A SECURE
ZONE-BASED ROUTING PROTOCOL
FOR
MOBILE AD HOC NETWORKS

A Thesis Report Submitted in Partial Fulfillment

of the Requirements for the Degree of

Master of Technology

in

Computer Science (Information Security)

By

Niroj Kumar Pani

Department of Computer Science and Engineering

National Institute of Technology, Rourkela

Rourkela, Orissa 769008, India

May 2009

A SECURE
ZONE-BASED ROUTING PROTOCOL
FOR
MOBILE AD HOC NETWORKS

A Thesis Report Submitted in Partial Fulfillment
of the Requirements for the Degree of

Master of Technology
in
Computer Science (Information Security)
By

Niroj Kumar Pani

Under the Guidance of

Dr. Ashok Kumar Turuk

Department of Computer Science and Engineering
National Institute of Technology, Rourkela
Rourkela, Orissa 769008, India

May 2009

National Institute Of Technology
Rourkela

CERTIFICATE

This is to certify that the thesis entitled, **"A Secure Zone-Based Routing Protocol for Mobile Ad Hoc Networks"** submitted by **Mr. Niroj Kumar Pani** in partial fulfillment of the requirements for the award of Master of Technology Degree in **Computer Science and Engineering** with specialization in **"Information Security"** at National Institute of Technology, Rourkela is an authentic work carried out by him under my supervision and guidance.

To the best of my knowledge, the matter embodied in the thesis has not been submitted to any other University / Institute for the award of any Degree or Diploma.

Date:

Dr. Ashok Kumar Turuk
Assistant Professor
Department of Computer Science and Engineering
National Institute of Technology
Rourkela-769008

- Dedicated to My Family

ACKNOWLEDGEMENTS

This project is by far the most significant accomplishment in my life and it would be impossible without the people who supported me and believed in me.

I would like to express my sincere heartfelt gratitude to my honorable, esteemed supervisor Dr. Ashok Kumar Turuk, Assistant Professor, Department of Computer Science and Engineering for his kind and valuable guidance for the completion of the thesis work. His consistent support and intellectual guidance inspired me to innovate new ideas. I am glad to work under his supervision

I am grateful to Dr. B. Majhi, Professor and Head, Department of Computer Science and Engineering for his excellent support during my work. I am also thankful to Dr. S. K. Jena, Dr. S. K. Rath, Dr. D. P. Mohapatra, Dr. R. Baliarsingh and Prof. B. D. Sahoo of CSE Department, for providing me support and advice in preparing my thesis work.

Thanks all my friends and classmates for their love and support. I have enjoyed their company so much during my stay at NIT, Rourkela. I would like to thank all those who made my stay in Rourkela, an unforgettable and rewarding experience.

Last, but not least I would like to thank my parents for supporting me to do complete my master’s degree in all ways.

Niroj Kumar Pani

CONTENTS

LIST OF FIGURES

LIST OF TABLES

ABBREVIATIONS

MANET	Mobile Ad Hoc Network
MAC	Medium Access Control
DARPA	Defense Advanced Research Projects Agency
DSDV	Destination-Sequenced Distance-Vector Routing
CGSR	Clustered Gateway Switch Routing
WRP	Wireless Routing Protocol
OLSR	Optimized Link State Routing
DSR	Dynamic Source Routing
AODV	Ad hoc On Demand Distance Vector Routing
TORA	Temporally Ordered Routing Algorithm
ABR	Associativity Based Routing
ZRP	Zone Routing Protocol
ZHLS	Zone-Based Hierarchal Link State Routing Protocol
MRP	Multipoint Relay
RREQ	Route Request
RREP	Route Reply
IARP	Intra-Zone Routing Protocol
IERP	Inter-Zone Routing Protocol
NDP	Neighbor Discovery Protocol
BRP	Bordercast Resolution Protocol
CA	Certification Authority
CN	Common Node
ARAN	Authenticated Routing for Ad Hoc Networks
RDP	Route Discovery Packet
REP	Reply Packet
SZRP	Secure Zone Routing Protocol

KMP	Key Management Protocol
SIARP	Secure Intra-Zone Routing Protocol
SIERP	Secure Inter-Zone Routing Protocol
MBRP	Modified Boarder Resolution Protocol
SKREQ	Session Key Request Packet
SKREP	Session Key Reply Packet
SRD	Secure Route Discovery Packet
SRR	Secure Route Reply Packet
ERR	Error Packet

ABSTRACT

An ad hoc network is the cooperative engagement of a collection of mobile nodes without the required intervention of any centralized access point or existing infrastructure. There is an increasing trend to adopt ad hoc networking for commercial uses; however, their main applications lie in military, tactical and other security-sensitive operations. In these and other applications of ad hoc networking, secure routing is an important issue.

Designing a foolproof security protocol for ad hoc network is a challenging task due to its unique characteristics such as, lack of central authority, frequent topology changes, rapid node mobility, shared radio channel and limited availability of resources. A number of protocols have been proposed in the literature for secure routing. However, most of these protocols are either proactive or reactive in approach. Both the approaches have their own limitations, for example, the proactive protocols use excess bandwidth in maintaining the routing information while, the reactive ones have long route request delay.

In this thesis, we proposed a secure hybrid ad hoc routing protocol, called Secure Zone Routing Protocol (SZRP), which aims at addressing the above limitations by combining the best properties of both proactive and reactive approaches. The proposed protocol is based on the concept zone routing protocol (ZRP). It employs an integrated approach of digital signature and both the symmetric and asymmetric key encryption techniques to achieve the security goals like message integrity, data confidentiality and end to end authentication at IP layer. The thesis details the design of the proposed protocol and analyses its robustness in the presence of multiple possible security attacks that involves impersonation, modification, fabrication and replay of packets caused either by an external advisory or an internal compromised node within the network. The security and performance evaluation of SZRP through simulation indicates that the proposed scheme successfully defeats all the identified threats and achieves a good security at the cost of acceptable overhead. Together with existing approaches for securing the physical and MAC layer within the network protocol stack, the Secure Zone Routing Protocol (SZRP) can provide a foundation for the secure operation of an ad hoc network.

