M.Sc. Engg. Thesis

Internet Gateway Discovery and Selection Scheme in Mobile Ad Hoc Network

by

Shahid Md. Asif Iqbal

Submitted to

Department of Computer Science and Engineering
in partial fulfillment of the requirements for the degree of
Master of Science in Computer Science and Engineering

Department of Computer Science and Engineering
Bangladesh University of Engineering and Technology (BUET)
Dhaka 1000

May 2011

M.Sc. Engg. Thesis

Internet Gateway Discovery and Selection Scheme in Mobile Ad Hoc Network

by

Shahid Md. Asif Iqbal

Roll No.: 100705010P

Supervised by

Dr. Md. Humayun Kabir

Associate Professor

Department of Computer Science and Engineering

Bangladesh University of Engineering and Technology (BUET)

Department of Computer Science and Engineering

Bangladesh University of Engineering and Technology (BUET)

Dhaka 1000

May 2011

The thesis titled "**Internet Gateway Discovery and Selection Scheme in Mobile Ad Hoc Network**," submitted by Shahid Md. Asif Iqbal, Roll No. 100705010P, Session October 2007, to the Department of Computer Science and Engineering, Bangladesh University of Engineering and Technology, has been accepted as satisfactory in partial fulfillment of the requirements for the degree of Master of Science in Computer Science and Engineering and approved as to its style and contents. Examination held on May 04, 2011

Board of Examiners

1. _____

Name: Dr. Md. Humayun Kabir — Chairman
Address: Associate Professor — (Supervisor)
Department of Computer Science and Engineering
Bangladesh University of Engineering and Technology
Dhaka-1000

2. _____

Name: Dr. Md. Monirul Islam — Member
Address: Professor and Head — (Ex-officio)
Department of Computer Science and Engineering
Bangladesh University of Engineering and Technology
Dhaka-1000

3. _____

Name: Dr. Reaz Ahmed — Member
Address: Associate Professor
Department of Computer Science and Engineering
Bangladesh University of Engineering and Technology
Dhaka-1000

4. _____

Name: Dr. Syed Faisal Hasan — Member
Address: Assistant Professor — (External)
Department of Computer Science and Engineering
Dhaka University
Dhaka-1000

Candidate's Declaration

It is hereby declared that this thesis or any part of it has not been submitted elsewhere for the award of any degree or diploma.

_____

Shahid Md. Asif Iqbal
Roll. No. 100705010P
(Candidate)

Contents

List of Figures

List of Tables

List of Algorithms

Abstract

Global connectivity to Mobile Ad Hoc Network (MANET) is necessary to access the Internet services from the MANET. Nodes in a MANET that connect it to the Internet are called Internet gateways. Internet gateways need to be discovered and selected in an appropriate way to deliver more packets to the Internet and reduce end-to-end delay. Currently, there are proactive, reactive, and hybrid schemes to discover and select Internet gateways in MANET. However, these schemes do not scale well with the number of nodes, traffic load, and speed of the nodes in MANET. To make it scalable, we proposed a new gateway discovery and selection scheme. In our scheme, the gateways advertise gateway advertisement messages only on-demand. Moreover, it contains the advertisements within a limit in order to make our scheme scalable. We also considered the interface queue length and the total number of neighbors along a route in addition to the hop count to bypass the loaded and dense route to the gateways in order to reduce the delay and packet loss. Simulation results show that our scheme scales well with the number of nodes, traffic load and the speed of the nodes in MAENT compared to that of other schemes. It also confirms that our scheme has less delay and packets drop than that of other schemes.

Acknowledgement

I would like to express my gratitude to my thesis supervisor, Dr. Md. Humayun Kabir, for his intellectual assistance, continual encouragement, and valuable guidance throughout the work. His invaluable feedback and critical analysis of my thoughts helped me a lot to improve the excellence of the work.

I would also like to express my thanks to the members of my thesis committee, Professor Dr. Md. Monirul Islam, Dr. Reaz Ahmed, and Dr. Syed Faisal Hasan, for their valuable suggestions.

Moreover, I wish to thank my colleague Mohammad Iftekhar Monir for giving feedback on the report.

I also want to thank my wife for her patience and support, particularly during the less exciting stage of my studies.

Mere words are insufficient to express my gratitude to my loving parents, for their love and outstanding support, without which this would not have been possible.

Chapter 1

Introduction

A Mobile Ad Hoc Network (MANET) is formed by a group of mobile nodes without the aid of any centralized administration or established infrastructure. A pair of mobile nodes may communicate with each other either directly or indirectly with the help of the intermediate nodes. Since these kinds of networks are very spontaneous and self-organizing, many useful applications such as multimedia streaming, collaborative work, information dissemination and jungle telemetry can be supported by these networks and that's why they are very demanding in commercial arena specially in the emergency services like hospitals, ambulance, police and military applications etc.

1.1 MANET and Internet

In future, the Internet is likely to be different from its present state because mobile devices with various computational resources will dominate it. Wireless communication technology and the Internet are developing so quickly that there are numerous mobile devices around us and multiple wireless networks are serving these mobile devices all the time. A MANET is generally considered as a stand-alone network i.e. communication is only supported among the nodes within the ad hoc domain. This stand-alone nature limits the applicability of MANET to the scenarios those require external connectivity. Integration of MANET and Internet can provide global connectivity to MANET so that it no longer remains stand-alone. This integration allows mobile users in MANET to access

the popular Internet applications such as e-mail, chat, instant messaging, file transfer etc. The integration expands both MANET and the Internet coverage range.

Integration of MANET with the Internet has recently become an active research area. To access the Internet from a MANET a subset of its nodes must have the interfaces to connect to the Internet directly. These nodes work as the Internet gateway, which facilitates other nodes to communicate outside the MANET. There might be multiple Internet gateways in a MANET. A mobile node in a MANET may be multi-hop away from the Internet gateways. In this case, the node has to use the Internet gateway through other intermediate nodes.

1.2 Background and the Problems

When a mobile node in a MANET wants to access the Internet, it needs to discover the available Internet gateways and selects the best one among them if multiple gateways are found. Therefore, it needs an efficient Internet gateway discovery and selection scheme that achieves high throughput, low delay and less network overhead. Two types of schemes, reactive and proactive, have been proposed to discover and select Internet gateways in MANET. In proactive schemes [1-9], Internet gateways periodically broadcast gateway advertisement messages in the MANET. Each node that receives the advertisement message forwards the advertisement to other nodes until the message is flooded over the whole network. These schemes cost heavy routing load since the gateway advertisements are broadcasted periodically throughout the entire ad hoc network even if there is no such demand from the nodes in the MANET. However, the proactive schemes are blessed with higher rate of successful delivery and lower delay. In reactive schemes [1-3] [9-12], a mobile node broadcasts a gateway discovery message to discover Internet gateways in the network. Whenever a gateway receives the discovery message, it unicasts a gateway advertisement message back to the requestor. These schemes suffer from higher delay and lower packet delivery ratio since the nodes have to send a gateway discovery message every time they need a gateway. Reactive schemes scale poorly regarding the number of sources willing to access the Internet. Few research

works [9] [13-19] proposed hybrid gateway discovery schemes where the dissemination of gateway advertisements is kept limited to a small area by setting appropriate Time to Live (TTL). Nodes outside the TTL coverage area reactively find their gateways. The performance of these schemes degrades if TTL is not adapted properly. Most of the existing hybrid schemes [9] [13-14] [16] do not adjust TTL value dynamically.

Gateway selection scheme selects the best gateway when it receives multiple gateway advertisements from multiple gateways. Gateway selection schemes proposed in [1-3] [5] [7] [9] [13-18] use hop count only to select a gateway. In these schemes, all the nodes always select the nearest gateway, a gateway may become a bottleneck under heavy traffic load and there is no remedy for this problem.

1.3 Solutions

To deal with the problems in existing Internet gateway discovery and selection schemes we propose a new hybrid gateway discovery scheme where gateways will act reactively, however, broadcast a gateway advertisement message when they receive a gateway discovery message from a mobile node. The TTL of the gateway advertisement message will also be set to a value equal to the distance of the gateway from the requestor. Each mobile node will configure its gateway after receiving the gateway advertisement message. In our scheme, a node selects a gateway that promises optimal performance, after receiving the advertisement messages from multiple gateways. While selecting the best gateway, the node will consider the interface queue size and the total number of neighbors of each node along the route in addition to the hop count. We consider the number of packets waiting in the interface queue of a mobile node as its interface queue size. The use of interface queue size in the selection of a gateway, allows us to redirect a mobile node from a heavily loaded gateway to a less loaded one and the inclusion of total number of neighbors of each node helps us to avoid a crowded area to reach the gateway.

1.4 Deposition

The rest of the thesis has been organized as follows. In Chapter 2 we describe Mobile Ad Hoc Network basics briefly. Chapter 3 reviews the current solutions for Internet gateway discovery and selection in MANET. Chapter 4 depicts our new hybrid gateway discovery scheme. We also introduce the new metric used in the gateway selection scheme in this chapter. Simulation setup and analysis of simulation results comes in Chapter 5 and finally in Chapter 6 we conclude our thesis with some future research guidelines.

Chapter 2

Mobile Ad Hoc Network Basics

In this chapter, we describe Mobile Ad Hoc Network (MANET) and some essential properties that are required to connect MANET to the Internet. Section 2.1 gives an overview of MANET, the protocol stack used in MANET, its applications and characteristics. In Section 2.2, we discuss the IEEE 802.11 protocol in brief. We talk about various routing protocols developed for ad hoc networks in Section 2.3. Section 2.4 describes some basic operations needed to access the Internet from MANET.

2.1 Mobile Ad Hoc Network

A Mobile Ad Hoc Network is an autonomous collection of mobile nodes connected by wireless links. Unlike the fixed networks, Mobile Ad Hoc Networks are characterized by the lack of infrastructure. Each node operates as an end device as well as a router for all other nodes in the network. Figure 2.1 shows a Mobile Ad Hoc Network.

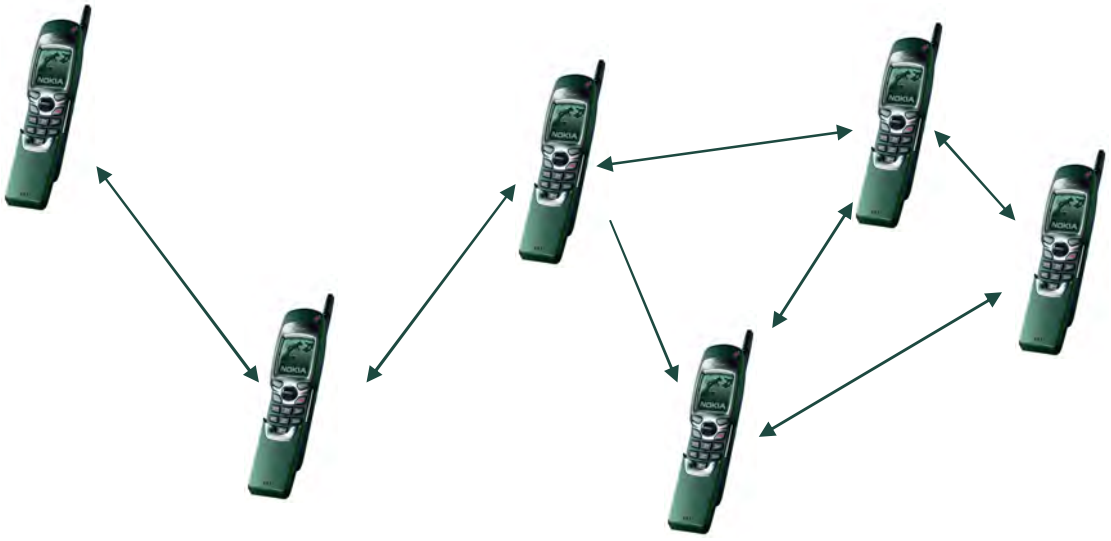


Figure 2.1: A Mobile Ad Hoc Network

Nodes in a MANET are free to move and organize themselves in an arbitrary fashion. Therefore, such networks have dynamic, rapidly-changing, and multihop network topologies. Also the network is decentralized, therefore the network activities like discovering the network topology, forwarding the data packets, dissemination of the routing messages must be executed by the nodes themselves, i.e., routing functionality is incorporated into the mobile nodes. Mobile Ad Hoc Networks can operate in a stand-alone fashion or could possibly be connected to a larger network such as the Internet.

2.1.1 The Protocol Stack

In this section the protocol stack for mobile ad hoc networks is described. Figure 2.2 shows the OSI model, TCP/IP suite and MANET protocol stack. The MANET protocol stack consists of five layers: physical layer, data link layer, network layer, transport layer and application layer.

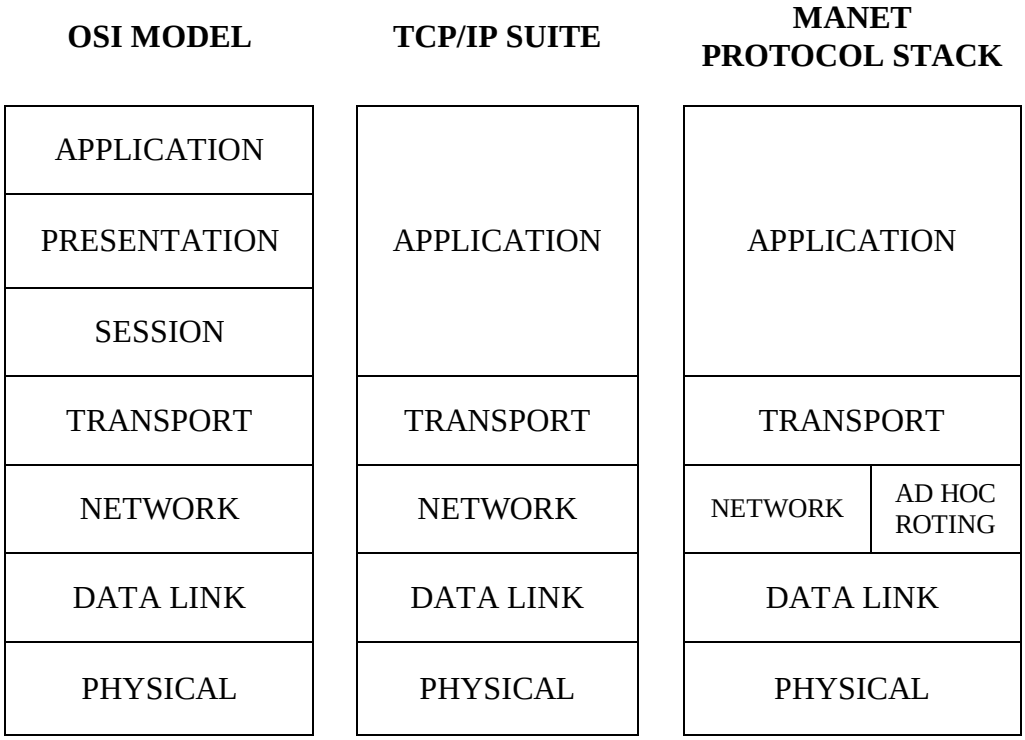


Figure 2.2: The OSI model, TCP/IP suite and MANET protocol stack.

Session, presentation and application layers of OSI model are merged into the application layer of MANET. MANET protocol stack is similar to the TCP/IP suite. The main difference between these two protocol stacks lies in the network layer. Mobile nodes (which are both hosts and routers) use an ad hoc routing protocol to route packets. In the physical and data link layers, mobile nodes run protocols that have been designed for wireless channels. Some options are the IEEE standard for wireless LANs, IEEE 802.11, the European ETSI standard for a high-speed wireless LAN, HIPERLAN 2, and finally an industry approach toward wireless personal area networks, i.e. wireless LANs at an even smaller range, Bluetooth [9]. The network layer is divided into two parts: Network and Ad Hoc Routing. The protocol used in the network part is Internet Protocol (IP) and the protocol used in the ad hoc routing part is Ad hoc On-Demand Distance Vector (AODV) [20]. Other ad hoc routing protocols that can be used in this part of the network layer are discussed in Sections 2.3. One of the reasons to why AODV has been selected for this layer is that it is one of the most developed routing protocols for mobile

ad hoc networks. A second reason is that it is easy to extend the AODV routing protocol to implement Internet gateway discovery and selection schemes.

In the transport layer, we consider the User Datagram Protocol (UDP) for this thesis. The Transmission Control Protocol (TCP) is not used because TCP does not perform well in mobile ad hoc networks [9]. One reason to this is that in wired networks, lost packets are almost always due to congestion but in mobile ad hoc networks lost packets are more often caused by other reasons like route changes or transmission errors.

2.1.2 Applications of MANETs

Ad Hoc networks can be operated as robust, inexpensive alternatives or enhancements to cell-based mobile network infrastructures. One of many possible uses of Mobile Ad Hoc Networks is in some business environments, where the need for collaborative computing might be more important outside the office environment than inside, such as in a business meeting outside the office to brief clients on a given assignment.

When properly combined with satellite-based information delivery, MANET technology can provide an extremely flexible method for establishing communications for fire/safety/rescue operations or other scenarios requiring rapidly-deployable communications with survivable and efficient dynamic networking.

Another application example of a mobile ad-hoc network is Bluetooth, which is designed to support a personal area network by eliminating the need of wires between various devices, such as printers and personal digital assistants.

2.1.3 Characteristics of MANETs

A MANET consists of mobile nodes which are permitted to move freely. The nodes may be located in or on airplanes, ships, trucks, cars, perhaps even on people or very small devices. A MANET is an autonomous system of mobile nodes. The system may operate in isolation, or may have gateways to and interface with a fixed network. In the latter operational mode, it is typically envisioned to operate as a "stub" network connecting to a

fixed internetwork. Stub networks carry traffic originating from or destined to internal nodes, but do not permit outside traffic to "transit" through the stub network.

MANET nodes are equipped with wireless transmitters and receivers using antennas which may be omni-directional (broadcast), highly-directional (point-to-point), or some combination thereof. At a given point in time, depending on the nodes' positions and their transmitter and receiver coverage patterns, transmission power levels and co-channel interference levels, a wireless connectivity in the form of a random and multihop graph exists among the nodes. This ad hoc topology may change with time as the nodes move or adjust their transmission and reception parameters.

MANETs have several salient characteristics:

Dynamic topologies Nodes are free to move arbitrarily; thus, the network topology which is typically multihop may change randomly and rapidly at unpredictable times, and may consist of both bidirectional and unidirectional links.

