
<td>IP Security Protocol</td>
</tr>
<tr>
<td>ISP</td>
<td>Internet Service Provider</td>
</tr>
<tr>
<td>LAN</td>
<td>Local Area Network</td>
</tr>
<tr>
<td>MAC</td>
<td>Media Access Control</td>
</tr>
<tr>
<td>MH</td>
<td>Mobile Host</td>
</tr>
<tr>
<td>MIP</td>
<td>Mobile Internet Protocol</td>
</tr>
<tr>
<td>MIPv4</td>
<td>Mobile IP for IPv4</td>
</tr>
<tr>
<td>MIPv6</td>
<td>Mobile IP for IPv6</td>
</tr>
<tr>
<td>M-MIP</td>
<td>Multihomed Mobile Internet Protocol</td>
</tr>
<tr>
<td>NBMA</td>
<td>Non Broadcast Multiple Access</td>
</tr>
<tr>
<td>NDP</td>
<td>Neighbour Discovery Protocol</td>
</tr>
<tr>
<td>OSPF</td>
<td>Open Shortest Path First</td>
</tr>
<tr>
<td>PDA</td>
<td>Personal Digital Assistant</td>
</tr>
<tr>
<td>QoS</td>
<td>Quality of Service</td>
</tr>
<tr>
<td>RNL</td>
<td>Relative Network Load</td>
</tr>
<tr>
<td>RVM</td>
<td>Running Variance Metric</td>
</tr>
<tr>
<td>RREP</td>
<td>Route reply</td>
</tr>
<tr>
<td>RREQ</td>
<td>Route request</td>
</tr>
<tr>
<td>Acronym</td>
<td>Full Form</td>
</tr>
<tr>
<td>---------</td>
<td>-----------</td>
</tr>
<tr>
<td>RERR</td>
<td>Route error</td>
</tr>
<tr>
<td>RIP</td>
<td>Routing Information Protocol</td>
</tr>
<tr>
<td>RTT</td>
<td>Round Trip Time</td>
</tr>
<tr>
<td>SNR</td>
<td>Signal-to-Noise Ratio</td>
</tr>
<tr>
<td>TCP</td>
<td>Transmission Control Protocol</td>
</tr>
<tr>
<td>UDP</td>
<td>User Datagram Protocol</td>
</tr>
<tr>
<td>UMTS</td>
<td>Universal Mobile Telecommunications System</td>
</tr>
<tr>
<td>VLAN</td>
<td>Virtual Local Area Network</td>
</tr>
<tr>
<td>VPN</td>
<td>Virtual Private Network</td>
</tr>
<tr>
<td>WLAN</td>
<td>Wireless Local Area Network</td>
</tr>
</tbody>
</table>