Chapter 1

INTRODUCTION

Motivation

Objective

Organization of Thesis

Chapter 1

Introduction

In this new era of communication, the advent of mobile computing has revolutionized our information society. The proliferation of new, powerful, efficient and compact communicating devices like personnel digital assistants (PDAs), pagers, laptops and cellular phones, having extraordinary processing power paved the way for advance mobile connectivity. We are moving from the Personal Computer age to the Ubiquitous Computing age in which a user utilizes, at the same time, several electronic platforms through which he can access all the required information whenever and wherever needed. The nature of ubiquitous devices makes wireless networks the easiest solution for their interconnection and, as a consequence, the wireless arena has been experiencing exponential growth in the past decade [4].

Among the myriad of applications and services run by mobile devices, network connections and corresponding data services are without doubt the most demanding ones. Currently, most of the connections among the wireless devices are achieved via fixed infrastructure-based service provider, or private networks. For example, connections between two cell phones are setup by BSC and MSC in cellular networks; laptops are connected to Internet via wireless access points. While infrastructure-based networks provide a great way for mobile devices to get network services, it takes time and potentially high cost to set up the necessary infrastructure. There are, furthermore, situations where user required networking connections are not available in a given geographic area, and providing the needed connectivity and network services in these situations becomes a real challenge.

For all these reasons, combined with significance advances in technology and standardization, new alternative ways to deliver mobile connectivity have been emerging. These are focused around having the mobile devices connect to each other in the transmission range through automatic configuration, setting up an *ad hoc mobile network* that is both flexible and powerful.

A mobile ad hoc network (MANET) sometimes called a *wireless ad hoc network* or a *mobile mesh network* is a wireless network, comprised of mobile computing devices (nodes) that use wireless transmission for communication, without the aid of any established infrastructure or centralized administration such as a base station or an access point [1, 2, 3, 4]. Unlike traditional mobile wireless networks, mobile ad hoc networks do not rely on any central coordinator but communicate in a self organized way. Mobile nodes that are within each other's radio range communicate directly via wireless links, while those far apart rely on other nodes to relay messages as routers. In ad hoc network each node acts both as a host (which is capable of sending and receiving) and a router which forwards the data intended for some other node. Ad hoc wireless networks can be deployed quickly anywhere and anytime as they eliminate the complexity of infrastructure setup.

Applications of ad hoc network range from military operations and emergency disaster relief, to commercial uses such as community networking and interaction between attendees at a meeting or students during a lecture. Most of these applications demand a secure and reliable communication.

Mobile wireless networks are generally more vulnerable to information and physical security threats than fixed wired networks. Vulnerability of channels and nodes, absence of infrastructure and dynamically changing topology, make ad hoc networks security a difficult task [4]. Broadcast wireless channels allow message eavesdropping and injection (vulnerability of channels). Nodes do not reside in physically protected places, and hence can easily fall under the attackers' control (node vulnerability). The absence of infrastructure makes the classical security solutions based on certification authorities and

on-line servers inapplicable. In addition to this, the security of routing protocols in the MANET dynamic environment is an additional challenge.

Most of the previous research on ad hoc networking has been done focusing only upon the efficiency of the network. There are quite a number of routing protocols proposed [5, 16, 21] that are excellent in terms of efficiency. However, they were generally designed for a non-adversarial network setting, assuming a trusted environment; hence no security mechanism has been considered. But in a more realistic setting such as a battle field or a police rescue operation, in which, an adversary may attempt to disrupt the communication; a secure ad hoc routing protocol is highly desirable.

The unique characteristics of ad hoc networks present a host of research areas related to security, such as, key management models, secure routing protocols, instruction detection systems and trust based models. This thesis work is based on the research done in the area of secure routing.

1.1 Motivation

Secure routing in the field of mobile ad hoc networks is one of the most emerging areas of research. Designing a foolproof security protocol for ad hoc routing is a challenging task due to the unique network characteristics such as, lack of central authority, rapid node mobility, frequent topology changes, insecure operational environment, shared radio channel and limited availability of resources. A number of protocols have been proposed in the literature for secure routing. A survey of these protocols is given in [1, 2, 22, 23]. Most of these protocols are either proactive or reactive in approach. However, both the approaches have their own limitations [16, 17]. For example, the proactive protocols use excess bandwidth in maintaining the routing information while, the reactive ones have long route request delay. Reactive routing also inefficiently floods the entire network for route determination.

In this thesis, we proposed a secure hybrid ad hoc routing protocol, called Secure Zone Routing Protocol (SZRP), which aims at addressing the above limitations by combining the best properties of both proactive and reactive approaches. The proposed protocol is based on the concept zone routing protocol (ZRP). It employs an integrated approach of digital signature and both the symmetric and asymmetric key encryption techniques to achieve the security goals like message integrity, data confidentiality and end to end authentication at IP layer. The thesis details the design of the proposed protocol and analyses its robustness in the presence of multiple possible security attacks that involves impersonation, modification, fabrication and replay of packets caused either by an external advisory or an internal compromised node within the network.

The security and performance evaluation of SZRP through simulation indicates that the proposed scheme successfully defeats all the identified threats and achieves a good security at the cost of acceptable overhead. Together with existing approaches for securing the physical and MAC layer within the network protocol stack, the Secure Zone Routing Protocol (SZRP) can provide a foundation for the secure operation of an ad hoc network.

1.2 Objective

The goal of this thesis is two fold.

- It aims towards suggesting, designing and implementing a highly efficient security solution for mobile ad hoc networks by establishing secure routing and effective key management mechanism.

- The proposed protocol should be built upon such a platform that it is not only efficient in terms of meeting the security requirements like message integrity, data confidentiality and end to end authentication but is also cost effective and applicable in practical environment.

1.3 Organization of Thesis

The thesis is divided into six chapters. Chapter 1, which is here, gives some introduction and motivation for proposing the Secure Zone Routing Protocol. In Chapter 2, we look at mobile ad hoc networking in closer details, covering their specific characteristics, complexities and design constraints. This is followed by a classification of existing routing algorithms in it. Chapter 3 examines the security issues and challenges associated with mobile ad hoc networks. In this chapter, we identify the different kinds of threats an ad hoc network faces and explore new approaches to secure its communication. In Chapter 4, we detail the design of Secure Zone Routing Protocol and analyze its robustness in the presence of multiple security attacks. Chapter 5 presents the possible implementation and performance evaluation of the proposed protocol through simulation work. We conclude the thesis in Chapter 6 with proposal for possible extension of the work done.