Bandwidth-constrained Wireless links will continue to have significantly lower capacity than their wired counterparts. In addition, the realized throughput of wireless communications after accounting for the effects of multiple access, fading, noise, and interference conditions, etc. is often much less than a radio's maximum transmission rate. One effect of the relatively low link capacities is that congestion is typically the norm rather than the exception, i.e. aggregate application demand will likely approach or exceed network capacity frequently. As the mobile network is often simply an extension of the fixed network infrastructure, mobile ad hoc users will demand similar services. These demands will continue to increase as multimedia computing and collaborative networking applications rise.

Energy-constrained operation Some or all of the nodes in a MANET may rely on batteries or other exhaustible means for their energy. For these nodes, the most important system design criteria for optimization may be energy conservation.

Limited physical security Mobile wireless networks are generally more prone to physical security threats than are fixed-cable nets. The increased possibility of eavesdropping, spoofing, and denial-of-service attacks should be carefully considered. Existing link security techniques are often applied within wireless networks to reduce security threats. As a benefit, the decentralized nature of

network control in MANETs provides additional robustness against the single points of failure of more centralized approaches.

These characteristics create a set of underlying assumptions and performance concerns for protocol design with MANET which extends beyond those guiding the design of routing within the higher-speed, semi-static topology of the fixed Internet.

2.2 IEEE 802.11 Standard

IEEE 802.11 standard provides physical (PHY) and MAC layer solutions for wireless local area networks. With the popularity of IEEE 802.11 standard family used in computers, laptops, and Personal Digital Assistants (PDAs), this standard is considered to be one of the solutions used in ad hoc networks. Especially in the simulations, IEEE 802.11 standard is preferred in ad hoc networks by most of the people because of its availability and easiness.

2.2.1 IEEE 802 Family

IEEE 802 specifications focus on the data link layer and physical layer of the Open System Interconnection (OSI) reference model. Some of the main family members of IEEE 802 are listed in Table 2-1.

Table 2.1: IEEE 802 Family

IEEE Standard	**Network Definition**	**Known as**
802.3	Wired Local Area Network	Ethernet
802.11	Wireless Local Area Network(WLAN)	WiFi
802.15.1	Wireless Personal Area Network(WPAN)	Bluetooth
802.15.4	Low Rate-Wireless Personal Area Network	Zigbee
802.16	Wireless Metropolitan Area Network	WiMax
802.20	Mobile Broadband Wireless Access (MBWA)	

2.2.2 IEEE 802.11 Family

IEEE 802.11 specification can be divided into two parts, which are 802.11 MAC and 802.11 PHY [21]. Part of the IEEE 802.11 family members are shown in Table 2-2.

Table 2.2: IEEE 802.11 Family

IEEE 802.11 MAC	IEEE 802.11 PHY				
802.11 Medium Access Control (CSMA/CA)	PHY Type	802.11 PHY (FHSS/DSSS)	802.11a PHY (OFDM)	802.11b PHY (DSSS)	802.11g PHY (OFDM)
	Data Rate	1, 2 Mbps	6, 9, 12, 18, 24, 36, 48, 54 Mbps	1, 2, 5.5, 11 Mbps	6, 9, 12, 18, 24, 36, 48, 54 Mbps
	Operating Frequency	2.4 GHz	5 GHz	2.4 GHz	2.4 GHz

802.11 PHY has a few physical layer designs. It includes Frequency-Hopping Spread-Spectrum (FHSS) PHY and Direct-Sequence Spread-Spectrum (DSSS) PHY in IEEE 802.11. Later versions of PHY layer schemes are orthogonal frequency division multiplexing (OFDM) PHY (specified in IEEE 802.11a) and High-Rate Direct-Sequence Spread Spectrum (HR/DSSS) PHY (specified in 802.11b). OFDM used in IEEE 802.11a helps to improve the data rate up to 54 Mbps. IEEE 802.11b is a very popular standard used in mobile wireless networks and its products hit the market in 1999. It is used widely in WLAN. IEEE 802.11 MAC is used to access to the mobile network. It follows Carrier Sensing Multiple Access/Collision Avoidance (CSMA/CA) mechanism with the random back-off mechanism.

2.2.3 Basic Service Set in IEEE 802.11

The Basic Service Set (BSS) is the basic building block of 802.11 networks. It is composed of several stations which could communication with each other. The area in which they can communicate is called the basic service area. There are basically two

configuration modes provided by IEEE 802.11 for the BSS. They are independent BSS and infrastructure BSS. The two configuration modes are shown in Figure 2.3.

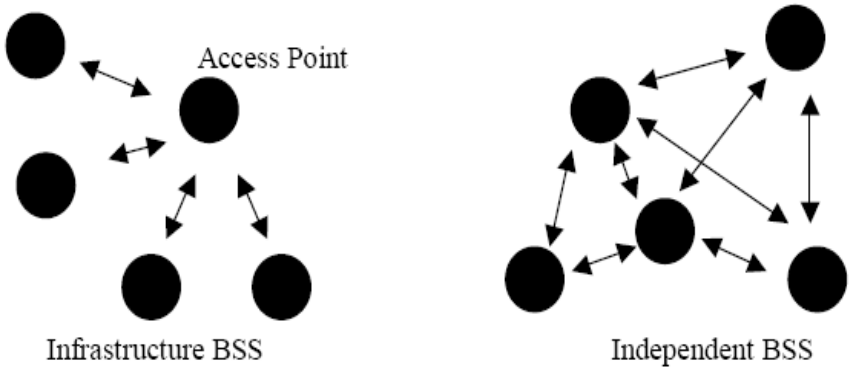


Figure 2.3: Two configuration modes in IEEE 802.11.

In the independent BSS, stations can communicate directly with each other when they are in each others transmission range. It is always called ad hoc networks. This kind of network is used when there is temporary need for wireless network between stations. The advantage of using ad hoc networks is that there is no need of infrastructure during the set up of the network.

In the infrastructure BSS, there is an access point in each BSS. Stations communicate with each other through the access point. That is, mobile station should first transmit the frames to the access point, and it is the responsibility of the access point to transmit those frames to the destination station. The transmission range of the access point is the radius of the service area of this wireless network. Because of this, the destination station does not need to be in the transmission range of the source station, but only need to be in the transmission range of the access point. There is no restriction to the distance between the source and destination station.

2.3 Routing Protocols in MANET

Routing protocols in ad hoc networks vary depending on the type of the network. Typically, ad hoc network routing protocols are classified into three major categories based on the routing information updated mechanism. They are proactive (table driven

routing protocols), reactive (on-demand routing protocols) and hybrid routing protocols. Both proactive and reactive routing protocols have specific advantages and disadvantages that make them suitable for certain types of scenarios. Proactive routing protocols have their routing tables updated at all the times, thus the delay before sending a packet is minimal.

However, routing tables that are always updated require periodic control messages that are flooded through the whole network - an operation that consumes a lot of time, bandwidth and energy. On the other hand, reactive routing protocols determine routes between nodes only when they are explicitly needed to route packets. However, whenever there is a need for sending a packet, the mobile node must first find the route if the route is not already known. This route discovery process may result in considerable delay.

2.3.1 Proactive Routing Protocols

Traditional distance-vector and link-state routing protocols are proactive in that they maintain routes to all nodes, including nodes to which no packets are sent. For that reason they require periodic control messages, which lead to scarcity of resources such as power and link bandwidth being used more frequently for control traffic as mobility increases. Examples of proactive routing protocols are Optimized Link State (OLSR) Routing Protocol [22], and Destination-Sequenced Distance-Vector (DSDV) [23] routing protocol.

2.3.1.1 OLSR Routing Protocol

OLSR is a proactive routing protocol for mobile ad hoc networks. The protocol inherits the stability of a link state algorithm and has the advantage of having routes immediately available when needed due to its proactive nature. OLSR is an optimization over the classical link state protocol, tailored for mobile ad hoc networks.

OLSR minimizes the overhead from flooding of control traffic by using only selected nodes, called MPRs (Multipoint Relays), to retransmit control messages. This technique significantly reduces the number of retransmissions required to flood a message to all

nodes in the network. Secondly, OLSR requires only partial link state to be flooded in order to provide shortest path routes. The minimal set of link state information required is, that all nodes, selected as MPRs, must declare the links to their MPR selectors. Additional topological information, if present, may be utilized e.g., for redundancy purposes.

OLSR may optimize the reactivity to topological changes by reducing the maximum time interval for periodic control message transmission.

Furthermore, as OLSR continuously maintains routes to all destinations in the network, the protocol is beneficial for traffic patterns where a large subset of nodes are communicating with another large subset of nodes, and where the [source, destination] pairs are changing over time. The protocol is particularly suited for large and dense networks, as the optimization done by using MPRs works well in this context. The larger and more dense a network, the more optimization can be achieved as compared to the classic link state algorithm.

OLSR is designed to work in a completely distributed manner and does not depend on any central entity. The protocol does not require reliable transmission of control messages: each node sends control messages periodically, and can therefore sustain a reasonable loss of some such messages. Such losses occur frequently in radio networks due to collisions or other transmission problems.

Also, OLSR does not require sequenced delivery of messages. Each control message contains a sequence number which is incremented for each message. Thus the recipient of a control message can, if required, easily identify which information is more recent - even if messages have been re-ordered while in transmission.

Furthermore, OLSR provides support for protocol extensions such as sleep mode operation, multicast-routing etc. Such extensions may be introduced as additions to the protocol without breaking backwards compatibility with earlier versions. OLSR does not require any changes to the format of IP packets. Thus any existing IP stack can be used as is: the protocol only interacts with routing table management.

Multipoint Relays

The idea of multipoint relays is to minimize the overhead of flooding messages in the network by reducing redundant retransmissions in the same region. Each node in the network selects a set of nodes in its symmetric 1-hop neighborhood which may retransmit its messages. This set of selected neighbor nodes is called the "Multipoint Relay" (MPR) set of that node. The neighbors of node N which are not in its MPR set, receive and process broadcast messages but do not retransmit broadcast messages received from node N.

Each node selects its MPR set from among its 1-hop symmetric neighbors. This set is selected such that it covers (in terms of radio range) all symmetric strict 2-hop nodes. The MPR set of N, denoted as MPR(N), is then an arbitrary subset of the symmetric 1-hop neighborhood of N which satisfies the following condition:

- Every node in the symmetric strict 2-hop neighborhood of N must have a symmetric link towards MPR(N). The smaller a MPR set, the less control traffic overhead results from the routing protocol.

Each node maintains information about the set of neighbors that have selected it as MPR. This set is called the "Multipoint Relay Selector set" (MPR selector set) of a node. A node obtains this information from periodic HELLO messages received from the neighbors.

A broadcast message, intended to be diffused in the whole network, coming from any of the MPR selectors of node N is assumed to be retransmitted by node N, if N has not received it yet. This set can change over time (i.e., when a node selects another MPR-set) and is indicated by the selector nodes in their HELLO messages.

2.3.1.2 DSDV Routing Protocol

Destination sequenced distance vector routing is adapted from the conventional Routing Information Protocol (RIP) to ad hoc networks routing. It adds a new attribute, sequence number, to each route table entry of the conventional RIP. Using the newly added sequence number, the mobile nodes can distinguish stale route information from the new and thus prevent the formation of routing loops.

Each node maintains a list of all destinations and number of hops to each destination. Each entry is marked with a sequence number. It uses full dump or incremental update to reduce network traffic generated by route updates. The broadcast of route updates is delayed by settling time. The only improvement made here is the avoidance of routing loops in a mobile network of routers. With this improvement, routing information can always be readily available, regardless of whether the source node requires the information or not. DSDV solves the problem of routing loops and count to infinity by associating each route entry with a sequence number indicating its freshness. In DSDV, a sequence number is linked to a destination node, and usually is originated by that node (the owner). The only case that a non-owner node updates a sequence number of a route is when it detects a link break on that route. An owner node always uses even-numbers as sequence numbers, and a non-owner node always uses odd-numbers. With the addition of sequence numbers, routes for the same destination are selected based on the following rules:

- A route with a newer sequence number is preferred
- In the case that two routes have a same sequence number, the one with a better cost metric is preferred.

The routing tables of the mobile nodes contain the following fields:

- All available destinations' IP address
- Next hop IP address
- Number of hops to reach the destination
- Sequence number assigned by the destination node
- Install time

The sequence number is used to distinguish stale routes from new ones and thus avoid the formation of loops. The stations periodically transmit their routing tables to their immediate neighbors. A station also transmits its routing table if a significant change has occurred in its table from the last update sent. So, the update is both time-driven and event-driven.

As stated above one of "full dump" or an "incremental update" is used to send routing table updates for reducing network traffic. A full dump sends the full routing table to the

neighbors and could span many packets whereas in an incremental update only those entries from the routing table are sent that has a metric change since the last update and it must fit in a packet. If there is space in the incremental update packet then those entries may be included whose sequence number has changed. When the network is relatively stable, incremental updates are sent to avoid extra traffic and full dump are relatively infrequent. In a fast-changing network, incremental packets can grow big so full dumps will be more frequent. Each routing update packet contains a unique sequence number assigned by the transmitter in addition to the routing table information. The route labeled with the highest (i.e. most recent) sequence number is used. If two routes have the same sequence number then the route with the best metric (i.e. shortest route) is used.

Based on the past history, the stations estimate the settling time of routes. The stations delay the transmission of a routing update by settling time to eliminate those updates that would have not been occurred if a better route were found very soon. Each row of the update send is of the following form:

<Destination IP address, Destination sequence number, Hop count>

After receiving an update neighboring nodes utilizes it to compute the routing table entries.

To damp the routing fluctuations due to unsynchronized nature of periodic updates, routing updates for a given destination can propagate along different paths at different rates. To prevent a node from announcing a routing path change for a given destination while another better update for that destination is still in route, DSDV requires node to wait for a settling time before announcing a new route with higher metric for a destination.

2.3.2 Reactive Routing Protocols

Reactive routing protocols operate only when there is a need of communication between two nodes. This approach allows the nodes to focus either on routes that are being used or on routes that are in process of being set up. Examples of reactive routing protocols are Ad hoc On-Demand Distance Vector (AODV) [20], and Dynamic Source Routing (DSR) [24].

2.3.2.1 AODV Routing Protocol

AODV is a distance vector routing protocol that operates on-demand. There are no periodic exchanges of routing tables; routes are only set up when a node wants to communicate with some other node.

Route Discovery

Whenever a mobile node wishes to communicate with a destination for which it has no routing information, it initiates route discovery by flooding the network with a route request (RREQ) message. The aim of route discovery is to set up a bidirectional route from source to the destination. Each node that receives the RREQ message looks in its routing table to see whether it is the destination or it has a fresh enough route to the destination. If it is the destination or it finds a route to the destination, it responds by sending route a reply (RREP) message back to the requesting node; otherwise it rebroadcasts the RREQ message. The RREP message is routed back on a reverse route that was created by the RREQ. Once the requestor receives the RREP, it can start using the route for data transmission. Each node maintains a routing table containing one route entry for each destination that the node is communicating. The route discovery process is depicted in Figure 2.4.

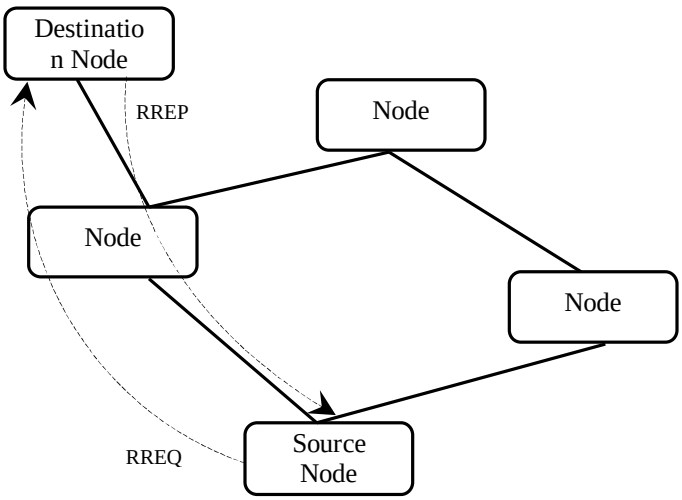


Figure 2.4: Ad Hoc On-Demand Distance Vector Routing Protocol

Message Formats

Route Request (RREQ) Message Format The format of the Route Request message is illustrated in Figure 2.5.

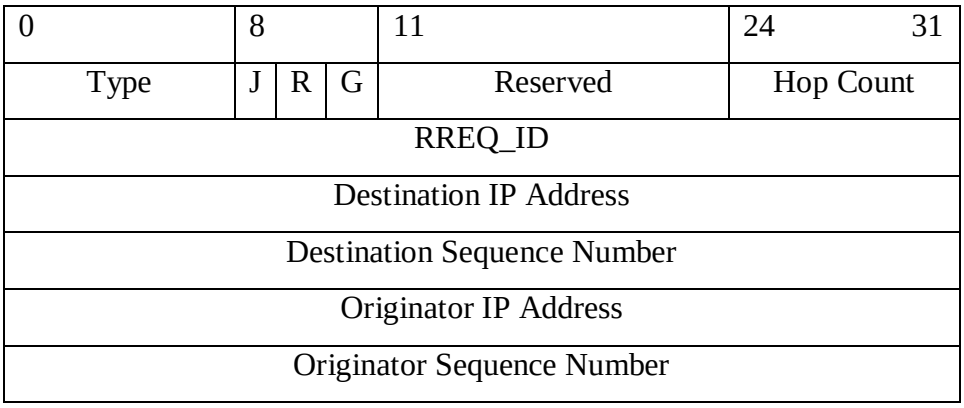


Figure 2.5: Route Request (RREQ) Message Format

RREQ message contains the following fields:

- Type: 1
- J: Join flag; reserved for multicast.
- R: Repair flag; reserved for multicast.
- G: Gratuitous RREP flag; indicates whether a gratuitous RREP should be unicast to the node specified in the Destination IP Address field.
- Reserved: Sent as 0; ignored on reception.
- Hop Count: The number of hops from the Originator IP Address to the node handling the request.
- RREQ ID: A sequence number uniquely identifying the particular RREQ when taken in conjunction with the originating node's IP address.
- Destination IP Address: The IP address of the destination for which a route is desired.
- Destination Sequence Number: The latest sequence number received in the past by the originator for any route towards the destination.
- Originator IP Address: The IP address of the node which originated the Route Request.

- Originator Sequence Number: The current sequence number to be used in the route entry pointing towards the originator of the route request.