Chapter 2

AD HOC NETWORKING

Mobile Ad Hoc Networks

MANET Applications

Routing approaches in Mobile Ad Hoc Network

Conclusion

Chapter 2

Ad Hoc Networking

Mobility is becoming increasingly important for users of computing systems. Technology has made possible smaller, less expensive and more powerful wireless communicating devices and computers. As a result users gain flexibility and the ability to exchange information and maintain connectivity while roaming through a large area. The necessary mobile computing support is being provided in some areas by installing base stations and access points. Mobile users can maintain their connectivity by accessing this infrastructure from home, from the office, or while on the road.

Such mobility support is not available in all locations where mobile communication is desired. Access points may not be set up due to high cost, low expected usage, or poor performance. This may happen during outdoor conferences or in emergency situations like natural disasters and military maneuvers in enemy territory. If mobile users want to communicate in the absence of a support structure, they must form an *ad hoc network.* In this chapter, we look at mobile ad hoc networking in closer details. We present their characteristics, analyze the complexities and design constraints associated with them and classify the existing routing algorithms in it.

2.1 Mobile Ad Hoc Networks

A mobile ad hoc network (MANET), sometimes called a *wireless ad hoc network* or a *mobile mesh network* is a wireless network, comprised of mobile computing devices (nodes) that use wireless transmission for communication, without the aid of any

established infrastructure or centralized administration such as a base station in cellular network or an access point in wireless local area network [1, 2, 3, 4]. The nodes are free to move randomly and organize themselves arbitrarily; thus, the network's wireless topology may change rapidly and unpredictably. Such a network may operate in a standalone fashion, or may be connected to the larger Internet. Unlike traditional mobile wireless networks, mobile ad hoc networks do not rely on any central coordinator but communicate in a self organized way. Mobile nodes that are within each other's radio range communicate directly via wireless links, while those far apart rely on other nodes to relay messages as routers. In ad hoc network each node acts both as a host (which is capable of sending and receiving) and a router which forwards the data intended for some other node. Hence it is appropriate to call such networks as "multi-hop wireless ad hoc networks". Figure 2.1 shows an example of mobile adhoc network and its communication technology.

As shown in Figure 2.1, an ad hoc network might consist of several home-computing devices, including laptops, cellular phones, and so on. Each node will be able to communicate directly with any other node that resides within its transmission range. For communicating with nodes that reside beyond this range, the node needs to use intermediate nodes to relay the messages hop by hop.

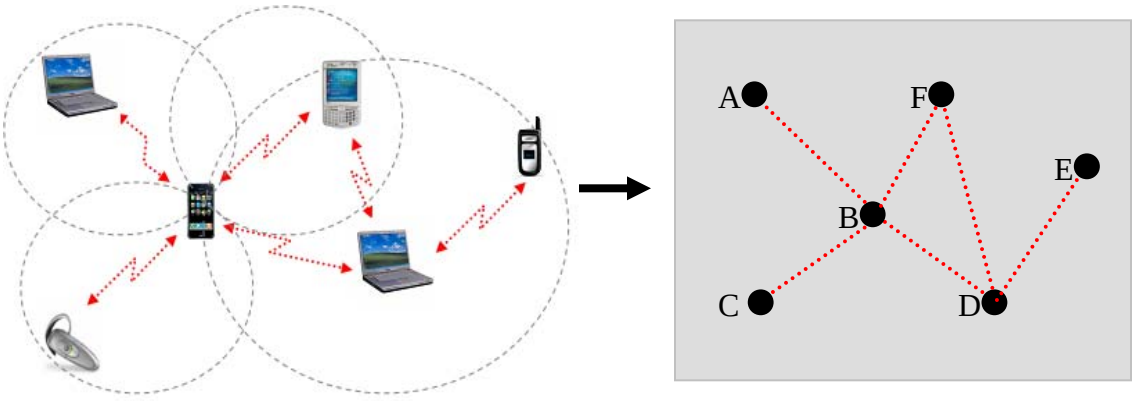


Fig 2.1: A Typical Mobile Ad Hoc Network

2.1.2 Characteristics, Complexities and Design Constraints

Mobile ad hoc networks eliminate the constraint of infrastructure set up and enable devices to create and join networks on the fly, any where, any time and virtually for any application. However, these flexibilities and convenience do come at a price. Mobile ad hoc networks inherit the common problems of wireless networking in general, and add their own constraints specific to ad hoc routing [2]. Some of the notable characteristics, complexities and design constraints of MANETs are presented below:

- **Wireless medium:** In an ad hoc environment, nodes communicate wirelessly and share the same media (radio, infrared etc.). The wireless medium has neither absolute, nor readily observable boundaries outside of which the stations are unable to receive network frames. Thus the channel is unprotected from outside signals and hence it is significantly less reliable than wired media.

- **Autonomous and infrastructureless:** MANET does not depend on any established infrastructure or centralized administration. Each node operates in distributed peer-to-peer mode, acts as an independent router and generates independent data. Network management has to be distributed across different nodes, which brings added difficulty in fault detection and management

- **Dynamic and changing network topology:** In mobile ad hoc networks, because nodes can move arbitrarily, the network topology, which is typically multi-hop, can change frequently and unpredictably, resulting in route changes, frequent network partitions, and possibly packet losses**.**

- **Limited availability of resources:** Because batteries carried by each mobile node have limited power supply, processing power is limited, which in turn limits services and applications that can be supported by each node. This becomes a bigger issue in MANET because, since each node is acting as both an end system and a router at the same time, additional energy is required to forward packets.

2.2 MANET Applications

Ad hoc wireless networks, due to their quick and economically less demanding deployment, find applications in several areas [1]. Some of these include:

- Military applications, such as establishing communication among a group of soldiers for tactical operations when setting up a fixed wireless communication infrastructure in enemy territories or in inhospitable terrains may not be possible.
- Emergency systems, for example, establishing communication among rescue personnel in disaster-affected area that need quick deployment of a network.
- Commercial uses such as community networking and interaction between attendees at a meeting or students during a lecture
- Collaborative and distributed computing.
- Wireless mess networks and wireless sensor networks.

2.3 Routing approaches in Mobile Ad Hoc Network

Since the advent of Defense Advanced Research Projects Agency (DARPA) packet radio networks in the early 1970s [1], numerous routing protocols have been developed for ad hoc mobile networks [2, 5]. As shown in Fig. 2.2, these are generally categorized as table-driven or proactive, on-demand or reactive and hybrid routing protocols.

Table-driven or Proactive Protocols: Proactive routing protocols attempt to maintain consistent, up-to-date routing information between every pair of nodes in the network by propagating, proactively, route updates at fixed intervals. As the resulting information is usually maintained in tables, the protocols are sometimes referred to as table-driven protocols. Representative proactive protocols include: Destination-Sequenced Distance-Vector (DSDV) routing [7], Clustered Gateway Switch Routing (CGSR) [8], Wireless Routing Protocol (WRP) [9], and Optimized Link State Routing (OLSR) [10].

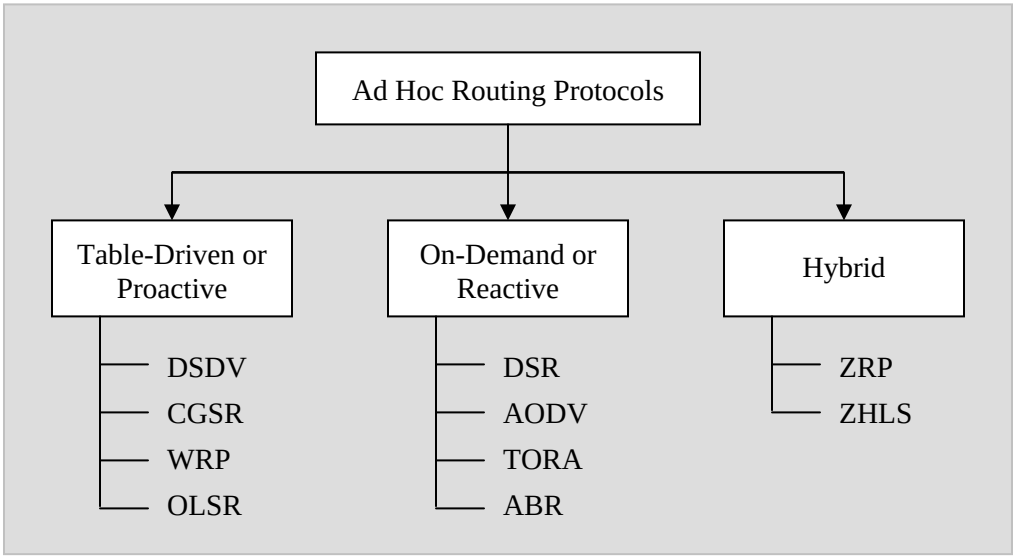


Fig 2.2: Classifications of Ad Hoc Routing Protocols

On-demand or Reactive Protocols: A different approach from table-driven routing is reactive or on-demand routing. These protocols depart from the legacy Internet approach. Reactive protocols, unlike table-driven ones, establish a route to a destination when there is a demand for it, usually initiated by the source node through discovery process within the network. Once a route has been established, it is maintained by the node until either the destination becomes inaccessible or until the route is no longer used or has expired. Representative reactive routing protocols include: Dynamic Source Routing (DSR) [12], Ad hoc On Demand Distance Vector (AODV) routing [13], Temporally Ordered Routing Algorithm (TORA) [14] and Associativity Based Routing (ABR) [15].

Hybrid Routing Protocols: Purely proactive or purely reactive protocols perform well in a limited region of network setting. However, the diverse applications of ad hoc networks across a wide range of operational conditions and network configuration pose a challenge for a single protocol to operate efficiently [3]. For example, reactive routing protocols are well suited for networks where the call-to-mobility ratio is relatively low. Proactive routing protocols, on the other hand, are well suited for networks where this ratio is relatively high. The performance of either class of protocols degrades when the protocols are applied to regions of ad hoc networks space between the two extremes.

Researchers advocate that the issue of efficient operation over a wide range of conditions can be addressed by a *hybrid* routing approach, where the proactive and the reactive behavior is mixed in the amounts that best match these operational conditions. Representative hybrid routing protocols include: Zone Routing Protocol (ZRP) [16] and Zone-based Hierarchal Link state routing protocol (ZHLS) [21].

In the following sub sections we examine three protocols, the optimized link state routing (OLSR) protocol, adhoc on-demand distance vector (AODV) routing protocol and zone routing protocol (ZRP), as they were found useful for the thesis work. OLSR and AODV fall under proactive and reactive family, where as, ZRP is a hybrid routing protocol.

2.3.1 Optimized Link State Routing (OLSR) Protocol

The Optimized Link State Routing (OLSR) protocol [10] is a variation of traditional link state routing, modified for improved operation in ad hoc networks. The key feature of OLSR is its use of *multipoint relays* (MPRs) to reduce the overhead of network floods and the size of link state updates. Each node computes its MPRs from its set of neighbours. The MPR set is selected such that when a node broadcasts a message, the retransmission of that message by the MPR set will ensure that the message is received by each of its two-hop neighbours. Hence, when ever a node broadcasts a message, only those neighbours in its MPR set rebroadcast the message. Other neighbours that are not in the MPR set process the message but not rebroadcast it. Further, when exchanging link state routing information, a node only lists its connections to those neighbours that have selected it as an MPR. That set of neighbours is termed as *MPR Selectors*.

The MPR set for a given node is the set of neighbours that covers the two-hop neighbourhood of the node, as shown in Figure 2.3. Nodes learn their set of two-hop neighbours through the periodic exchange of *Hello messages*. Each node periodically transmits a Hello message that contains a list of neighbours. Associated with each neighbour is an attribute including the directionality of the link to that neighbour. The node is labeled *symmetric* if the link to the neighbour is bidirectional, or *asymmetric* if a

Hello has been received from that node but the link has not been confirmed as bidirectional. When a node receives this Hello message from each of its neighbours, it obtains complete knowledge of its two-hop neighbour set at that point in time. Further, if its own address is listed in the Hello message, it knows the link with that neighbour is bidirectional. It can then update the status of that neighbour to be symmetric.