Route Reply (RREP) Message Format The format of the Route Reply message is illustrated in Figure 2.6

<table>
<tr><td>0</td><td colspan="2">8</td><td>10</td><td>19</td><td>24 31</td></tr>
<tr><td>Type</td><td>R</td><td>A</td><td>Reserved</td><td>Pref. Sz.</td><td>Hop Count</td></tr>
<tr><td colspan="6">Destination IP Address</td></tr>
<tr><td colspan="6">Destination Sequence Number</td></tr>
<tr><td colspan="6">Originator IP Address</td></tr>
<tr><td colspan="6">Lifetime</td></tr>
</table>

Figure 2.6: Route Reply (RREP) Message Format

RREP message contains the following fields:

- Type: 2
- R: Repair flag; used for multicast.
- A: Acknowledgment required.
- Reserved: Sent as 0; ignored on reception.
- Prefix Size: If nonzero, the 5-bit Prefix Size specifies that the indicated next hop may be used for any nodes with the same routing prefix (as defined by the Prefix Size) as the requested destination.
- Hop Count: The number of hops from the Originator IP Address to the Destination IP Address. For multicast route requests this indicates the number of hops to the multicast tree member sending the RREP.
- Destination IP Address: The IP address of the destination for which a route is supplied.
- Destination Sequence Number: The destination sequence number associated to the route.
- Originator IP Address: The IP address of the node which originated the RREQ for which the route is supplied.

- Lifetime: The time in milliseconds for which nodes receiving the RREP consider the route to be valid.

Route Maintenance

Routes in AODV are considered active as long as they are in use. If a route is no longer in use, it will expire and eventually be expunged from the routing table. Therefore, the routing table maintained at each node does not often supply a stale route when the node needs to set up a connection with a node in the MANET. Rather the node initiates the route discovery process instead of using a stale route.

Each node keeps track of its local connectivity, i.e. its neighbors. This is performed by using periodic exchange of Hello messages, or by using feedback from the link-layer upon unsuccessful transmission. Route maintenance in AODV makes use of route error (RERR) messages. When a link breaks along an active route, the node upstream of the break sends a RERR message to each neighbor that was using that link to reach the destination. The RERR message lists each destination that is now unreachable owing to the loss of the link.

When a source node receives a RERR, it may initiate a route discovery again if it still needs the route. AODV guarantees loop-free routes by using sequence numbers that indicate how fresh a route is. Each route entry keeps track of certain fields. Some of these fields are:

- Destination IP Address: The IP address of the destination for which the route is kept
- Destination Sequence Number: The destination sequence number associated to the route
- Next Hop: Either the destination itself or an intermediate node designated to forward packets to the destination
- Hop Count: The number of hops from the Originator IP Address to the Destination IP Address
- Lifetime: The time span for which the route to the destination is considered to be valid
- Routing Flags: The state of the route; up (valid), down (not valid) or in repair.

2.3.2.2 DSR Routing Protocol

Dynamic Source Routing, DSR, is a reactive routing protocol that uses *source routing* to send packets [24]. It is reactive like AODV which means that it only requests a route when it needs one and does not require to maintaining routes to the destinations that are not in use. It uses source routing which means that the source must know the complete hop sequence to the destination in order to send packets to it. Each node maintains a route cache, where all routes it knows are stored. The route discovery process is initiated only if the desired route cannot be found in the route cache.

To limit the number of route request propagations, a node processes the route request message only if it has not already received the message and its address is not present in the route record of the message.

As mentioned before, DSR uses source routing, i.e. the source determines the complete sequence of hops that each packet should traverse. This requires that the sequence of hops is included in each packet's header. A negative consequence of this is the routing overhead every packet has to carry. However, one big advantage is that intermediate nodes can learn routes from the source routes in the packets they receive. Since finding a route is generally a costly operation in terms of time, bandwidth and energy, this is a strong argument for using source routing. Another advantage of source routing is that it avoids the need for up-to-date routing information in the intermediate nodes through which the packets are forwarded since all necessary routing information is included in the packets. Finally, it avoids routing loops easily because the complete route is determined by a single node instead of making the decision hop-by-hop.

Route Discovery

Route Discovery is used whenever a source node desires a route to a destination node. First, the source node looks up in its route cache to determine if it already contains a route to the destination. If the source finds a valid route to the destination, it uses this route to send its data packets. If the node does not have a valid route to the destination, it initiates the route discovery process by broadcasting a *route request* message. The route request message contains the address of the source and the destination, and a unique identification number.

An intermediate node that receives a route request message searches its route cache for a route to the destination. If no route is found, it appends its address to the route record of the message and forwards the message to its neighbors. The message propagates through the network until it reaches either the destination or an intermediate node with a route to the destination. Then a *route reply* message, containing the proper hop sequence for reaching the destination, is generated and unicast back to the source node.

Route Maintenance

Route Maintenance is used to handle route breaking ups. When a node encounters a fatal transmission problem at its data link layer, it removes the route from its route cache and generates a *route error* message. The route error message is sent to each node that has sent a packet routed over the broken link. When a node receives a route error message, it removes the hop in error from its route cache. Acknowledgment messages are used to verify the correct operation of the route links.

In wireless networks acknowledgments are often provided as the link-layer acknowledgment frame defined by IEEE 802.11. If a built-in acknowledgment mechanism is not available, the node transmitting the message can explicitly request a DSR-specific software acknowledgment to be returned by the next node along the route.

2.3.3 Hybrid routing Protocol

Hybrid routing protocol is the combination of the proactive and reactive routing protocols. A hybrid routing protocol not only minimizes the disadvantages, but also takes the advantages of the proactive and reactive routing protocols. Here, each mobile node proactively maintains routes within a local region (referred to as the routing zone). Mobile nodes residing outside the zone can be reached with reactive routing.

2.3.3.1 Zone Routing Protocol (ZRP)

Zone Routing Protocol, ZRP [25], is a routing protocol that is designed for mobile ad hoc networks . It is a hybrid protocol which is divided in two parts: proactive and reactive. The proactive part uses a modified distance vector scheme within the *routing zone* of each node. The routing zone is determined by a *zone radius*, which is the minimum number of hops it should take to get to any node. Thus, each node has a routing zone, which is composed of nodes within its local area. This proactive component is called *Intrazone Routing Protocol* (IARP). The reactive component is called *Interzone Routing Protocol* (IERP), and uses queries to get routes when a node is to send a packet to a node outside of its routing zone.

ZRP uses a method called *bordercasting* in which a node asks all nodes on the border of its routing zone to look for the node outside of its routing zone.

Intrazone Routing Protocol (IARP)

The Intrazone Routing Protocol (IARP) proactively maintains routes to destinations within a local neighborhood, which is referred to as a routing zone. More precisely, a node's routing zone is defined as a collection of nodes whose minimum distance in hops from the node in question is no greater than a parameter referred to as the zone radius. Note that each node maintains its own routing zone. An important consequence is that the routing zones of neighboring nodes overlap.

Interzone Routing Protocol (IERP)

The operation of the reactive Interzone Routing Protocol (IERP) is quite similar to standard route discovery process of reactive routing protocols. An IERP route discovery is initiated when no route is locally available to the destination of an outgoing data packet. The source generates a route query message, which is uniquely identified by a combination of the source node's address and request number. The query is then relayed to a subset of neighbors as determined by the bordercast algorithm. Upon receipt of a route query message, a node checks if the destination lies in its zone or if a valid route to

it is available in its route cache. If the destination is found, a route reply is sent back to the source. If not, the node bordercasts the query again.

Bordercast Resolution Protocol (BRP)

Since the topology of the local zone of each mobile node is known (this information is provided by IARP), global route discovery is simplified. Rather than broadcasting a route query from neighbor to neighbor, ZRP uses a concept called *bordercasting*. Bordercasting means that the route query is directed toward regions of the network that have not yet been covered by the query. A covered node is the one that belongs to the routing zone of a node that has received a route query. Hence, the route query traffic is reduced by directing route queries outwards from the source and away from covered routing zones.

2.4 Internet Connectivity to MANET

Although an autonomous, stand-alone mobile ad hoc network is useful in many cases, a mobile ad hoc network connected to the Internet is much more desirable. So far, most of the research concerning mobile ad hoc networking has been done on protocols for autonomous mobile ad hoc networks. However, during the last decade, some works have been done concerning the integration of mobile ad hoc networks and the Internet.

To achieve this network interconnection, gateways that understand the protocols of both the mobile ad hoc network stack and the TCP/IP suite are needed.

Whenever a mobile node is to send packets to the Internet, it must forward the packets to a gateway. Thus, all communication between a mobile ad hoc network and the Internet must pass through the gateways.

The protocol stacks involved during communication between a MANET node and the Internet node are shown in Figure 2.7.

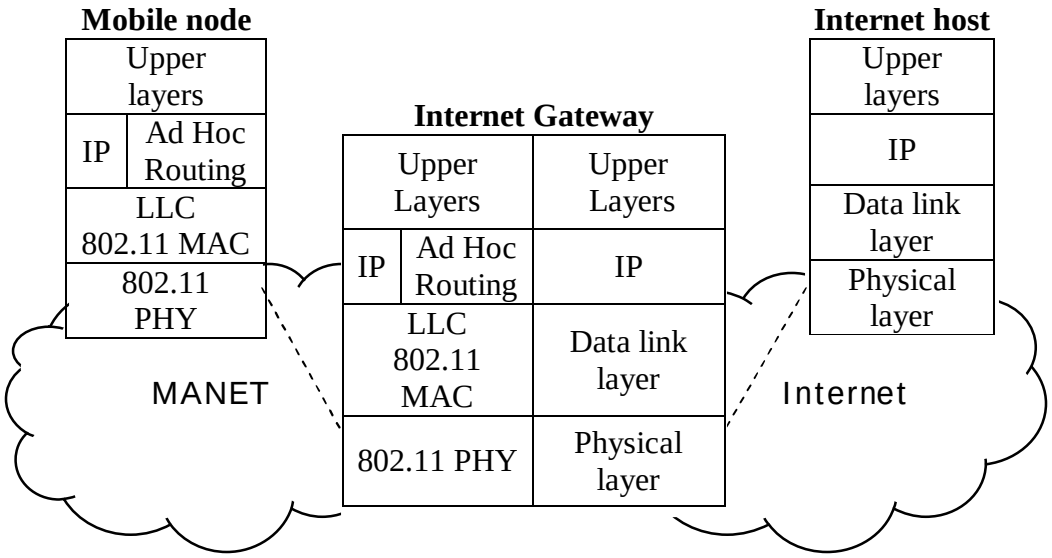


Figure 2.7: The protocol stacks used by mobile nodes, gateways and Internet hosts.

A gateway acts as a bridge between a MANET and the Internet. Therefore, it has to implement both the MANET protocol stack and the TCP/IP suite. A gateway must understand the both protocol stacks to exchange packets between the two networks.

2.4.1 Internet Gateway Discovery

To communicate with the Internet a MANET node has to discover an Internet gateway to which traffic destined for the Internet can be forwarded, and from which traffic returned from the Internet can be received. The gateway discovery schemes used in MANET can be classified into three categories: proactive, reactive and hybrid. Messages that are used in these schemes to discover Internet gateways are: Gateway Discovery (GWDSC) Message, Gateway Reply (GWREP) Message, and Gateway Advertisement (GWADV) Message [9].

2.4.1.1 The Gateway Discovery Message

It contains exactly the same fields with the same functions as the ordinary RREQ message, except a flag. This flag is called *Internet-Global Address Resolution Flag* and is referred to as the I-flag. Hence, the RREQ message extended with the I-flag is referred to as the GWDS**C** message. Figure 2.8 shows the format of the GWDSC message.

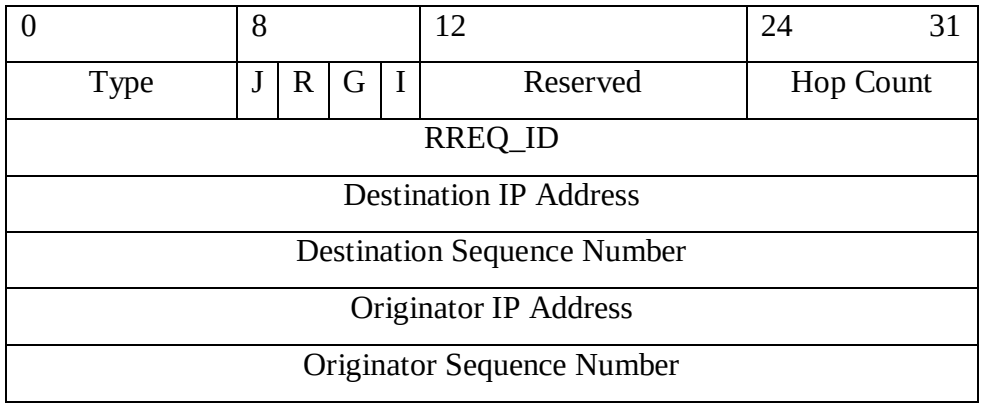


Figure 2.8: The Gateway Discovery (GWDSC) Message Format

The I-flag is used for global address resolution and it indicates that the source node requests global connectivity. Section 2.4.3 describes how the GWDSC message is used to discover a gateway reactively.

2.4.1.2 The Gateway Reply Message

It contains exactly the same fields with the same functions as the ordinary RREP message except the I-flag. Hence, the RREP message extended with the I-flag is referred to as the GWREP message. Figure 2.9 shows the format of the GWREP message.

<table>
<tr><td>0</td><td colspan="4">8</td><td>11</td><td>19</td><td>24 31</td></tr>
<tr><td>Type</td><td>R</td><td>A</td><td>I</td><td>Reserved</td><td>Pre. Sz.</td><td colspan="2">Hop Count</td></tr>
<tr><td colspan="8">Destination IP Address</td></tr>
<tr><td colspan="8">Destination Sequence Number</td></tr>
<tr><td colspan="8">Originator IP Address</td></tr>
<tr><td colspan="8">Lifetime</td></tr>
</table>

Figure 2.9: The Gateway Reply (GWREP) Message Format

The I-flag is used for global address resolution and, if set, it indicates that this RREP contains information about a gateway. Section 2.4.3 describes how the GWREP message is used to unicast a gateway advertisement message in case of reactive gateway discovery.

2.4.1.3 The Gateway Advertisement Message

A GWADV message is basically a RREP message extended with one field similar to the RREQ_ID of the RREQ message. The new field is named as Broadcast_ID. The Broadcast_ID field is used to prevent duplicate broadcasting of the same GWADV message.

Figure 5.1 illustrates the GWADV message format which can solve the problem of duplicated broadcast messages.

<table>
<tr><td>0</td><td>8</td><td>19</td><td>24 31</td></tr>
<tr><td>Type</td><td>Reserved</td><td>Pref. Sz.</td><td>Hop Count</td></tr>
<tr><td colspan="4">Broadcast_ID</td></tr>
<tr><td colspan="4">Destination IP Address</td></tr>
<tr><td colspan="4">Destination Sequence Number</td></tr>
<tr><td colspan="4">Originator IP Address</td></tr>
<tr><td colspan="4">Lifetime</td></tr>
</table>

Figure 2.10: The Gateway Advertisement (GWADV) Message Format

When a mobile node receives a GWADV message, it first checks to determine whether a GWADV message with the same originator IP address and Broadcast_ID has already been received. If such a GWADV message has not been received, the message is rebroadcasted. Otherwise, the newly received GWADV message is discarded. Hence, duplicate GWADV messages are not forwarded.

2.4.2 Proactive Gateway Discovery

In this scheme, the gateway discovery is initiated by a gateway itself. A gateway periodically broadcasts Gateway Advertisement (GWADV) message which is transmitted after the expiration of the gateway's timer. All mobile nodes residing in the gateway's transmission range receive the advertisement. Upon receipt of the advertisement, the mobile nodes that do not have a gateway route, can create one. Mobile nodes that already have a gateway route, can update the route if the corresponding gateway seems better. The advertisement is forwarded by the mobile nodes to the other nodes. In this way, the message is flooded through the whole network. Advertisement interval must be chosen carefully to stop unnecessary flooding. This scheme achieves high throughput and less delay because of the availability of gateway information. The main disadvantage of the scheme is that the advertisement message is flooded through the whole MANET periodically even if there is no such demand from the nodes. This is a very costly operation in terms of node energy and network bandwidth.

2.4.3 Reactive Gateway Discovery

In this scheme, no periodic flooding of gateway advertisement messages is used. A mobile node that wants to access the Internet initiates the reactive gateway discovery by broadcasting a Gateway Discovery (GWDSC) Message in the MANET. Intermediate mobile nodes that receive the message re-broadcast it. Upon receipt of a gateway discovery message, a gateway unicasts a Gateway Reply (GWREP) message to the requestor. A mobile node in the MANET continues to use a selected gateway for a predefined time. If a better gateway appears within this time, the mobile node does not

discover or switch to it. The disadvantage of reactive gateway discovery is that the load on intermediate nodes, especially on those close to a gateway might increase.

2.4.4 Hybrid Gateway Discovery

Hybrid schemes selectively use proactive and reactive gateway discovery. Proactive gateway discovery is used for the mobile nodes within a certain distance around a gateway. Mobile nodes residing outside this distance use reactive gateway discovery. It minimizes the disadvantages of proactive and reactive gateway discovery. However, the scheme needs some intelligent adaptation of TTL value for the gateway advertisement message in order to contain the proactive discovery within an optimum distance.

2.4.5 Internet Gateway Selection

If the mobile nodes discover multiple gateways, they need to select the best gateway. A metric is normally needed to select the best one. Different metrics are be used to select the best gateway in different schemes [1-19] [26]. Some of these are:

- Minimum hop count to the nearest gateway
- Traffic load along the route to the gateway
- Service classes provided and supported by each gateway
- Spatial distance between a MANET node and a gateway
- Speed of the nodes
- Node’s available energy
- Hybrid metric, a combination of two or more of the above metrics.

2.4.6 Handoff

A node performs a handover if it changes its Internet gateway while communicating with a host in the Internet. In case of conventional wireless networks, like WLAN, the quality of the wireless link between a mobile node and the neighboring access points (APs) determines when to handover from one AP to another. However, in MANET, the situation is more complicated; often nodes do not have a direct wireless link to an Internet gateway, most of the times they are connected to a gateway via intermediate nodes. Thus, nodes (sources) cannot use handover policies that are based on the link quality to the AP; rather the complete multi-hop path to the Internet gateway must be taken into consideration. Internet gateway discovery scheme and the ad hoc routing protocol both have enormous influence on the multi-hop handover performances [10]. Two types of handover can occur in case of multi-hop handover. First, a handover can occur if a source itself or any of the intermediate nodes moves and breaks the connection. Therefore, a new connection to the Internet has to be setup that may result in the selection of a new gateway, consequently results in a handover. Second, if a node discovers a better gateway while communicating with a host in the Internet it switches to the new gateway, hence make a handover.