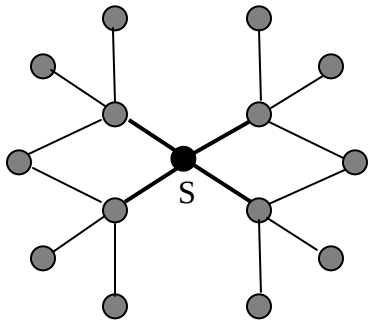


Fig 2.3: Multipoint Relays

The MPRs set may be calculated according to the algorithm [11] as follows. Each node starts with an empty MPR set. The N is defined to be the set of one-hop neighbours with which there exists bidirectional connectivity and the set of N_2 is the set of two-hop bidirectional neighbours. The first nodes that are selected for the MPR set are those nodes in N that are the only neighbours of some node in N_2. Next, the degree of each node n in N that is not in the MPR set is calculated., where the degree is the number of nodes in N_2 that are not covered by nodes in the MPR set, the node in N that has the highest degree is included in the MPR set. Once all the nodes in N_2 are covered, the process terminates.

Once each node's MPR set is selected, routing path within the network can be determined. Because OLSR is a proactive protocol, each node maintains a route to every other node in the network. To diffuse topology information, nodes periodically exchange topology control (YC) messages with their neighbours. The TC message for a given node lists the set of neighbours that have selected the sending node as an MPR. This is called the *multi point relay selector* set of the node. Only this set of nodes is advertised within the network. As a node receives TC messages from other network nodes, it can create or

modify routing entries to each node in the network using any shortest path routing algorithm, such as a variation of Dijkstra's algorithm.

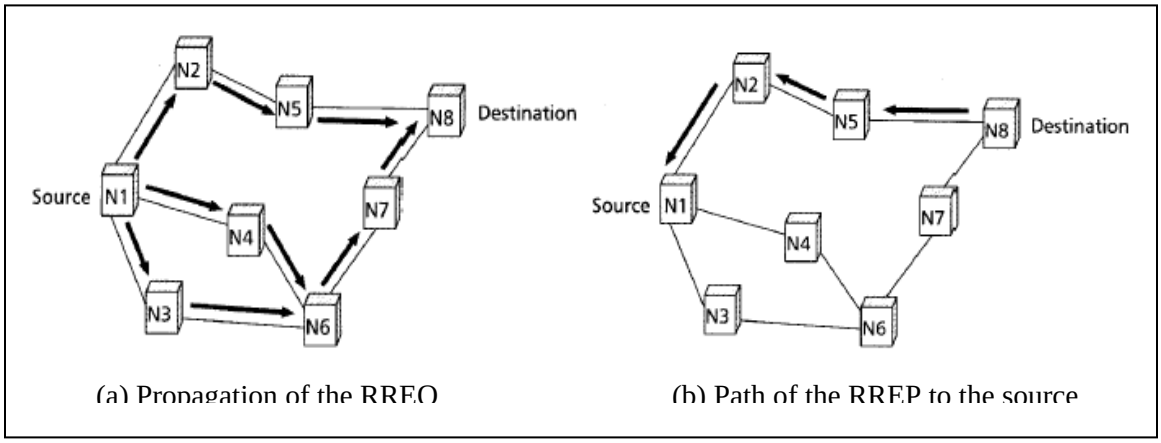
2.3.2 Ad hoc On Demand Distance Vector (AODV) routing

The Ad Hoc On-Demand Distance Vector (AODV) routing protocol [13] is based on the DSDV algorithm described in [7]. AODV is an improvement on DSDV because it typically minimizes the number of required broadcasts by creating routes on a demand basis, as opposed to maintaining a complete list of routes as in the DSDV algorithm. The authors of AODV classify it as a pure on-demand route acquisition system, since nodes that are not on a selected path do not maintain routing information or participate in routing table exchanges.

When a source node desires to send a message to some destination node and does not already have a valid route to that destination, it initiates a path discovery process to locate the other node. It broadcasts a route request (RREQ) packet to its neighbors, which then forward the request to their neighbors, and so on, until either the destination or an intermediate node with a "fresh enough" route to the destination is located. Figure 2.4a illustrates the propagation of the broadcast RREQs across the network. AODV utilizes destination sequence numbers to ensure all routes are loop-free and contain the most recent route information. Each node maintains its own sequence number, as well as a broadcast ID. The broadcast ID is incremented for every RREQ the node initiates, and together with the node's IP address, uniquely identifies an RREQ. Along with its own sequence number and the broadcast ID, the source node includes in the RREQ the most recent sequence number it has for the destination. Intermediate nodes can reply to the RREQ only if they have a route to the destination whose corresponding destination sequence number is greater than or equal to that contained in the RREQ.

During the process of forwarding the RREQ packets, intermediate nodes record in their route tables the address of the neighbor from which the first copy of the broadcast packet is received, thereby establishing a reverse path. If additional copies of the same RREQ

are later received, these packets are discarded. Once the RREQ reaches the destination or an intermediate node with a fresh enough route, the destination or the intermediate node responds by unicasting a route reply (RREP) packet back to the neighbor from which it first received the RREQ (Fig. 2.4b). As the RREP packet is routed back along the reverse path, nodes along this path set up forward route entries in their route tables which point to the node from which the RREP packet came. These forward route entries indicate the active forward route. Associated with each route entry is a route timer which will cause the deletion of the entry if it is not used within the specified lifetime. Because the RREP packet is forwarded along the path established by the RREQ packet, AODV only supports the use of symmetric links in the network.



(a) Propagation of the RREQ (b) Path of the RREP to the source

Fig 2.4: AODV route discovery

Routes are maintained by AODV as follows. If a source node moves, it is able to reinitiate the route discovery protocol to find a new route to the destination. If a node along the route moves, its upstream neighbor notices the move and propagates a link failure notification message (an RREP packet with infinite metric) to each of its active upstream neighbors to inform them of the erasure of that part of the route [13]. These nodes in turn propagate the link failure notification to their upstream neighbors, and so on until the source node is reached. The source node may then choose to reinitiate route discovery for that destination if a route is still desired.