2.5 Summary

In this chapter, we have given the overview of mobile ad hoc networks along with its protocol stacks, salient features, and popular applications. We also talked about how to connect the Internet to MANET. A MANET node first discovers the available Internet gateways in the MANET and selects the best one among them. There are proactive, reactive and hybrid Internet gateway discovery schemes. In the next chapter, we will discuss the pros and cons of current Internet gateway discovery and selection schemes.

Chapter 3

Related Works

During the last decade, many works have been devoted to the study of ad hoc routing protocols, but the decade lacks adequate works to provide Internet connectivity to the nodes in MANET. Since Internet has made information more available and easier to access, the desire for having a MANET connected to the Internet is increasing. Typically, several gateways in a MANET connect the network to the Internet. The rest of the nodes discover the available gateways and select the best one among them.

3.1 Internet Gateway Discovery Schemes

Recently the issue of Internet connectivity to MANET has been addressed by [1-19] [26-27]. MIPMANET [3] was designed to provide nodes in the ad hoc networks with access to the Internet and the mobility services of IP. A foreign agent (FA) in MIPMANET [3] acts as an access point and provides Internet connectivity to an entire ad hoc network. It uses a single IP address as a care-of-address and reverse tunneling to provide Internet access to the nodes. Each FA in the MANET broadcasts foreign agent advertisement messages periodically. Mobile nodes in the network use ad hoc on-demand distance vector (AODV) routing protocol for routing within the MANET. FAs have the MIPMANET Internetworking Unit (MIWU) that is inserted between the FA and the ad hoc network. MIPMANET uses MIPMANET Cell Switching (MMCS) algorithm to handover between foreign agents. Belding-Royer et al. [28] proposed Mobile IP for IPv4 ad hoc networks using AODV routing protocol. In that proposal, a node first has to

determine the location of the destination node before it starts sending packets to that destination. Here, a FA unicasts a route reply (F-RREP) message when it receives a FA discovery message from a mobile node. Mobile nodes use the F-RREP messages to determine the location of the destination nodes. It is capable of routing packets to FA using default route. A disadvantage of this proposal is that, a mobile node has to know that the destination of a packet is not within the ad hoc network before sending it to the FA, which in turn increases the delay for connection setup.

In [1], the authors discussed the technique to provide global Internet connectivity to IPv6 MANET environment using on-demand routing. The paper proposed two Internet gateway discovery schemes: proactive gateway discovery scheme using periodic gateway advertisement messages from the gateway and reactive gateway discovery scheme by flooding gateway discovery messages from the nodes. Lee at el. [13] proposed two gateway advertisement schemes based on the observation of traffic and mobility pattern of nodes to avoid unnecessary routing overhead in MANET. However, the scheme relies on source routing protocol that limits the applicability and scalability of the solution.

In addition to the reactive or proactive gateway discovery schemes [1-12] there are some research works [9] [13-19] that proposed hybrid gateway discovery schemes. In the hybrid schemes, the time-to-live (TTL) value of the gateway advertisements is kept limited to certain boundary in order to contain the proactive discovery within an optimum range. These schemes are mainly designed to minimize the disadvantages of proactive and reactive schemes i.e. to provide good connectivity and low overhead. However, these schemes require some intelligent adaptation of the TTL value. In [19] authors proposed a load-adaptive access gateway discovery protocol that defined a proactive range for the gateway advertisement which is dynamically adjusted according to the changing network conditions. Nevertheless, the gateway advertisement scheme is effective when there are only fixed sized packets in the network. Here the authors used the network size and the number of nodes in the network to compute the initial proactive range, which is unlikely because there may be no good technique to know the size and the number of nodes in a MANET.

3.2 Internet Gateway Selection Schemes

If a node discovers multiple gateways then it has to decide which one is to use. Majority of current gateway selection schemes [1-3] [5] [7] [9] [13-18] [28] use hop count to select the best gateway, and they always select the nearest gateway with the hop count metric. If all the mobile nodes always select their nearest gateway then the nearest gateway may become bottleneck under heavy traffic load, also there might be congested nodes along the route to the gateway. That is, hop count based selection schemes choose a gateway that might have less capacity and difficult to reach. As a result, network performance degrades with the hop count metric.

Few research works [4] [6] [8] [10-12] [19] [26] considered traffic load in addition to the hop count to select the best gateway. Each of these research works treated traffic load differently than the others. Kumar et al. [4] considered the number of packets waiting in the interface queue of the nodes to select a gateway. Khan et al. [6] considered the number of packets waiting in the routing queue of the nodes to select a gateway. However, both of these works converted the number of packets into equivalent hop count without proper justification, which may not provide the actual traffic load information. Le-Trung et al. [10] proposed a hybrid metric for Internet gateway selection that provides load-balancing of intra/inter-MANET traffic. However, the selection scheme introduces extra routing load and requires high processing power consumption to compute the hybrid metric. Li et al. [11] considered the speed of the nodes along with node's available energy and traffic load to select a gateway. Zhanyang et al. [12] also considered the speed of the nodes to compute the gateway selection metric. Nevertheless, obtaining the speed of a node impose additional cost which may limit the applicability of the work. QoS-enabled access gateway selection scheme proposed in [19] considered the packet arrival rate of a gateway in an interval as the traffic load. It uses a Decision Function (DF) that considered the traffic load and hop count to select a gateway. In this case, each intermediate node needs to piggyback its load information periodically on data packets, which increase the header size of the data packets. In [26], the authors proposed a gateway selection scheme based on hop count, gateway load and path quality, and make use of a hybrid search approach which is based on orthogonal genetic algorithm and

sensitivity analysis. The authors have used the maximum packet queue size, average packet queue size and an index α to compute the gateway load. However, the computation of average packet queue size depends on the periodical gateway advertisement and better average can only be obtained for smaller advertisement interval. The authors did not talk about how to select the value of α either. In [26], the authors used the variance in arrival times of periodical gateway advertisement broadcast messages in order to evaluate the quality of the path between mobile nodes and the gateway. However, the computation of the variance needs an intelligent selection of a history window in order to express how long history needs to be considered when calculating the mean value and variance. This makes their selection scheme effective for periodical gateway advertisement only with small advertisement interval. Nevertheless, periodical gateway advertisement with small advertisement interval results in tremendous routing load in the network.

3.3 Summary

In this chapter, we have reviewed the current solutions for Internet gateway discovery and selection. Proactive schemes achieve good connectivity but increase routing load in the MANET. Reactive schemes reduce the routing overhead at the expense of higher delay and lower throughput. Challenge with the hybrid schemes is to choose the appropriate proactive boundary. Internet gateway selection schemes use hop count, traffic load or the mix of these two to select the best gateway. Current gateway selection schemes have the possibility to create the bottleneck and increase routing overhead. In the next chapter, we will introduce our new Internet gateway discovery and selection scheme in MANET.

Chapter 4

Proposed Internet Gateway Discovery and Selection Scheme

In this chapter, we describe our proposed Internet gateway discovery and selection scheme for MANET. At first, we present the network architecture that our scheme is based on. After that, we describe our Internet gateway discovery scheme. We also show the computation of the metrics that are used in our Internet gateway selection scheme.

4.1 Network Architecture

We assume a regular MANET consists of two types of nodes. One type of nodes have Internet connectivity, we call them Internet gateways, and the other type of nodes that don't have Internet connectivity but they can access the Internet through the Internet gateways. We call this second type of nodes simply, mobile nodes.

We assume all the nodes in our MANET have equal transmission range. Nodes can communicate directly with each other if they fall in each other's transmission range. Nodes who are not within each other's transmission range can also communicate indirectly via one or more intermediate nodes. Nodes can join or leave the network anytime. Nodes are free to move in any direction. We did not impose any Internet bandwidth limitation on the Internet gateways. Figure 4.1 shows the schematic diagram of our network architecture.

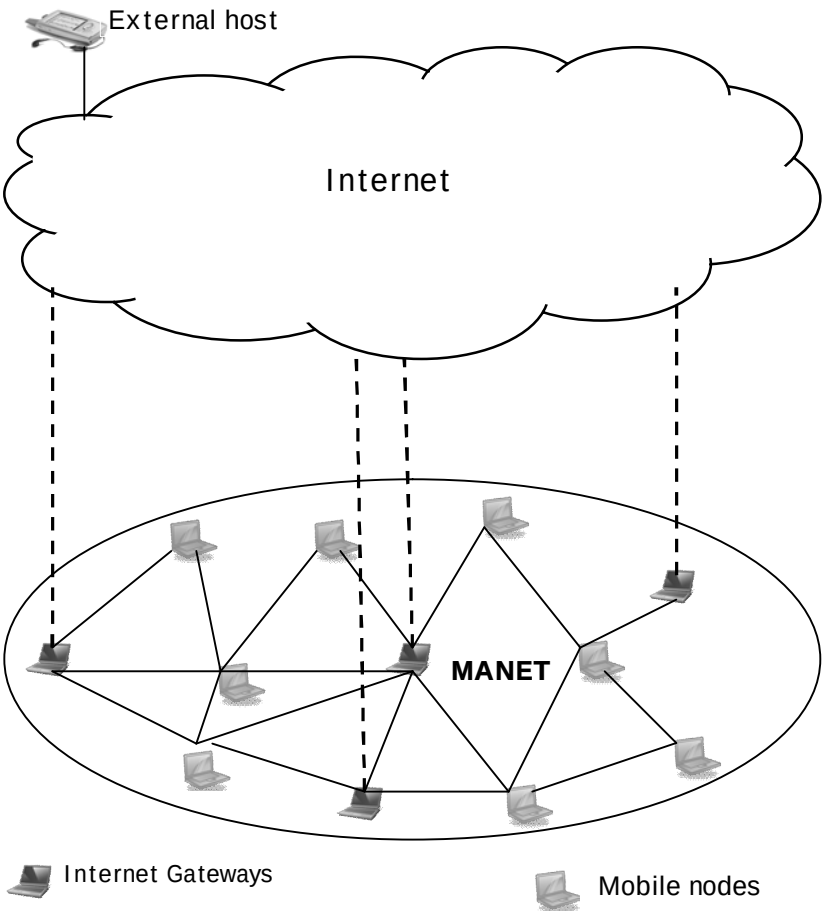


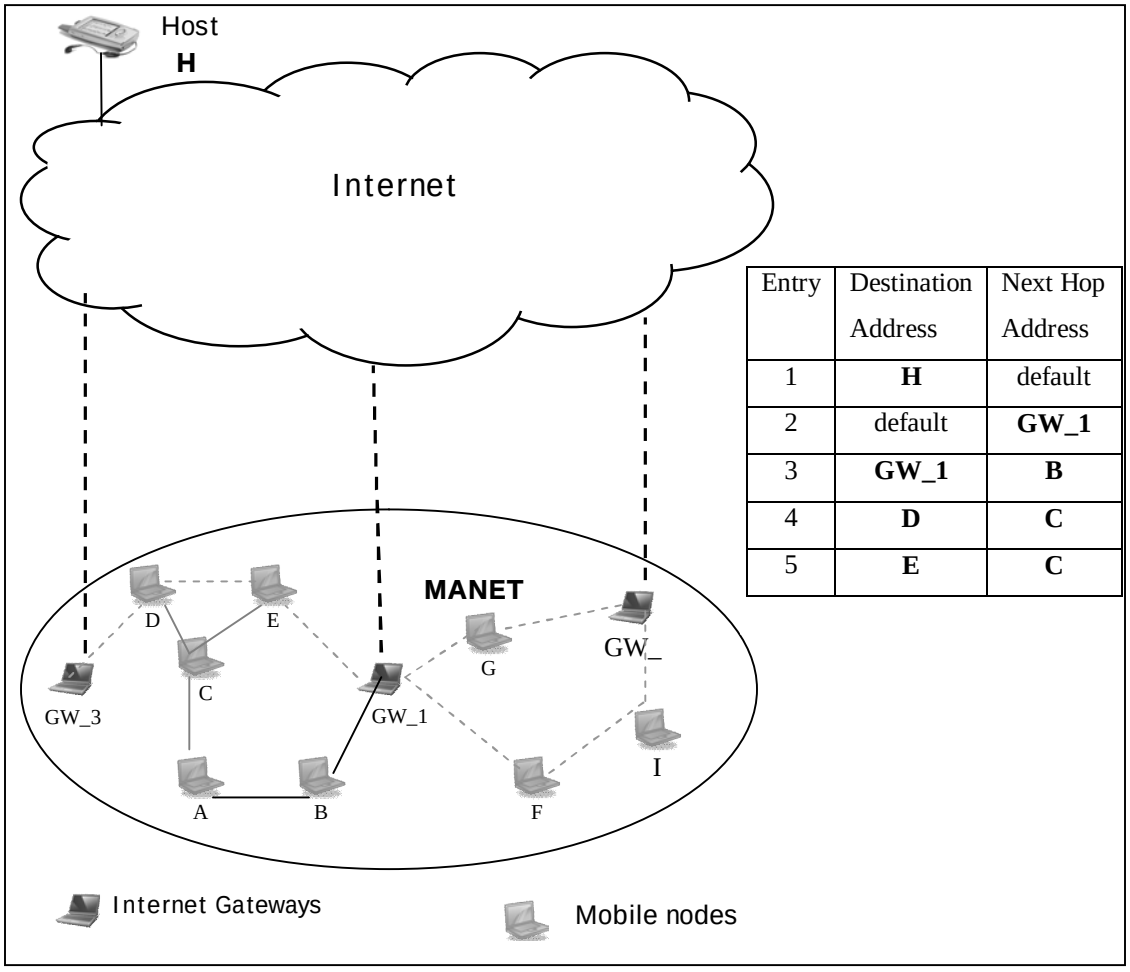
Figure 4.1: Network architecture to connect MANET to the Internet

Internet gateways in our MANET can access the Internet themselves. However, the mobile nodes have to access the Internet through an Internet gateway. For this reason, mobile nodes have to discover the gateways first. We describe our gateway discovery scheme in Section 4.2. If multiple gateways are discovered by a mobile node, the best gateway must be selected to access the Internet. We describe our gateway selection scheme in Section 4.3. Any MANET routing protocol such as AODV [20], OLSR [22] and DSR [24] can be used to route the packets within our network.

4.2 Internet Gateway Discovery

When a mobile node in the MANET wants to access the Internet, at first it has to find a gateway. Like [4] [6] [9], a mobile node in our gateway discovery scheme looks in its

routing table to find a default route i.e. a route to a gateway. In Figure 4.2, we show the routing table of a mobile node **A** containing route entries to a host in the Internet and mobile nodes in MANET. In the figure, **H** represents a host in the Internet, **GW_1** represents an Internet gateway in the MANET, and **B**, **C**, **D**, and **E** represent mobile nodes in the MANET.



Entry	Destination Address	Next Hop Address
1	**H**	default
2	default	**GW_1**
3	**GW_1**	**B**
4	**D**	**C**
5	**E**	**C**

Figure 4.2: Routing table of a mobile node in our scheme

The first entry in the table indicates that the destination is a host in the Internet since the next hop entry for host **H** is set to default. The second entry indicates that gateway **GW_1** has chosen for its Internet connection. The third entry indicates that node **B** is the next hop towards the gateway **GW_1**. The rest of the entries indicate that the destinations are all in the MANET since their next hop entry is not set to default. Mobile node **A**

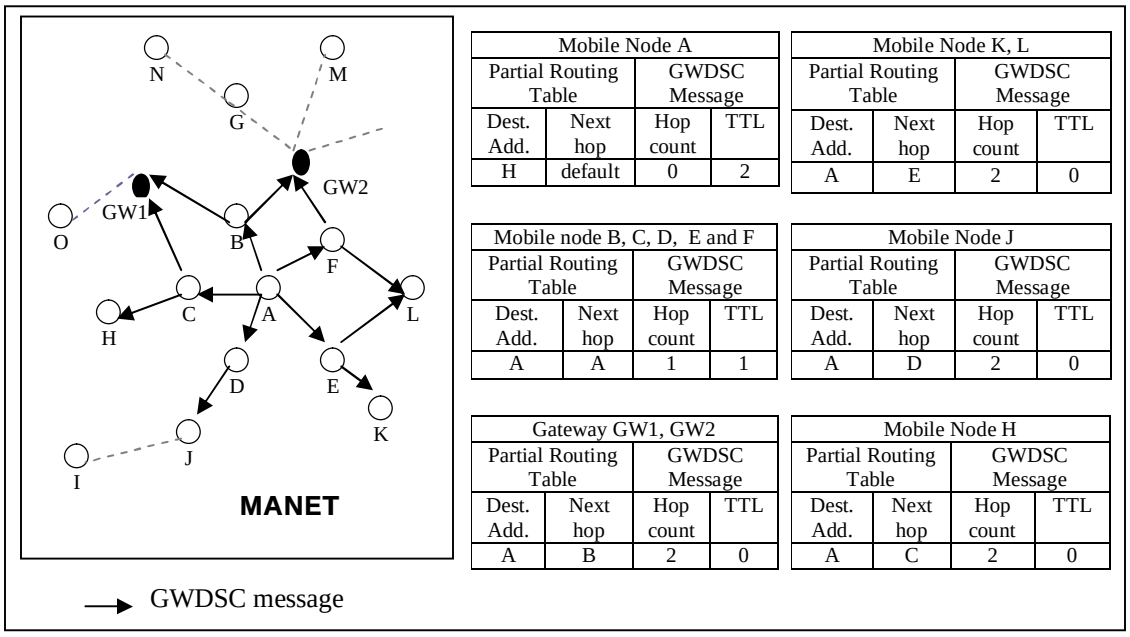
makes recursive lookup into the routing table to find gateway **GW_1** when it wants to communicate with the host **H** in the Internet. If the mobile node finds a default route, it uses the route to send packets to the gateway i.e. to the Internet.

However, if the mobile node does not find a route to a gateway in its routing table, we propose it to start a gateway discovery process by broadcasting a gateway discovery (GWDSC) message in the MANET. While broadcasting the GWDSC message, we propose the requesting mobile node to set an initial time to live (TTL) value for the message and start a timer to wait for the reception of the gateway advertisement message from the gateways. Figure 4.3 shows the format of the GWDSC message.