An additional aspect of the protocol is the use of hello messages, periodic local broadcasts by a node to inform each mobile node of other nodes in its neighborhood. Hello messages can be used to maintain the local connectivity of a node. Nodes listen for retransmission of data packets to ensure that the next hop is still within reach. If such a retransmission is not heard, the node may use any one of a number of techniques, including the reception of hello messages, to determine whether the next hop is within communication range. The hello messages may list the other nodes from which a mobile has heard, thereby yielding greater knowledge of network connectivity.

2.3.3 Zone Routing Protocol (ZRP)

As explained earlier, either a purely proactive or purely reactive approach to implement a routing protocol for a MANET has their disadvantages. The Zone Routing Protocol (ZRP) as described in [16] aims at addressing these limitations by combining the best properties of both proactive and reactive approaches and hence it can be classed as a hybrid proactive/reactive routing protocol.

In a MANET it can be safely assumed that most communication takes palace between nodes close to each other. Therefore, ZRP reduces the proactive scope to a z*one* centered on each node and reactive approach outside the zone. When a node has a data packet for a particular destination, it checks whether the destination is within its zone or not. If it is within the zone, the packet is routed proactively. Reactive routing is used if the destination is outside the zone.

A *zone* (routing zone) of a node is nothing but the area of local neighbourhood of that node. The “size” of a zone is not determined by geographical measurement, as one might expect, but is given by a radius of length β where, β is the number of hops to the perimeter of the zone. Each node may be within multiple overlapping zones, and each zone may be of a different size. An example routing zone is shown in Figure 2.5, where the routing zone of S includes the nodes A–I, but not K. In the illustrations, the radius is marked as a circle around the node in question.

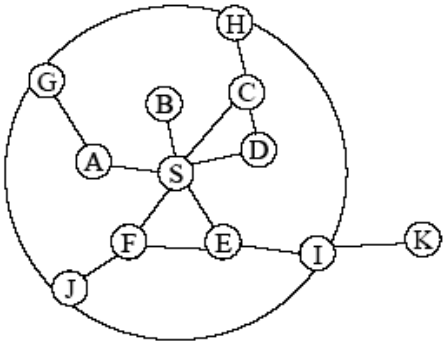


Fig 2.5: Routing zone of node S with zone radius β =2

The nodes of a zone are divided into peripheral nodes and interior nodes. Peripheral nodes are nodes whose minimum distance to the central node is exactly equal to the zone radius β. The nodes whose minimum distance is less than β are interior nodes. In Figure 2.5, the nodes A–F are interior nodes, the nodes G–J are peripheral nodes and the node K is outside the routing zone. Note that node H can be reached by two paths, one with length 2 and one with length 3 hops. The node is however within the zone, since the shortest path is less than or equal to the zone radius.

ZRP refers to the locally proactive routing component as the IntrA-zone Routing Protocol (IARP) [18]. The globally reactive routing component is named IntEr-zone Routing Protocol (IERP) [19]. IERP and IARP are not specific routing protocols. Instead, IARP is a family of limited-depth, proactive link-state routing protocols like OLSR (Refer Section 2.3.1). It periodically computes the route to all intrazone nodes (nodes that are within the routing zone of a node) and maintains this information in a data structure called IARP routing table. Correspondingly, IERP is a family of reactive routing protocols like DSR [12] or AODV (Refer Section 2.3.2) that offer enhanced route discovery and route maintenance services based on local connectivity monitored by IARP.

Since IARP employees a proactive link state routing protocol for maintaining intrazone routing information, the first thing which becomes necessary for IARP is to know about the neighbours of a node. In order to learn about a node's direct neighbors and possible

link failures, IARP relies on a Neighbor Discovery Protocol (NDP) provided by the MAC layer. NDP transmits “HELLO” beacons at regular intervals. Upon receiving a beacon, the neighbor table [18] is updated. Neighbors, for which no beacon has been received within a specified time, are removed from the table

For route discovery by IERP, the notion *bordercasting* [20] is introduced. Bordercasting utilizes the topology information provided by IARP to direct query request to the border of the zone. The bordercast packet delivery service is provided by the Bordercast Resolution Protocol (BRP) [20]. BRP uses a map of an extended routing zone to construct bordercast trees for the query packets. BRP employs query control mechanisms, to direct route requests away from areas of the network that already have been covered. [20]. The relationship between the components is illustrated in Figure 2.6.

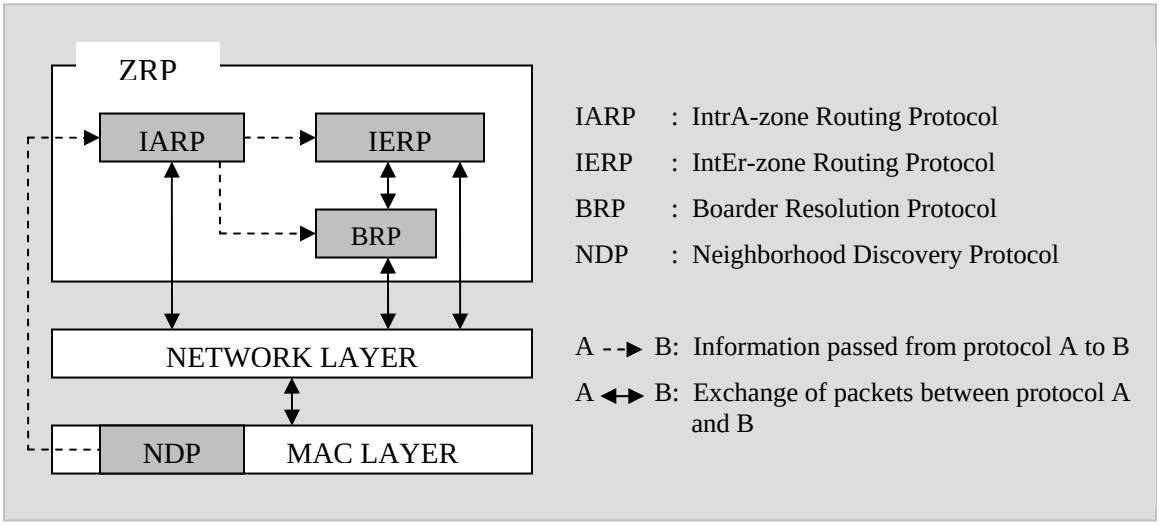


Fig 2.6: ZRP Architecture

Routing: A node that has a packet to send first checks whether the destination is within its local zone or not using information provided by IARP routing table. If the destination is within the zone, then the IARP routing table must have a valid route to the destination. So in this case, the packet is routed proactively to the intrazone destination. Reactive routing is used if the destination is outside the zone.

The reactive routing process is divided into two phases: the route request phase and the route reply phase. In the route request, the source sends a route request packet to its peripheral nodes using BRP. If the receiver of a route request packet knows the destination, it responds by sending a route reply back to the source. Otherwise, it continues the process by bordercasting the packet. In this way, the route request spreads throughout the network. If a node receives several copies of the same route request, these are considered as redundant and are discarded [19, 20].

The reply is sent by any node that can provide a route to the destination. To be able to send the reply back to the source node, routing information must be accumulated when the request is sent through the network. The information is recorded either in the route request packet (source routing approach [12]), or as next-hop addresses in the nodes along the path similar to AODV. In the first case, the nodes forwarding a route request packet append their address and relevant node/link metrics to the packet. When the packet reaches the destination, the sequence of addresses is reversed and copied to the route reply packet. The sequence is used to forward the reply back to the source. In the second case, the forwarding nodes records routing information as next-hop addresses, which are used when the reply is sent to the source. This approach can save transmission resources, as the request and reply packets are smaller [19].

2.4 Conclusion

In this chapter, we presented mobile ad hoc networks as a new paradigm for wireless communication. We identified the characteristics, complexities and design constraints associated with them, discussed some of their deployment scenarios and classify the existing routing algorithms in it. In the next chapter, we look into ad hoc networking from security view point. We identify the different kinds of security attacks an ad hoc network faces and explore new approaches to secure its communication.

Chapter 3

SECURITY IN MOBILE AD HOC NETWORKS

Security goals
Issues and challenges in security provisioning
Security attacks on ad hoc routing protocols
Security mechanisms and solutions
Secure routing
Observation

Chapter 3

Security in Mobile Ad Hoc Networks

Wireless mobile ad hoc nature of MANET brings new security challenges to network design. Mobile adhoc networks, due to their unique characteristics, are generally more vulnerable to information and physical security threats than wired networks or infrastructure-based wireless networks. In this chapter, we explore the various security requirements (goals) for wireless ad hoc network and the different types of threats an ad hoc network faces. We identify the new challenges and opportunities posed by this new networking environment and explore new approaches to secure its communication.

3.1 Security Goals

To secure an ad hoc network, a security protocol must satisfy the following attributes: confidentiality, integrity, availability, authenticity and non-repudiation [22, 23].

Confidentiality ensures that classified information in the network is never disclosed to unauthorized entities. Sensitive information, such as strategic military decisions or location information requires confidentiality. Leakage of such information to enemies could have devastating consequences.

Integrity guarantees that a message being transferred between nodes is never altered or corrupted. Data can be altered either intentionally by malicious nodes in the network or accidentally because of benign failures, such as radio propagation impairment or through hardware glitches in the network.

Availability implies that the requested services (e.g. bandwidth and connectivity) are available in a timely manner even though there is a potential problem in the system. Availability of a network can be tempered for example by dropping off packets and by resource depletion attacks.

Authenticity is a network service to determine a user's identity. Without authentication, an attacker can impersonate any node, and in this way, one by one node, it can gain control over the entire network.

Finally, ***non-repudiation*** ensures that the information originator cannot deny having sent the message. Non-repudiation is useful for detection and isolation of compromised nodes.

3.2 Issues and Challenges in Security Provisioning

Designing a foolproof security protocol for ad hoc routing is a very challenging task due its unique characteristics such as, shared radio channel, insecure operational environment, lack of central authority and association rules among nodes and limited availability of resources [1]. A brief discussion on how each of the above mentioned characteristics causes difficulty in providing security in ad hoc wireless network is given below.

- **Shared radio channel:** Unlike the wired networks where a separate dedicated transmission line can be provided between a pair of end users, the radio channel used for communication in ad hoc networks is broadcast in nature and shared by all nodes in the network. Data transmitted by a node is received by all the nodes within its direct transmission range. So a malicious node can easily obtain data being transmitted in the network.

- **Insecure operational environment:** The operational environment in which MANETs are generally used may not be always secure, for example, a battle field. In such environment, nodes may move in and out of hostile and insecure enemy territory, where they would be highly vulnerable to security attacks.

- **Lack of central authority:** In wired networks or infrastructure based wireless networks it would be possible to monitor the network traffic through routers or base stations and implement security mechanisms at those points. Since MANETs don't have any such central points, these mechanisms can't be applicable to them.

- **Lack of association rules:** In MANET, **s**ince nodes can leave or join the network at any point of time, if no proper authentication mechanism is used for associating nodes with the network intruders can easily join the network and carry out attacks.

- **Limited availability of resources:** Resources such as bandwidth, battery power and computational power are scare in ad hoc networks. Hence, it is difficult to implement complex cryptography-based security mechanisms in such networks.

3.3 Security Attacks on Ad Hoc Routing Protocols

The complexity and uniqueness of MANETs make them more vulnerable to security threats than their wired counterparts. Attacks on ad hoc wireless networks can be classified as *passive* and *active* attacks, depending on whether the normal operation of the network is disrupted or not. [2].

- **Passive attacks:** A passive attack does not disrupt the normal operation of the network; the attacker snoops the data exchanged in the network without altering it. Here the requirement of confidentiality gets violated. Detection of passive attack is very difficult since the operation of the network itself doesn't get affected. One of the solutions to the problem is to use powerful encryption mechanism to encrypt the data being transmitted, there by making it impossible for the attacker to get useful information from the data overheard.