<table>
<tr><td>0</td><td colspan="4">8</td><td>12</td><td>24 31</td></tr>
<tr><td>Type</td><td>J</td><td>R</td><td>G</td><td>I</td><td>Reserved</td><td>Hop Count</td></tr>
<tr><td colspan="7">RREQ_ID</td></tr>
<tr><td colspan="7">Destination IP Address</td></tr>
<tr><td colspan="7">Destination Sequence Number</td></tr>
<tr><td colspan="7">Originator IP Address</td></tr>
<tr><td colspan="7">Originator Sequence Number</td></tr>
</table>

Figure 4.3: Format of GWDSC messages in our scheme

In our scheme, upon receipt of a GWDSC message, an intermediate node creates a reverse route entry for the requestor in its routing table and forwards the GWDSC message to its neighbors. In this way, a GWDSC message reaches one or more Internet gateways in the network if there is any. Figure 4.4 shows an example of how a GWDSC message is flooded in the network when a mobile node **A** wants to communicate with a host **H** (not shown in the figure) in the Internet. The figure also shows how reverse entries are created in the intermediate nodes. Here, the GWDSC message hits the gateways **GW1** and **GW2** in the network.



Mobile Node A			
Partial Routing Table		GWDSC Message	
Dest. Add.	Next hop	Hop count	TTL
H	default	0	2

Mobile Node K, L			
Partial Routing Table		GWDSC Message	
Dest. Add.	Next hop	Hop count	TTL
A	E	2	0

Mobile node B, C, D, E and F			
Partial Routing Table		GWDSC Message	
Dest. Add.	Next hop	Hop count	TTL
A	A	1	1

Mobile Node J			
Partial Routing Table		GWDSC Message	
Dest. Add.	Next hop	Hop count	TTL
A	D	2	0

Gateway GW1, GW2			
Partial Routing Table		GWDSC Message	
Dest. Add.	Next hop	Hop count	TTL
A	B	2	0

Mobile Node H			
Partial Routing Table		GWDSC Message	
Dest. Add.	Next hop	Hop count	TTL
A	C	2	0

Figure 4.4: Broadcast of GWDSC messages in our scheme.

In Figure 4.4, we can see that mobile node **A** broadcasts a GWDSC message with a TTL value equal to 2 and Hop count equal to zero. The node also sets the next hop to default for host **H** while broadcasting the GWDSC message. The mobile nodes in the MANET that are within 2 hops distance from node **A** create a reverse route entry for **A**, increment the Hop count field, decrement the TTL field of the GWDSC message and forward it to the neighbors if the TTL is not zero. The Hop count value of the GWDSC message will be used by the gateways GW1 and GW2 to broadcast gateway advertisements.

We propose an Internet gateway to broadcast a gateway advertisement (GWADV) message when triggered by a GWDSC message. We also propose to set the TTL value of the GWADV message equal to the distance of the gateway from the requesting mobile node. In our scheme, we control the TTL value of the GWADV message to contain the dissemination of the GWADV message to a certain range, which helps to reduce the routing overhead to an extent. We allow gateways to broadcast GWADV messages only in response to GWDSC messages in order to avoid unnecessary flooding of GWADV messages in the network.

In addition to the conventional fields, we have added two new fields in the GWADV message header. We name these new fields ***Q*** and ***N*** respectively. We use the ***Q*** field to represent the total interface queue size of nodes along a route from a gateway to a mobile node. We use ***N*** field to represent the total number of neighbors of the nodes along a route from a gateway to a mobile node. We use the Hello messages of AODV routing protocol to obtain the neighbor information of a gateway or a mobile node.

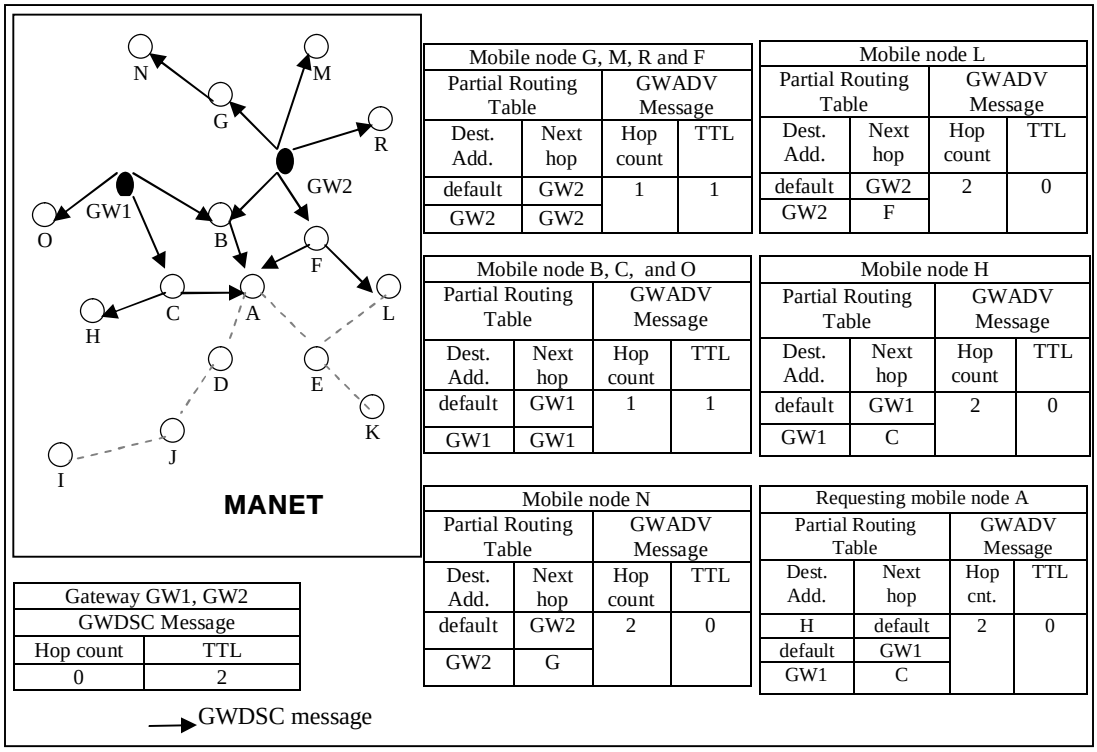
We propose an Internet gateway to populate these two fields before flooding a GWADV message. We also propose intermediate nodes to update these two fields while forwarding the message to the next nodes. The modified structure of a gateway advertisement message header in our scheme is given in Figure 4.5.

<table>
<tr><td>0</td><td>8</td><td>19</td><td>24 31</td></tr>
<tr><td>Type</td><td>Reserved</td><td>Pref. Sz.</td><td>Hop Count</td></tr>
<tr><td colspan="4">Broadcast_ID</td></tr>
<tr><td colspan="4">Destination IP Address</td></tr>
<tr><td colspan="4">Destination Sequence Number</td></tr>
<tr><td colspan="4">Source IP Address</td></tr>
<tr><td colspan="4">Lifetime</td></tr>
<tr><td colspan="4">Q</td></tr>
<tr><td colspan="4">N</td></tr>
</table>

Figure 4.5: Format of GWADV Message in our scheme

Upon receipt of a GWADV message, we propose a mobile node to decrement the TTL first and to configure the corresponding gateway if it does not have a gateway configured yet. In this way, more nodes in the network will have the opportunity to configure their gateway without broadcasting a GWDSC message, i.e., our scheme will reduce the GWDSC message broadcast to a significant level. Mobile nodes that already have their gateway configured should reconfigure the gateway if the corresponding gateway seems better. A GWADV message is forwarded to the neighbors if the TTL value is not zero. In this way, we allow the GWADV message to reach to the requesting mobile node. Therefore, in our scheme, a GWADV message helps not only the requesting mobile node

but also the other nodes in the network to configure their gateway. As a result, our proposed scheme helps a mobile node in a MANET to hand off to a better gateway even before its current Internet connection is broken. Figure 4.6 shows how we flood the GWADV messages in the MANET in response to GWDSC messages. The figure also shows how the intermediate nodes create forward entries for the gateways and update their default route. Here, the gateways GW1 and GW2 broadcast GWADV messages in the network with TTL value 2 (Hop count of the GWDSC message received from node A, taken from Figure 4.4). For simplicity, we do not show the ***Q*** and ***N*** field of the GWADV message in the figure.



Mobile node G, M, R and F

Partial Routing Table		GWADV Message	
Dest. Add.	Next hop	Hop count	TTL
default	GW2	1	1
GW2	GW2		

Mobile node L

Partial Routing Table		GWADV Message	
Dest. Add.	Next hop	Hop count	TTL
default	GW2	2	0
GW2	F		

Mobile node B, C, and O

Partial Routing Table		GWADV Message	
Dest. Add.	Next hop	Hop count	TTL
default	GW1	1	1
GW1	GW1		

Mobile node H

Partial Routing Table		GWADV Message	
Dest. Add.	Next hop	Hop count	TTL
default	GW1	2	0
GW1	C		

Mobile node N

Partial Routing Table		GWADV Message	
Dest. Add.	Next hop	Hop count	TTL
default	GW2	2	0
GW2	G		

Requesting mobile node A

Partial Routing Table		GWADV Message	
Dest. Add.	Next hop	Hop cnt.	TTL
H	default	2	0
default	GW1		
GW1	C		

Gateway GW1, GW2

GWDSC Message	
Hop count	TTL
0	2

Figure 4.6: Broadcast of GWADV message in our scheme

In Figure 4.6, we can see that both the gateways broadcast GWADV messages with a TTL value equal to 2. This TTL value is equal to the Hop count value of the GWDSC message they receive from the mobile node **A**. The mobile nodes that are within the 2 hops distance from **GW1** or **GW2** create their corresponding default route and a forward route for **GW1** or **GW2**. The mobile nodes also increment the Hop count field, decrement the TTL field of the GWADV message and forward it to the neighbors if the

TTL is not zero. The figure also shows that the requesting mobile node **A** creates a default route i.e. a route to gateway **GW1** with the next hop set to **C**.

However, if the requesting mobile node does not receive any GWADV message before the timer expires, we propose the node to broadcast a new GWDSC message with an increased TTL value. We propose the requesting mobile node to increase the TTL value linearly. We increment the TTL value linearly to experience less routing overhead (GWDSC messages). We allow this process to continue until either the requesting mobile node receives a GWADV message or it broadcasts a GWDSC message with a pre-defined maximum TTL value

An algorithmic depiction of our Interne gateway discovery scheme is given below:

Algorithm 1: Algorithm for Internet Gateway Discovery

{This algorithm is invoked when a mobile node ***s*** wants to send packets to a host in the Internet}

1. **if** (default_route)
2. forward (packets)
3. exit
4. **else**
5. enqueue (packets)
6. *send_GWDSC_message*()
7. **end if**

Algorithm 2: procedure *send_GWDSC_message*()

{This algorithm is invoked when a mobile node ***s*** needs to discover an Internet gateway to send packets to the Internet}

1. GWDSC_TTL = initial_TTL
2. last_GWDSC_TTL = GWDSC_TTL
3. GWDSC_Hopcount = 0
4. GWADV_receive = FALSE
5. **while** (GWADV_receive = = FALSE or last_GWDSC_TTL <= pre-defined_TTL)
6. broadcast (GWDSC_message)
7. increment GWDSC_TTL
8. GWDSC_Hopcount = 0
9. last_GWDSC_TTL = GWDSC_TTL
10. **end while**

Algorithm 3: procedure *node_receive_GWDSC_message*()

{This algorithm is invoked when an intermediate node ***p*** receives a GWDSC message from a requesting mobile node ***s***}

```
1.  if (lookup(GWDSC_message_source))
2.    drop(GWDSC_message)
3.    exit
4.  end if
5.  GWDSC_TTL--
6.  route_add (s)
7.  GWDSC_Hopcount ++
8.  if (GWDSC_TTL != 0)
9.    forward (GWDSC_message)
10. else
11.   drop (GWDSC_message)
12. end if
```

Algorithm 4: procedure *gateway_receive_GWDSC_message*()

{This algorithm is invoked when an Internet gateway ***q*** receives a GWDSC message from a requesting mobile node ***s***}

1. **if** (lookup(GWDSC_message_source))
2. drop(GWDSC_message)
3. exit
4. **end if**
5. GWDSC_TTL--
6. route_add (s)
7. GWDSC_Hopcount ++
8. *send_GWADV_message*()

Algorithm 5: procedure *send_GWADV_message*()

{This algorithm is invoked when an Internet gateway ***q*** broadcasts a GWADV message in response to a GWDSC message from a requestor ***s***}

1. GWADV_TTL=GWDSC_Hopcount
2. $\boldsymbol{Q} = \boldsymbol{int\_q\_size_q}$
3. $\boldsymbol{N} = \boldsymbol{n_q}$
4. GWADV_Hopcount = 0
5. broadcast (GWADV_message)

Algorithm 6: procedure *node_receive_GWADV_message*()

{This algorithm is invoked when an intermediate node ***p*** receives a GWADV message from a gateway ***q***}

```
1.  if (lookup(GWADV_message_source))
2.    drop(GWADV_message)
3.    exit
4.  end if
5.  GWADV_TTL--
6.  select_gateway()
7.  if (GWADV_TTL != 0)
8.    forward (GWADV_message)
9.  else
10.   drop(GWADV_message)
11. end if
```

Algorithm 7: procedure requestor_*receive_GWADV_message*()

{This algorithm is invoked when a requesting mobile node ***s*** receives a GWADV message from a gateway ***q***}

```
1.  if (lookup(GWADV_message_source))
2.    drop(GWADV_message)
3.    exit
4.  end if
5.  GWADV_receive = = TRUE
6.  GWADV_TTL--
7.  select_gateway()
8.  if (GWADV_TTL != 0)
9.  forward (GWADV_message)
10. else
11. drop (GWADV_message)
12. end if
```

Thus, our gateway discovery scheme consists of on-demand GWDSC messages like reactive scheme, broadcast of GWADV messages like proactive scheme and limited TTL value for GWADV messages like hybrid scheme. That is, our scheme combines the bests of the three conventional Internet gateway discovery schemes and can provide efficient and faster discovery of Internet gateways.

4.3 Internet Gateway Selection

We propose a new composite metric to select the best gateway when a mobile node receives multiple gateway advertisement messages from multiple gateways; we call this new metric ***gateway-cost*** (***gc***). Our metric ***gc*** is composed of three factors: hop count, interface queue size and total number of neighbors.

Like [1-3] [5] [7] [9] [13-18] [28], we consider hop count to select the best gateway. It denotes the number of nodes or routers between a mobile node and an Internet gateway. This factor allows a mobile node to reach the Internet using minimum number of hops which facilitates the rapid convergence and resource thriftiness of the network.

We consider the interface queue size of each node along a route to a gateway. Interface queue size of a node denotes the number of packets waiting in the interface queue of that node. If the size of the interface queue of each node along a route to a gateway is less, then more packets can be sent to the Internet using that route and the packets will have to wait less. Thus, we consider interface queue size of each node to allow fair distribution of the network load among the gateways and congestion prevention in the network.

We consider the total number of neighbors of each node along a route to a gateway. This factor helps a mobile node to select a gateway whose path is least dense. A least dense path is more likely to have least contention and best to use to reach the gateway. As far as we know, nobody used this factor to select a gateway in a MANET before us.

Whenever a node ***p*** in a MANET receives a GWADV message from a gateway ***q***, we propose it to calculate ***gc*** using eq. (1):

$$gc_q = hc_q + \frac{Q}{Q+1} + \frac{N}{N+1} \qquad (1)$$

$$q \in V_{GW}$$

$$Q = \sum_{i=1}^{hc_q} int\_q\_size_i$$

$$N = \sum_{i=1}^{hc_q} n_i$$

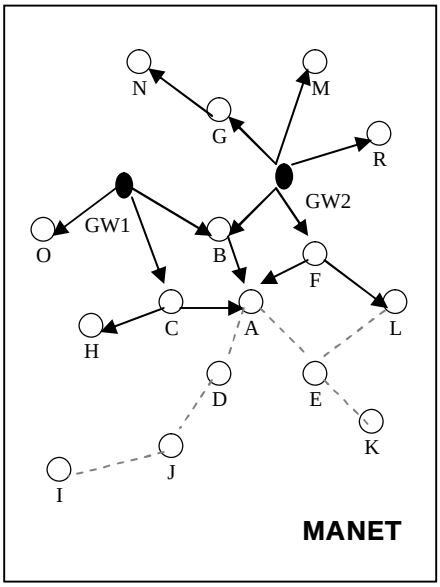
Where V_{GW} is the set of Internet gateways present in the network, hc_q is the number of hops from q to p, $int\_q\_size_i$ represents the interface queue size of node i along the route from gateway q to node p, n_i represents the number of neighbors of node i along the route from gateway q to node p.

When a mobile node receives multiple GWADV messages from multiple gateways, we select the gateway with the lowest gc.

We give more emphasis on the hop count because it is always better to select a shorter route to minimize network delay and to optimize network resource usage. A packet routing through a shorter path also have better chance to face less network adversaries, such as bit error and congestion. Although the queue size and the number of neighbors along the route help us to avoid the gateways having bad route to reach, these two are actually less significant factor compared to hop count. Thus, if the two factors are kept intact like the hop count in the computation of the metric gc, then our selection scheme may choose a gateway which is not closest in terms of hop count. As a result, a mobile node in a MANET has to travel a longer route to reach an Internet gateway in the MANET. A longer route not only increases delay or consumes network bandwidth and node energy but it also involves more intermediate nodes to forward packets to an Internet gateway. A route to a gateway with higher number of intermediate nodes has better chance to suffer from more congestion and collision compared to that of smaller routes. Consequently, this fact may cause more packets drop and route re-discoveries in the network. For this reason, we give less emphasis on these two factors. To do so, we individually adjust these two factors so that they can contribute positively in the

computation of the gateway selection metric ***gc*** but their individual contribution always remains less than 1. Therefore, our metric ***gc*** selects the gateway whose path is not only less loaded and less dense but also shortest.

Figure 4.7 shows an example of how the values of the ***hc, Q,*** and ***N*** factors are updated as the GWADV messages are flooded in the network. The figure also shows how the mobile nodes select the gateways based on these factors. Though the interface queue size of a node depends on the network conditions like traffic load, speed of the nodes, etc., for easy understanding we have considered interface queue size of a node is equal to its number of neighbors in this figure. In the figure, we can see that each mobile node computes its ***gc*** after receiving GWADV messages from gateway GW1 or GW2 and creates its gateway route based on ***gc***.



GWADV message

Mobile node R						
GWADV Message					***gc***	Gateway route
From Gateway	From Node	Hop count	***Q***	***N***		
GW2	GW2	1	6	6	2.71	R>GW2

Mobile node M						
GWADV Message					***gc***	Gateway route
From Gateway	From Node	Hop count	***Q***	***N***		
GW2	GW2	1	6	6	2.71	M>GW2

Mobile node F						
GWADV Message					***gc***	Gateway route
From Gateway	From Node	Hop count	***Q***	***N***		
GW2	GW2	1	8	8	2.78	F>GW2

Mobile node G						
GWADV Message					***gc***	Gateway route
From Gateway	From Node	Hop count	***Q***	***N***		
GW2	GW2	1	7	7	2.75	G>GW2

Gateway GW1			
GWADV Message			Gateway route
Hop count	***Q***	***N***	
0	3	3	N/A

Gateway GW2			
GWADV Message			Gateway route
Hop count	***Q***	***N***	
0	5	5	N/A

Mobile node L						
GWADV Message					***gc***	Gateway route
From Gateway	From Node	Hop count	***Q***	***N***		
GW2	F	2	9	9	3.8	L>F> GW2

Mobile node B							
GWADV Message						***gc***	Gateway route
From Gateway	No.	From Node	Hop count	***Q***	***N***		
GW1	1	GW1	1	6	6	**2.71**	B>GW1
GW2	2	GW2	1	8	8	2.78	

Mobile node N						
GWADV Message					***gc***	Gateway route
From Gateway	From Node	Hop count	***Q***	***N***		
GW2	G	2	8	8	3.78	N>G> GW2

Mobile node H						
GWADV Message					***gc***	Gateway route
From Gateway	From Node	Hop count	***Q***	***N***		
GW1	C	2	7	7	3.75	H>C> GW1

Mobile node C						
GWADV Message					***gc***	Gateway route
From Gateway	From Node	Hop count	***Q***	***N***		
GW1	GW1	1	6	6	2.71	C>GW1

Mobile node A							
GWADV Message						***gc***	Gateway route
From Gateway	No.	From Node	Hop count	***Q***	***N***		
GW1	1	C	2	11	11	**3.83**	A>C> GW1
GW1	2	B	2	11	11	3.83	
GW2	3	B	2	13	13	3.86	

Mobile node O						
GWADV Message					***gc***	Gateway route
From Gateway	From Node	Hop count	***Q***	***N***		
GW1	GW1	1	4	4	2.60	O>GW1

Figure 4.7: Example of Gateway selection in our scheme

In Figure 4.7 we can see that the gateways **GW1** and **GW2** populate the ***Q*** and ***N*** field while broadcasting their corresponding GWADV messages. Mobile nodes in the

MANET that receive the GWADV messages from **GW1** or **GW2** update the ***Q*** and ***N*** fields of the corresponding GWADV message and rebroadcasts the message if the TTL of the message is not zero. In the process the mobile nodes also compute their corresponding ***gc*** and select best gateway based on ***gc***.

An algorithmic depiction of our Internet gateway selection scheme is given below:

Algorithm 8: procedure *select_gateway*()

{This algorithm is invoked when a mobile node ***p*** creates or updates its gateway route after receiving a GWADV message from a gateway ***q***}

1. $Q = Q + int\_q\_size_p$
2. $N = N + n_p$
3. GWADV_Hopcount++
4. $gc_q = hc_q + \frac{Q}{Q+1} + \frac{N}{N+1}$
5. **if** (default_route!)
6. $gw\_rt\_cost_p = gc_q$ //***gw_rt_cost*** is the path cost to reach a gateway from node ***p***
7. route_add (***q***)
8. **while** (buffer_packets = dequeue(packets))
9. forward (packets)
10. **end while**
11. **end if**
12. **if** (default_route)
13. **if** ($gc_q < gw\_rt\_cost_p$)
14. route_update (***q***)
15. **end if**
16. **end if**

4.4 Summary

In this chapter, we have introduced our new Internet gateway discovery scheme, which uses a triggered broadcast of gateway advertisement in the MANET. The broadcast is triggered by gateway discovery request from the mobile nodes in MANET. Our scheme also limits the dissemination range of the gateway advertisement messages. The metric comprises of hop count, traffic load, and nodes neighbor information is used to select the best gateway. In the next chapter, we evaluate our scheme through simulation.

Chapter 5

Performance Evaluation

To evaluate the performance of our proposed Internet gateway discovery and selection scheme, we implemented our scheme in ns-2 [29] network simulator and compared the results with that of the proactive, reactive and hybrid schemes that were proposed in [30]. We also modified the MANET routing protocol AODV [20] to route packets between a gateway and a mobile node.

5.1 Performance Metrics

We compare all the Internet gateway discovery and selection schemes based on three performance metrics namely Internet Packet Delivery Ratio, Average End-to-End Delay, and Normalized Control Overhead. These are the standard performance metrics that are also used by many research works [4] [6-12] [19] to evaluate Internet Gateway Discovery and Selection Schemes.

- **The Internet Packet Delivery Ratio (IPDR)**: IPDR is defined as the ratio between the total number of data packets received by the corresponding destination hosts in the Internet and the total number of data packets sent to the Internet by all the mobile nodes in the MANET.

- **The Average End-to-End Delay**: It is defined as the average time needed to send a data packet from a node to a host in the Internet. It is computed in milliseconds (ms).

- **The Normalized Control Overhead (NCO):** NCO is defined as the ratio between the total number of AODV messages transmitted by the nodes in MANET and the total number of data packets received by the hosts in the Internet.

We vary the number of nodes in MANET from 10 to 30 to see the network behavior under different traffic load. The number of neighbors of each node also varies with the number of nodes in the MANET. We vary the speed of the nodes from 2 to 30 m/s which allows us to compare the performance of the schemes in different speeds, such as walking speed (2 m/s), downtown driving speed (10 m/s), suburban driving speed (20 m/s), and highway driving speed(30 m/s) [13].

5.2 Simulation Setup

This section describes the network scenario, the movement model, the communication model, and the simulation parameters that we have used in our study.

5.2.1 Scenario

Like [11] [14] [19] [26], our simulated network is spanning in a standard area of 1000x1000m^2. Each mobile node in our simulation has a wireless transmission range of 250 meter, which is the standard range and also used by the other research works [4] [6] [9] [10] [11] [15] [19] [26] [27]. This transmission range ensures no network partitioning.

We have considered 4 Internet gateways in the MANET in our simulation scenarios in order to load balance the Internet traffic. We assume a higher Internet bandwidth for gateways compared to that of the MANET nodes. We set the Internet bandwidth of each gateway to 10 Mbps.

We ran our simulations for 500 units of simulation time. According to our observation, 500 units of simulation time is high enough to see the steady behavior of the network in different scenarios. The seed time for each node to send data packets is considered 0.5 units of the simulation time. This seed time confirms that all the schemes start their gateway discovery process before the nodes start sending the data packets to the Internet.

A screenshot of a simulation scenario is given in Figure 5.1. In the figure, the red-colored hexagonal nodes represent the gateways, the blue-colored square nodes represent the Internet hosts and the green-colored circular nodes represent the mobile nodes.

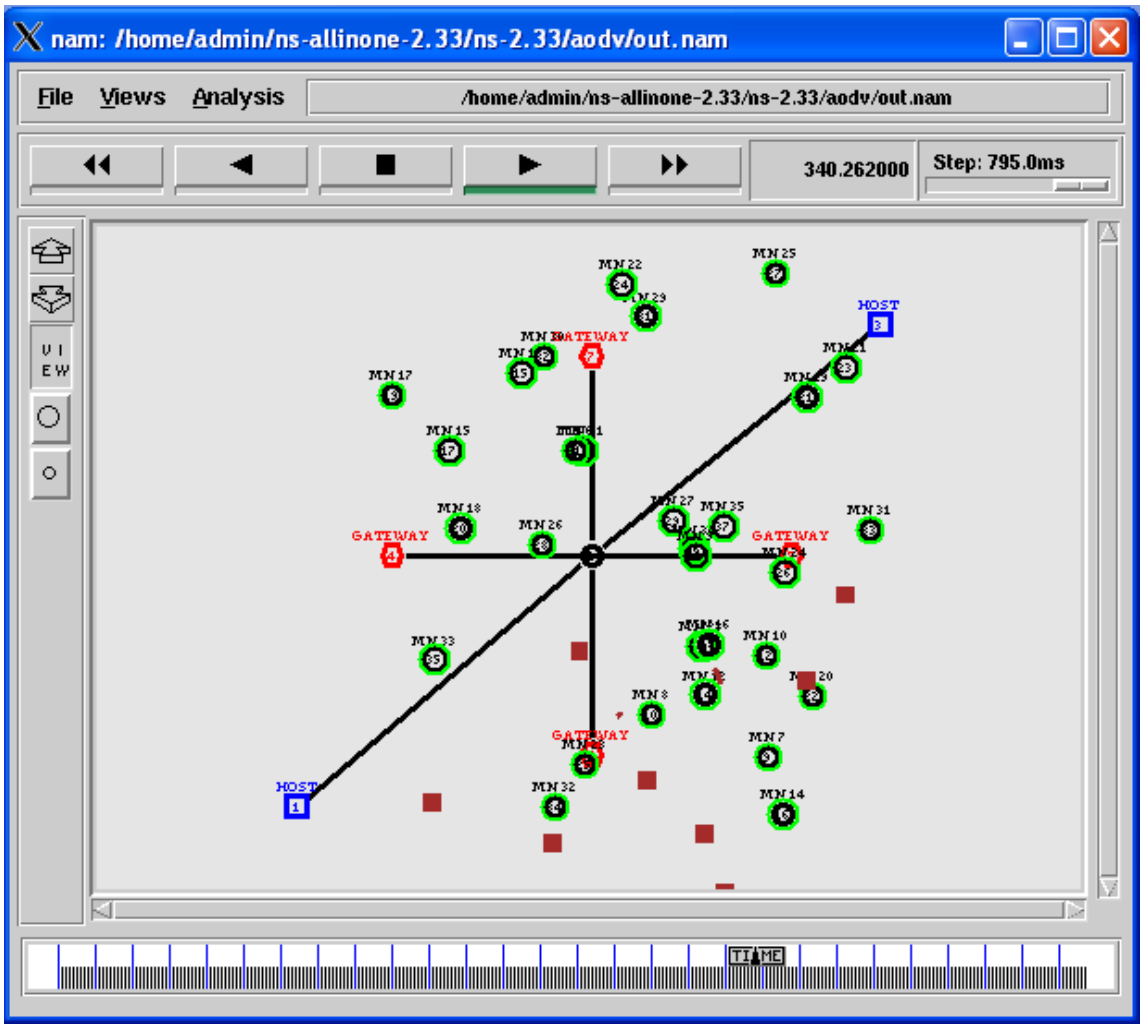


Figure 5.1**:** Screenshot of a Simulation Scenario

5.2.2 Movement Model

We used the Random Waypoint Movement Model [31] as the mobility model for our simulation. It is the benchmark mobility model that has been used in many research works [3-4] [6-9] [11-19] [26-27] in order to evaluate network protocols in MANET. According to this model, a mobile node remains stationary for a certain period called *pause time*. After the *pause time* is over the node selects a destination randomly and

moves to that destination at a random speed. The random speed is distributed uniformly between zero (zero not included) and some maximum speed. We set the maximum speeds between 2 to 30 m/s for different scenarios. When the node reaches the destination, it again remains stationary for the *pause time* period and repeats the same procedure until the end of the simulation. We set the *pause time* to 20 seconds in our simulations which is good enough for a node to change the movement direction.

5.2.3 Communication Model

We allowed all the mobile nodes in the network to access the Internet, i.e., each mobile node sends data packets to the hosts in the Internet. Each mobile node in our simulation uses Constant Bit Rate (CBR) traffic to send packets to the corresponding hosts in the Internet. We use CBR traffic due to the reason mentioned in section 2.1.1. We wish to see the performance of different schemes under heavy traffic load. For this reason, we allow each mobile node to generate 10 packets per second and send them to the Internet. Like [4] [6] [9-12] [14-15] [19] [26-27], we permit each mobile node in the MANET to generate packets of size 512 bytes. By varying the number of nodes, we actually varied the traffic load in different simulation scenarios.

5.2.4 Parameters

Table 5.1 gives the values of some simulation parameters that are used for most of the simulation scenarios.

Table 5.1**:** Common parameters used in most of the simulation scenarios

Parameter	**Value**
Number of Internet gateways	4
Number of hosts in the Internet	2
Topology size	1000 x 1000 m^2
Transmission range	250 m
Internet BW	10 Mbps
Mobility Model	Random waypoint
Traffic type	CBR
Packet size	512 bytes
Pause time	20 s
Simulation time	500 s

5.3 Result Analysis

Figures 5.2, 5.3 and 5.4 report IPDR, average end-to-end delay, and NCO respectively by varying the number of nodes but setting the maximum speed of a node to 30 m/s. In these figures we labeled our scheme as “interactive”. We have taken the average of 10 simulation run results for each data point plotted in the figures.

When there are fewer nodes (less than 20) in the network, the total traffic generated by them is comparatively less. As a result, there is less congestion in the network which helps the nodes to deliver the packets to the gateways with less dropout and the gateways can also forward the packets to the Internet with ease. However, when the number of nodes in the network increases, the traffic load in the network also starts to increase.

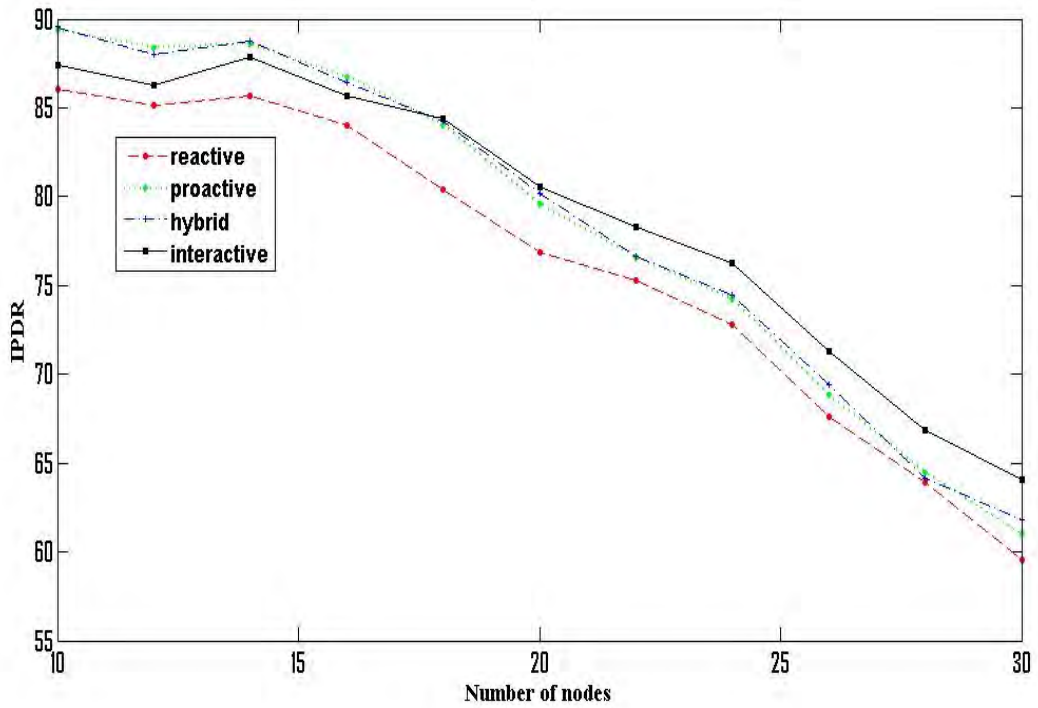


Figure 5.2: IPDR of all schemes against the number of nodes

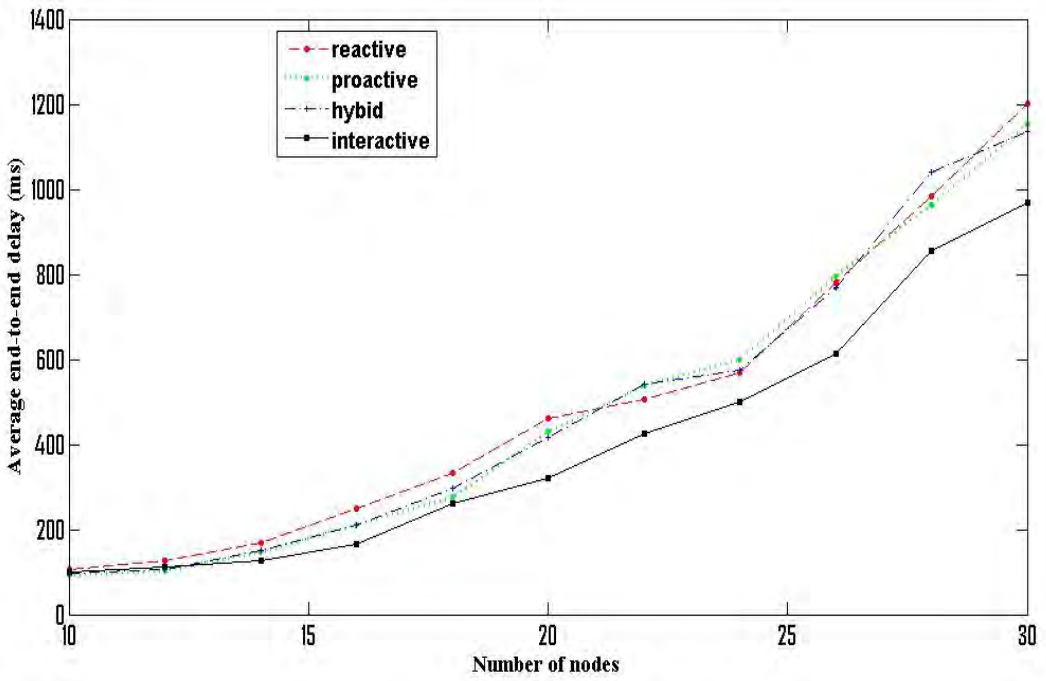


Figure 5.3: Average end-to-end delay of all schemes against the number of nodes.

Increased traffic load results in more congestion and more collisions in the network. As a result more packets are waiting in the interface queue of the forwarding nodes and getting dropped if the waiting time exceeds its limit. These facts reduce the packet delivery ratio and increase the end-to-end delay. Thus, IPDR decreases (Figure 5.2) and the average end-to-end delay increases (Figure 5.3) with the increase in the number of nodes in all the schemes. The periodic GWADV messages in the network in the other schemes help the nodes to have updated gateway information and achieve higher IPDR (Figure 5.2) with fewer nodes in the network. However, IPDR in our scheme started to exceed the IPDR of other schemes when the number of nodes is 20 or more. The average end-to-end delay obtained from our scheme is also better than that of other schemes (Figure 5.3). By avoiding the forwarding nodes having longer interface queue as well as the route to the gateway having higher concentration of neighbor nodes our scheme suffers from less packet drop and less waiting. For these reasons, IPDR is higher and the average end-to-end delay is lower in our scheme compared to that of other schemes while the number of nodes in MANET is increasing beyond 20.

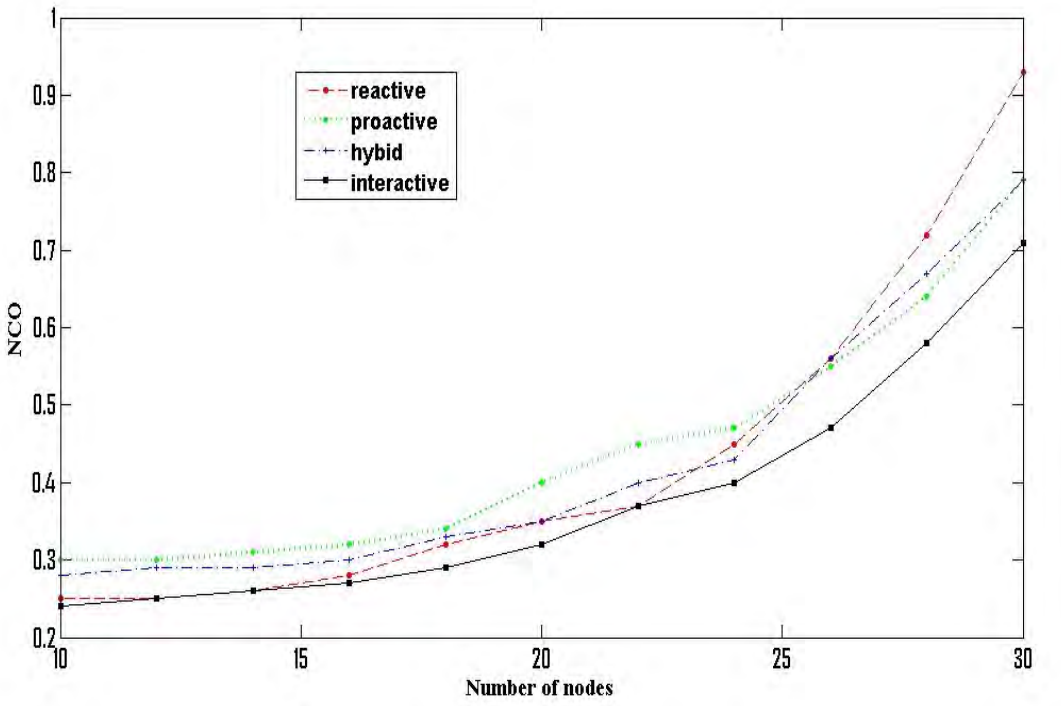


Figure 5.4: NCO of all schemes against the number of nodes.

From Figure 5.4 we can see that our scheme out performs the other schemes with respect to NCO performance metric. NCO obtained from all the schemes increase with the number of nodes in the network. Traffic load in the network increases as the number of nodes in MANET increases, which in turn increases the packet drop as explained earlier. Since NCO is the ratio between the number of routing packets and the number of successfully delivered data packets, it increases when there are less delivered data packets. As our scheme suffers from less packet drop than that of the others, it yields less NCO than that of others. Again, a gateway in our scheme broadcasts a GWADV message in response to a GWDSC message. Not only the requesting mobile node gets the gateway information from the GWADV message but also the other nodes get the same information without transmitting their own GWDSC messages. This technique allows many mobile nodes to bypass the gateway discovery phase. As a result, they do not overwhelm the network by broadcasting GWDSC messages. For this reason, we have less routing packets in our scheme than that of other schemes, i.e., less NCO.

Figures 5.5, 5.6, and 5.7 report the same performance metrics respectively by varying the speed of the nodes but using only 30 mobile nodes. We have taken the average of 10 simulation run results for each data point plotted in the figures.

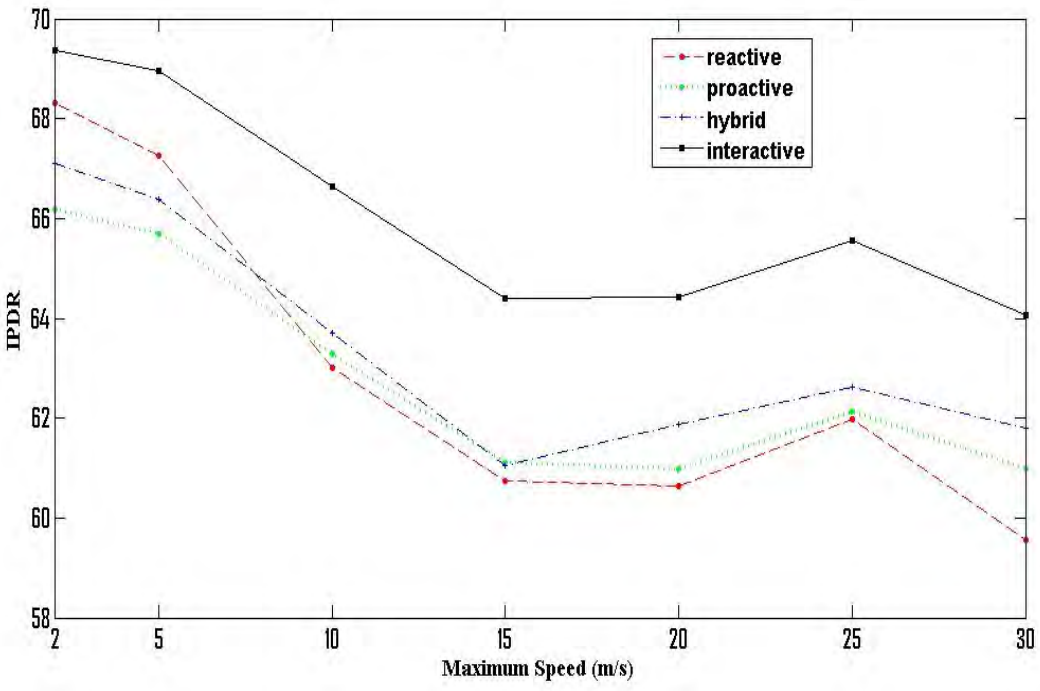


Figure 5.5: IPDR of all schemes against the speed of nodes.

Figure 5.5 shows that IPDR obtained from all the schemes is high at the low speed, i.e. at 2m/s; it starts to decrease with the increase in the speed. The reason behind this fact is that the routing tables of the mobile nodes become obsolete when the nodes move with the high speed. As a result, more packets are dropped by the nodes in the network due to having no routes or obsolete routes to the gateways and the IPDR is reduced.

Our scheme performs better than the other schemes by selecting gateways that have less dense route and the forwarding nodes on the route that have shorter queue lengths.

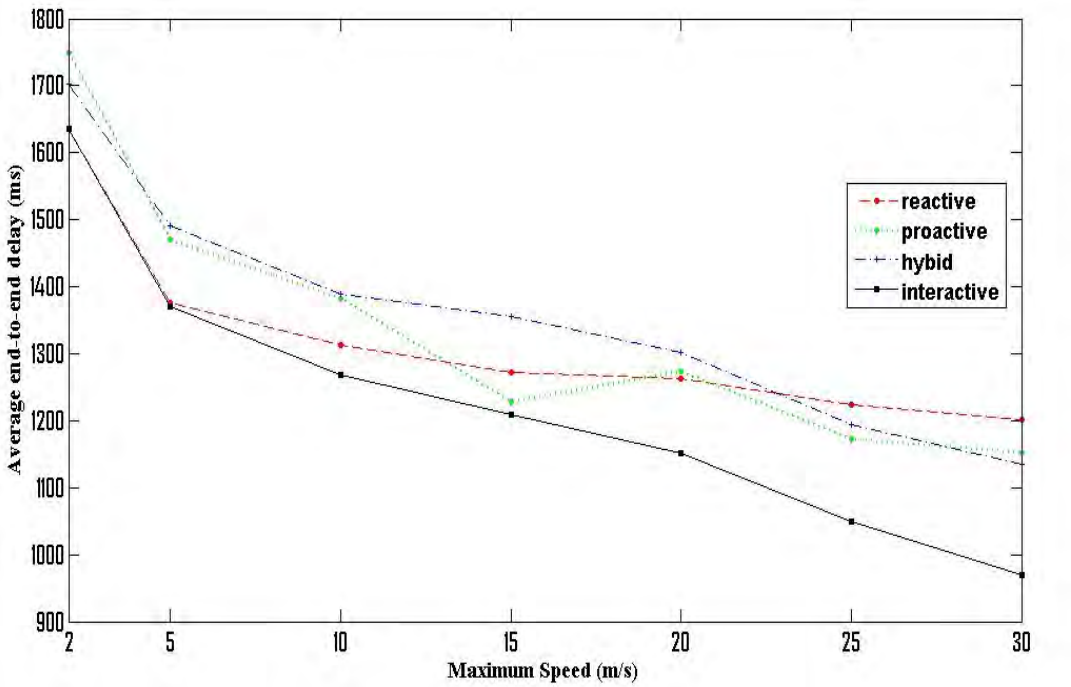


Figure 5.6: Average end-to-end delay of all schemes against the speed of nodes.

We can see from Figure 5.6 that the average end-to-end delay in all the schemes decreases at the higher speeds. At the higher speeds the entries in the routing tables become obsolete quickly. Higher number of packets are dropped in the network for not having the routing entry. This reduces the average length of the interface queue in the network. Because of these shorter queue lengths, packets do not need to wait much in the network to get delivered. Our scheme avoids the routes having longer queue lengths and higher concentration of neighbor nodes. For this reason, our scheme experiences the lowest end-to-end delay.

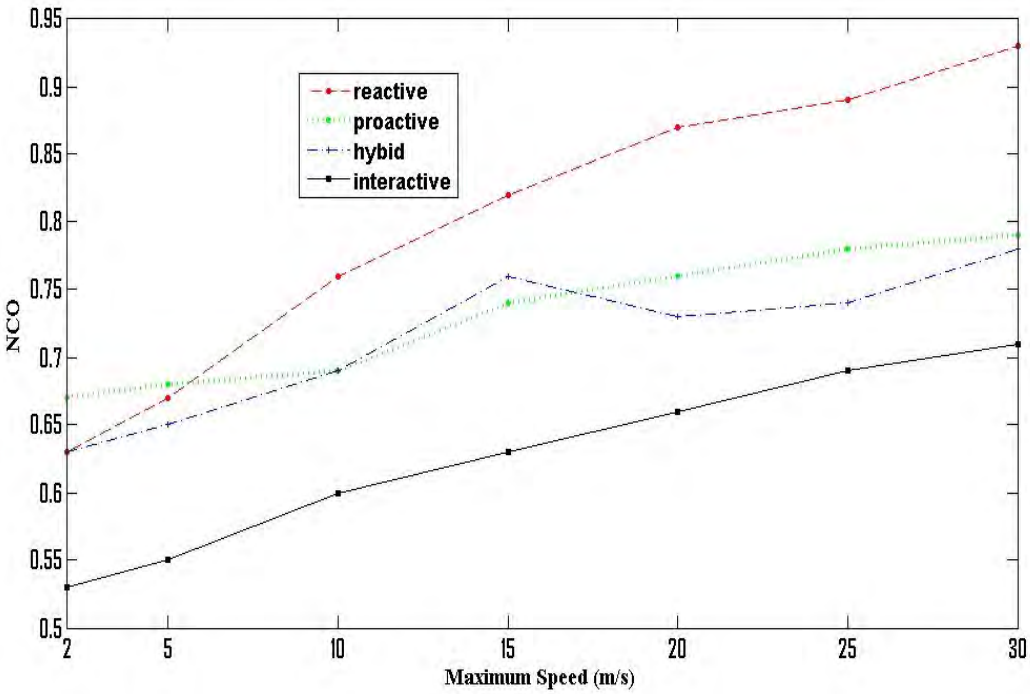


Figure 5.7: NCO of all schemes against the speed of nodes.

Figure 5.7 shows that NCO, which is the ratio between the number of routing packets and the number of packets successfully delivered, increases with the speed of the mobile nodes in every scheme. Since the routing tables of the mobile nodes become obsolete when the nodes move with the high speed, nodes in the network suffer from having no routes or obsolete routes to the gateways. This fact causes more packet drops and more route re-discoveries. As a result NCO of all the schemes increases as the speed of the mobile nodes increases. However, our scheme has less NCO than that of the other schemes because it has less packet dropouts and it requires less routing packets compared to that of the other schemes.

From the above analysis of the results, we can conclude that our gateway discovery and selection scheme performs better than all other existing schemes in terms of packet delivery ratio, end-to-end delay, and network overhead with different size of MANET and with different speed of mobile nodes in the MANET. Thus, the proposed gateway discovery and selection scheme will scale well with the number of nodes, the traffic load and the speed of the nodes.

5.4 Statistical Analysis of the Simulation Results

We perform a statistical test to show that our scheme provides significant performance improvement over the other schemes.

5.4.1 The T test

The t test evaluates whether the means of two groups are significantly different from each other. It calculates a t value using the difference in the means and variances of the two groups as follows.

t value = (**Difference between the group means**) / (**Variability of the groups**)

The greater the t value, the more the significance in the differences between two means. If the two groups being compared have a low degree of variance i.e. if the denominator in the above formula is low, then there is a higher chance that the two groups are distant. There are many variations of the t test. Each has its own specific formula for calculating a t value for the sampled data points.

5.4.1.1 Paired Two-Sample T test

The paired two-sample t test is used to find whether the "before" and "after" means of the samples have changed during an experiment. Here, a t value is calculated from the data points of both the samples. The formula to calculate the t value for a paired two-sample t test is given in eq. (2):

$$\mathbf{t} = \frac{\overline{\mathbf{M}}}{\left(\frac{\mathbf{S}}{\sqrt{\mathbf{n}}}\right)} \qquad \textbf{(2)}$$

$$\overline{M} = \frac{\sum_{i=1}^{n}\left(X_i - Y_i\right)}{n} = \frac{\sum_{i=1}^{n} Z_i}{n}$$

$$\mathbf{S} = \sqrt{\frac{\sum_{\mathbf{i}=1}^{\mathbf{n}}\left(\overline{\mathbf{M}} - \mathbf{Z_i}\right)^2}{\mathbf{n}-1}}$$

Where **X** and **Y** are two groups of data points taken from the same objects “before” and “after” of an experiment respectively and **n** is the sample size.

Null Hypothesis The null hypothesis states that there is no significant difference between the means of two groups "before" and "after" of an experiment.

Alternative Hypothesis It states that difference between the means of two groups "before" and "after" of an experiment are significantly different.

5.4.1.2 Level of Significance

Once the t value is calculated we need to look up in a table of significance to test whether the t value is large enough so that we can conclude that the difference between the means of two groups is significant. To test the significance, a risk level called the alpha level is set. A significance level of 0.05 or 0.01 is normal. For example, if 0.05 (5%) significance level is chosen to make a decision then we would accept five false results out of hundred results. Ninety five times out of hundred times, we would get true results i.e. we are 95% confident that we have made the right decision.

5.4.1.3 One or Two-Tailed Test

A t test can be a one-tailed test or a two-tailed test. A one-tailed test is used when we are interested to test the hypothesis that one scheme is better than another scheme. However, in case of two-tailed test we test whether one scheme is better or worse than the other.

5.4.2 Our T test

We use the paired two-sample two-tailed t test to determine whether the improvement in the performance metrics i.e. IPDR, average end-to-end delay, and NCO in our scheme is significantly better than that of the reactive scheme. We compare two schemes in each data points given in the figures from Figure 5.2 to 5.7. Since each data point is the average of ten simulation run results, we simply measure the results of the reactive and interactive schemes in each run as the before and the after means respectively in order to get our t test results.

In our t test, the level of significance (**alpha**) is 0.05, the sample size (**n**) is equal to 10 and the degrees of freedom (**df**) is equal to (**n** – 1) = 9.

From the standard table of significance in [32], we create a partial t table given in Table 5.2 to interpret the results of our t test.

Table 5.2: A partial T table

	df = 9, **alpha** = 0.05
Critical t value ($\mathbf{T_{critical}}$) for two tailed t test	2.262157

We compare the t values ($\mathbf{T_{value}}$) obtained from the t tests with the critical t value ($\mathbf{T_{critical}}$) to determine whether there is a significant difference between the “reactive” and our “interactive” schemes. If a $\mathbf{T_{value}}$ is greater than the $\mathbf{T_{critical}}$ then we reject the null hypothesis and if a $\mathbf{T_{value}}$ is smaller than the $\mathbf{T_{critical}}$ then we accept the null hypothesis.

5.4.2.1 T test for IPDR

To perform the t test on IPDR of Figure 5.2 and 5.5, our null hypothesis and the alternate hypothesis are as follows:

H_0: The two means of the IPDR of the reactive and our interactive schemes are not significantly different.

H_a: The two means of the IPDR of the reactive and our interactive schemes are significantly different.

T test results on IPDR of Figures 5.2 and 5.5 are given in Tables 5.3 and 5.4 respectively.

Table 5.3: T test results on IPDR in Figure 5.2

Speed (m/s)	**Schemes**	**Mean**	**Variance**	**T_{value}**	**T_{value} - $T_{critical}$**	**Remarks**
2	reactive	68.313	35.14162	2.252716	-0.00944	accept H_0
	interactive	69.384	36.42932			
5	reactive	67.25	24.86528	3.899695	1.637538	reject H_0
	interactive	68.965	17.72547			
10	reactive	63.004	10.56805	6.589359	4.327202	reject H_0
	interactive	66.634	13.93272			
15	reactive	60.745	7.684339	14.07569	11.81353	reject H_0
	interactive	64.4	7.034222			
20	reactive	60.653	9.665534	8.291875	6.029718	reject H_0
	interactive	64.421	4.958988			
25	reactive	61.981	9.239654	8.397143	6.134986	reject H_0
	interactive	65.551	5.000868			
30	reactive	59.55	13.85924	10.7013	8.439143	reject H_0
	interactive	64.061	9.760157			

Table 5.4: T test results on IPDR in Figure 5.5

No. of nodes	Scheme	Mean	Variance	T stat	T_{value} - $T_{critical}$	Remarks
10	reactive	86.075	24.62596	3.278258	1.016101	reject H_0
	interactive	87.402	18.21993			
12	reactive	85.181	8.546321	4.332188	2.070031	reject H_0
	interactive	86.304	6.662849			
14	reactive	85.692	12.84706	7.06506	4.802903	reject H_0
	interactive	87.891	13.06741			
16	reactive	84.01	5.872689	2.43274	0.170583	reject H_0
	interactive	85.701	8.154557			
18	reactive	80.374	12.1318	4.28	2.02	reject H_0
	interactive	84.404	10.37456			
20	reactive	76.882	10.8134	8.44477	6.18	reject H_0
	interactive	80.584	7.326204			
22	reactive	75.261	9.802143	7.111317	4.84916	reject H_0
	interactive	78.303	7.697712			
24	reactive	72.8	4.1636	13.83014	11.56798	reject H_0
	interactive	76.284	4.032316			
26	reactive	67.614	9.775161	5.8816	3.619443	reject H_0
	interactive	71.259	8.515699			
28	reactive	63.924	6.573827	6.61801	4.355853	reject H_0
	interactive	66.891	4.941062			
30	reactive	59.55	13.85924	10.7013	8.439143	reject H_0
	interactive	64.061	9.760157			

From Tables 5.3 and 5.4, we see that the difference between the $\mathbf{T_{value}}$ and the $\mathbf{T_{critical}}$ is positive for most of the cases (we reject the null hypothesis), i.e., our interactive scheme provides higher IPDR than the reactive scheme for most of the cases with a confidence level 95%.

5.4.2.2 T test for Average end-to-end delay

To perform the t test on average end-to-end delay of Figures 5.3 and 5.6, our null hypothesis and the alternate hypothesis are as follows:

H_0: The two means of the delay of the reactive and our interactive schemes are not significantly different.

H_a: The two means of the delay of the reactive and our interactive schemes are significantly different.

T test results on average end-to-end delay of Figures 5.3 and 5.6 are given in Tables 5.5 and 5.6 respectively.

Table 5.5: T test results on average end-to-end delay in Figure 5.3

Speed (m/s)	**Scheme**	**Mean**	**Variance**	**T stat**	**T_{value} - $T_{critical}$**	**Remarks**
2	reactive	1635.244	111306.2	0.00015	-2.26201	accept H_0
	interactive	1635.79	102155.4			
5	reactive	1375.798	168763.5	0.073861	-2.1883	accept H_0
	interactive	1370.436	118051.5			
10	reactive	1312.123	59996.24	0.878159	-1.384	accept H_0
	interactive	1269.379	63131.87			
15	reactive	1271.642	33214.52	1.243227	-1.01893	accept H_0
	interactive	1209.519	43261.78			
20	reactive	1263.175	39371.22	3.049494	0.787337	reject H_0
	interactive	1151.85	28371.26			
25	reactive	1223.248	12716.83	3.932532	1.670375	reject H_0
	interactive	1050.294	24786.76			
30	reactive	1201.496	33353.4	8.879538	6.617381	reject H_0
	interactive	970.025	21784.99			

Table 5.6: T test results on average end-to-end delay in Figure 5.6

No. of nodes	Scheme	Mean	Variance	T stat	T_{value} - $T_{critical}$	Remarks
10	reactive	106.402	905.7582	1.694366	-0.56779	accept H_0
	interactive	99.414	1294.542			
12	reactive	125.905	1023.625	3.23106	0.968903	reject H_0
	interactive	110.932	1015.462			
14	reactive	167.781	771.3989	7.235256	4.973099	reject H_0
	interactive	125.581	617.7584			
16	reactive	248.456	5355.004	4.078866	1.816709	reject H_0
	interactive	165.631	902.1428			
18	reactive	333.069	3229.179	4.859085	2.60	reject H_0
	interactive	260.976	3166.544			
20	reactive	461.234	4367.612	4.626608	2.36	reject H_0
	interactive	320.915	3509.737			
22	reactive	507.213	3056.937	2.085241	-0.17692	accept H_0
	interactive	426.488	9503.946			
24	reactive	568.126	2542.479	1.885491	-0.37667	accept H_0
	interactive	499.204	7242.416			
26	reactive	779.692	20805.49	5.547239	3.285082	reject H_0
	interactive	613.58	15840.75			
28	reactive	982.681	7900.833	3.393468	1.131311	reject H_0
	interactive	853.974	16227.58			
30	reactive	1201.496	33353.4	8.879538	6.617381	reject H_0
	interactive	970.025	21784.99			

From Tables 5.5 and 5.6, we see that our interactive scheme provides lower average end-to-end delay than the reactive scheme for most of the cases with a confidence level 95%.

5.4.2.3 T test for NCO

To perform the t test on NCO of Figures 5.4 and 5.7, our null hypothesis and the alternate hypothesis are as follows:

H_0: The two means of the NCO of the reactive and our interactive schemes are not significantly different.

H_a: The two means of the NCO of the reactive and our interactive schemes are significantly different.

T test results on NCO of Figures 5.4 and 5.7 are given in Tables 5.7 and 5.8 respectively.

Table 5.7: T test results on NCO in Figure 5.4

Speed (m/s)	**Scheme**	**Mean**	**Variance**	**T stat**	**T_{value} -$T_{critical}$**	**Remarks**
2	reactive	0.634	0.01785	4.03458	1.772423	reject H_0
	interactive	0.526	0.006582			
5	reactive	0.68	0.036778	4.24356	1.981403	reject H_0
	interactive	0.553	0.012779			
10	reactive	0.764	0.008649	7.38037	5.118213	reject H_0
	interactive	0.6	0.005756			
15	reactive	0.826	0.021404	5.95741	3.695253	reject H_0
	interactive	0.632	0.008951			
20	reactive	0.871	0.013877	7.81772	5.555563	reject H_0
	interactive	0.658	0.002929			
25	reactive	0.895	0.011072	7.70752	5.445363	reject H_0
	interactive	0.694	0.002761			
30	reactive	0.930	0.019662	6.30933	4.047173	reject H_0
	interactive	0.708	0.002529			

Table 5.8: T test results on NCO in Figure 5.7

No. of nodes	Scheme	Mean	Variance	T stat	T_{value} -$T_{critical}$	Remarks
10	reactive	0.248	0.000262	2.44949	0.187333	reject H_0
	interactive	0.244	0.000227			
12	reactive	0.251	0.000143	0.317999	-1.94416	accept H_0
	interactive	0.252	0.000196			
14	reactive	0.261	0.000521	1.86052	-0.40164	accept H_0
	interactive	0.256	0.000293			
16	reactive	0.278	0.000596	2.75085	0.488693	reject H_0
	interactive	0.265	0.000428			
18	reactive	0.321	0.00061	6.81516	4.55E+00	reject H_0
	interactive	0.287	0.000357			
20	reactive	0.353	0.001312	4.30187	2.04E+00	reject H_0
	interactive	0.324	0.000804			
22	reactive	0.374	0.000671	0.58277	-1.67939	reject H_0
	interactive	0.37	0.000778			
24	reactive	0.449	0.001588	3.8512	1.589043	reject H_0
	interactive	0.403	0.001446			
26	reactive	0.559	0.004766	3.53363	1.271473	reject H_0
	interactive	0.474	0.003329			
28	reactive	0.724	0.007604	7.96496	5.702803	reject H_0
	interactive	0.583	0.002934			
30	reactive	0.930	0.019662	6.30933	4.047173	reject H_0
	interactive	0.708	0.002529			

From Tables 5.5 and 5.6 it is evident that our interactive scheme provides lower NCO than that of the reactive scheme for most of the cases with a confidence level 95%. All the t test results prove that our scheme is significantly better than the reactive scheme in terms of packet loss, end-to-end delay, and network overhead.

5.5 Summary

In this chapter, we described our simulation scenarios and parameters. The performance from several simulation runs were compared and analyzed. Simulation results ensure the superiority of our scheme over the proactive, reactive, and hybrid schemes. We also performed t test to show that our scheme is statistically significant and better than the other schemes. We conclude our thesis in the next chapter.

Chapter 6

Conclusion and Future Works

Integration of mobile ad hoc network and the Internet allows ubiquitous Internet services for mobile users in a MANET. In this thesis, at first, we discussed the advantage and disadvantage of the current Internet gateway discovery and selection schemes. Proactive scheme requires considerable overheads but achieves good connectivity and lower delay because nodes instantly know better routes to gateways. In contrast, reactive scheme suffers from longer delay and lower packet delivery ratio but it achieves low routing overhead. The hybrid schemes minimize the disadvantages of reactive and proactive scheme but it needs an intelligent adaptation of the optimal proactive area. Only hop count is used in the existing schemes to select the gateway which forces the nodes in the network to select the closest gateway always. The closest gateways might have huge traffic than that of the other gateways and can turn into as a bottleneck to Internet traffic. To rescue the network from the problems of current Internet gateway discovery and selection schemes, we proposed a new gateway discovery and selection scheme. Our scheme uses a triggered broadcast of gateway advertisement messages at the gateways when hit by gateway discovery messages. We also bounded the dissemination of the gateway advertisement messages up to the requesting mobile node from the gateway. We combined hop count, traffic load (interface queue length), and the total number of neighbors along a route to the gateway in order to formulate a new metric for gateway selection. Our metric chooses the gateway which is not only closest but also has the route from the mobile node with less load and less dense.

We compared our gateway discovery and selection scheme with the other schemes in terms of three performance metrics: Internet Packet Delivery Ratio, Average End-to-End Delay and Normalized Control Overhead. Simulation results show that our scheme outperforms other schemes.

6.1 Future Works

A number of open issues remain. In this work, we consider the gateways to be stationary. In a hybrid environment, it is very likely that there will be a mixture of stationary and mobile gateways. Therefore, mobility of the gateways is an important issue in the gateway discovery and selection process and needs to be considered with due diligence. Thus, our next task is to develop a gateway discovery and selection scheme considering both stationary and mobile gateways.

In our present research work, we considered much higher Internet bandwidth for the gateways compared to that of the MANET. However, higher Internet bandwidth might not be available at the gateways and it might be a serious bottleneck for the Internet traffic of the MANET. In our future work, we can also consider the actual Internet bandwidth available at the gateways while selecting the best gateway.

In our present research work, we allowed the gateway to broadcast gateway advertisement message when it is being hit by a gateway discovery message without considering the current traffic load at the gateway. If the current load is higher and new Internet traffic is directed towards this gateway by a gateway selection algorithm at the MANET nodes which does not consider the current traffic at the gateway, the new Internet traffic at the heavily loaded gateway might increase serious congestion in the network. In our future work, we will consider the current traffic load at the gateways to select the best gateway.

References

[1] Wakikawa R., Malinen J. T., Perkins C. E., Nilsson A., Tuominen A. J., "Global Connectivity for IPv6 Mobile Ad Hoc Networks," *IETF Internet Draft*, draft-wakikawa-manet-globalv6-05.txt, March 2006.

[2] Sun Y., Belding-Royer E. M., Perkins C. E., "Internet Connectivity for Ad Hoc Mobile Networks," *International Journal of Wireless Information Networks, Special Issue on Mobile Ad Hoc Networks (MANETs): Standards, Research, Applications*, pp. 75-88, April 2002.

[3] Jonsson U., Alriksson F., Larsson T., Johansson P., Maguire Jr. G. Q., "MIPMANET – Mobile IP for Mobile Ad Hoc Networks," *Proc. of the First IEEE/ACM Annual Workshop on Mobile Ad Hoc Networking and Computing*, USA, pp. 75-85, August 2000.

[4] Kumar R., Misra M., Sarje A. K., "A Proactive Load-Aware Gateway Discovery in Ad Hoc Networks for Internet Connectivity," *International Journal of Computer Networks & Communications (IJCNC)*, vol. 2, no.5, pp. 120-139, September 2010.

[5] Engelstad P. E., Tønnesen A., Hafslund A., Egeland G. "Internet Connectivity for Multi-Homed Proactive Ad Hoc Networks," *Proc. of the International Conference on Communication*, France, pp. 4050-4056, June 2004.

[6] Khan K. U. R., Reddy A. V., Zaman R. U., Kumar M. "An Effective Gateway Discovery Mechanism in an Integrated Internet-MANET (IIM)," *Proc. of the International Conference on Advances in Computer Engineering*, India, pp. 24-28, June 2010.

[7] Yuste A. J., Triviño A., Trujillo F. D., Casilari E., "Improved Scheme for Adaptive Gateway Discovery in Hybrid MANET," *Proc. of the International Conference on Distributed Computing Systems Workshops*, Genova, pp. 270-275, June 2010.

[8] Kim Y., Ahn S., Yu H., Lee J., Lim Y., "Proactive Internet Gateway Discovery Mechanisms for Load-Balanced Internet Connectivity in MANET," *Proc. of International Conference on Information Networking. Towards Ubiquitous Networking and Services*, Portugal, pp. 285-294, January 2007.

[9] Hamidian A. A., "A Study of Internet Connectivity for Mobile Ad Hoc Networks in NS 2," *Master's Thesis*, Department of Communication Systems, Lund Institute of Technology, Lund University, January 2003.

[10] Le-Trung Q., Engelstad P. E., Skeie T., Taherkordi A., "Load-Balance of Intra/Inter-MANET Traffic over Multiple Internet Gateways," *Proc. of the International Conference on Advances in Mobile Computing and Multimedia*, Austria, pp. 50-57, November 2008.

[11] Li X. , Li Z., "A MANET Accessing Internet Routing Algorithm based on Dynamic Gateway Adaptive Selection," *Frontiers of Computer Science in China*, vol. 4 , no. 1, pp. 143-150, March 2010.

[12] Zhanyang X., Xiaoxuan H., Shunyi Z., "A Scheme of Multipath Gateway Discovery and Selection for MANET Using Multi-Metric," *Proc. of the International Conference on Information Science and Engineering*, China, pp. 2500-2503, December 2009.

[13] Lee J., Kim D., Garcia-Luna-Aceves J. J., Choi Y., Choi J., Nam S., "Hybrid Gateway Advertisement Scheme for Connecting Mobile Ad Hoc Networks to the Internet," *Proc. of the 57th IEEE Vehicular Technology Conference*, Korea , pp. 191-195, April 2003.

[14] Ratanchandani P., Kravets R., "A Hybrid Approach to Internet Connectivity for Mobile Ad Hoc Networks," *Proc. of IEEE Wireless Communications and Networking Conference*, USA, pp. 1522-1527, March 2003.

[15] Ruiz P. M., Gomez-Skarmeta A. F., “Adaptive Gateway Discovery Mechanisms to Enhance Internet Connectivity for Mobile Ad Hoc Networks,” *Ad Hoc and Sensor Wireless Networks*, vol. 1, pp. 159-177, March 2005.

[16] Jiang H., Jin S., “Adaptive Strategies for Efficiently Locating Internet-based Servers in MANETs,” *Proc. of ACM International Workshop on Modeling, Analysis, and Simulation of Wireless and Mobile Systems*, Canada, pp. 341-348, October 2005.

[17] Bin S., Bingxin S., Bo L., Zhonggong H., Li Z., “Adaptive Gateway Discovery Scheme for Connecting Mobile Ad Hoc Networks to the Internet,” *Proc. of International Conference on Wireless Communications, Networking and Mobile Computing*, China, vol. 2, pp. 795-799, September 2005.

[18] Zhuang L., Liu Y., Liu K., Zhai L., Yang M., “An Adaptive Algorithm for Connecting Mobile Ad Hoc Network to Internet with Unidirectional Links Supported,” *The Journal of China Universities of Posts and Telecommunications*, vol. 17, supplement 1, pp. 44-49, July 2010.

[19] Park B., Lee W., Lee C., “QoS-aware Internet access schemes for wireless mobile ad hoc networks,” *Computer Communications*, vol. 30, issue 2, pp. 369-384, January 2007.

[20] Perkins C. E., Das S. R., “Ad hoc On-Demand Distance Vector (AODV) Routing,” *IETF Internet draft*, draft-ietf-manet-aodv-09.txt, November 2001.

[21] Nurthy C. S. R., Manoj B. S., “Ad hoc Wireless Networks: Architectures and Protocols,” *Prentice Hall*, 2004.

[22] Jacquet P., Muhlethaler P., Qayyum A., “Optimized Link State Routing Protocol,” *IETF Internet Draft*, draft-ietf-manet-olsr-00.txt, November 1998.

[23] Perkins C. E., Bhagwat P., “Highly Dynamic Destination-Sequenced Distance-Vector Routing (DSDV) for Mobile Computers,” *Proc. of the ACM Conference on Communications Architectures, Protocols and Applications*, UK, pp. 234-244, 1994.

[24] Johnson D.B., Maltz D.A., “Dynamic Source Routing in Ad Hoc Wireless Networks,” *In Mobile Computing, Kluwer Academic Publishers*, chapter 5, pp. 153-181, 1996.

[25] Haas Z. J.; Pearlman M. R.; Samar P., "The Zone Routing Protocol (ZRP) for Ad Hoc Networks, *IETF Internet Draft*, draft-ietf-manet-zone-zrp-04.txt, July 2002.

[26] Ma W., Liu J., "A Gateway Selection Scheme for Internetworking of MANET and Internet using Improved Genetic Algorithm," *Proceedings of the International Conference on Wireless Communications, Networking and Mobile Computing*, USA, pp. 2668-2671, September 2009.

[27] Ruiz P. M., Ros F. J., Gomez-Skarmeta A. F., "Internet Connectivity for Mobile Ad Hoc Networks: Solutions and Challenges," *IEEE communication Magazine*, vol. 43, issue 10, pp. 118-125, October 2005.

[28] Belding-Royer E. M., Sun Y., Perkins C. E., "Global Connectivity for IPv4 Mobile Ad Hoc Networks," *IETF Internet Draft*, draft-royer-manet-globalv4-00.txt, November 2001.

[29] NS-2 home page http://www.isi.edu/nsnam/ns/index.html.

[30] AODV+: The Network Simulator: Contributed Code, available from: <http://www.isi.edu/nsnam/ns/ns-contributed.html/>.

[31] Hyytiä E., Koskinen H., Lassila P., Penttinen A., Virtamo J., "Random Waypoint Model in Wireless Networks," *Networks and Algorithms: Complexity in physics and Computer Science* , Helsinki, June 2005.

[32] Nehmzow U., "Statistical Tools for Describing Experimental Data," *Robot Behaviour: Design, Description, Analysis and Modelling*, Chapter 4, Springer, 2009.