- **Active attacks:** An active attack attempts to alter or destroy the data being exchanged in the network there by disrupting the normal functioning of the network. Active attacks can be *internal* or *external.* External attacks are carried

out by nodes that do not belong to the network. Internal attacks are from compromised nodes that are part of the network. Since the attacker is already part of the network, internal attacks are more severe and hard to detect than external attacks. Active attacks, whether carried out by an external advisory or an internal compromised node involves actions such as impersonation (masquerading or spoofing), modification, fabrication and replication.

Both passive and active attacks can be made on any layer of the network protocol stack [1]. This section however, focuses on network layer attacks only (routing attacks). Depending upon the various attacking behavior routing attacks can be classified into five categories: attacks using information disclosure, impersonation (masquerading or spoofing), modification, fabrication, and replay of packets. Among these information disclosure is a passive attack while the rest fall under the active category.

3.3.1 Information disclosure Attack

In this, a compromised node may leak confidential information to unauthorized nodes in the network. Such information may include information regarding the network topology, geographic location of nodes or, optimal routes to unauthorized nodes in the network [2]. Attacks such as location disclosure and traffic analysis come under this category.

3.3.2 Attacks using Impersonation

In impersonation attacks, the attacker assumes the identity and privileges of an authorized node, either to make use of the network resources that may not be available to it under normal circumstances, or to disrupt the normal functioning of the network by injecting false routing information into the network. Impersonation can be done by several ways. The attacker could by chance guess the identity and authentication details of the authentic node called the target node, or it could snoop for the authentication details of the target node from the previous communication. Some of the impersonation attacks include:

Man-in-the-Middle Attack: In this attack, a malicious node impersonates the receiver with respect to the sender, and the sender with respect to the receiver, without having either of them realize that they have been attacked with an intension to read or modify the messages between two parties [23].

Sybil Attack: In the Sybil attack [25], an attacker pretends to have multiple identities. A malicious node can behaves as if it were a larger number of nodes either by impersonating other nodes or simply by claiming false identities. Sybil attacks are classified into three categories: direct/indirect communication, fabricated/stolen identity, and simultaneity. In the direct communication, Sybil nodes communicate directly with legitimate nodes, whereas in the indirect communication messages sent to Sybil nodes are routed through malicious nodes. An attacker can fabricate a new identity or it can simply steal it after destroying or temporarily disabling the impersonated node. All Sybil identities can participate simultaneously in the network or they may be cycled through.

3.3.3 Attacks using Modification

This attack disrupts the routing function by having the attacker illegally modifying the content of the messages. Examples of such attacks include redirection by changing the route sequence number and redirection with modified hop count. Some of the attacks involving packet modification are given below:

Misrouting Attack: In the misrouting attack, a non-legitimate node redirects the routing message and sends data packet to the wrong destination [33]. This type of attack is carried out by modifying the final destination address of the data packet or by forwarding a data packet to the wrong next hop in the route to the destination.

Byzantine attack: Here a compromised intermediate node or a set of compromised intermediate nodes collectively carries out attacks such as creating routing loops, routing packets on non-optimal paths and selectively dropping packets [26]. Since in such attacks the network would seem to operate normally Byzantine failure are hard to detect.

Denial of service (DoS) attack: In this type of attack, an attacker attempts to prevent legitimate and authorized users of services offered by the network from accessing those services. A DoS attack can be carried out in many ways and against any layer in the network protocol stack. The classic way is to flood packets to any centralized resource used in the network by modifying the routes information in the packets so that the resource is no longer available to nodes in the network, resulting the network no longer operating in the manner it was designed to operate. This may lead to failure in the delivery of guaranteed services to the end users.

3.3.4 Attacks using Fabrication

In fabrication attacks, an intruder generates false routing messages, such as routing updates and route error messages, in order to disturb network operation or to consume other node resources. A number of fabrication based attacks are presented below:

Resource Consumption Attack: In this attack, a malicious node deliberately tries to consume the resources (e.g. battery power, bandwidth, etc.) of other nodes in the network [4]. The attacks could be in the form of unnecessary route request control messages, very frequent generation of beacon packets, or forwarding of stale information to nodes.

***Routing Table or Route Cache Poisoning*:** In this attack, a malicious node sends false routing updates to other uncompromised nodes [1]. Such an attack may result in sub-optimal routing, network congestion or even make some part of the network inaccessible.

Routing table overflow: Here, the attacker advertises routes to non-existing nodes, to the authorized nodes present in the network. The main objective of such an attack is to cause an overflow of the routing table, which would in turn prevent the creation of entries corresponding to new routes to authorized nodes [1]. Proactive routing protocols are more vulnerable to this attack compared to reactive routing protocols.

Rushing Attack: On demand routing protocols that use route discovery process are vulnerable to this type of attack [27]. An attacker node which receives a "route request" packet from the source node floods the packet quickly through out the network before other nodes which also receive the same "route request" packet can react. Nodes that receive the legitimate "route request" packet assume those packets to be the duplicates of the packet already received through the attacker node and hence discard those packets. Any route discovered by the source node would contain the attacker node as one of the intermediate nodes. Hence the source node would not be able to find secure routes.

Black Hole Attack: In this type of attack, a malicious node falsely advertises good path (e.g., shortest path or most stable path) to the destination node during the path finding process [1]. The intension of the malicious nodes could be to hinder the path finding process or to interrupt all the data packets being sent to the concerned destination node.

Gray Hole Attack: Under this attack, an attacker drops all data packets but it lets control messages to route through it [28]. This selective dropping makes gray hole attacks much more difficult to detect then blackhole attack.

3.3.5 Replay Attacks

In the replay attack, an attacker retransmits data to produce an unauthorized effect. Examples of replay attacks are wormhole attack and tunneling attack.

Wormhole Attack: In this attack [29], two compromised nodes can communicate with each other by a private network connection. The attacker can create a vertex cut of nodes in the network by recording a packet at one location in network, tunneling the packet to another location, and replaying it there. The attacker does not require key material as it only needs two transceivers and one high quality out-of-band channel. The wormhole can drop packets or it can selectively forward packets to avoid detection. It is particularly dangerous against different network routing protocols in which the nodes consider themselves neighbor after hearing a packet transmission directly from some node.

Tunneling Attack: In a tunneling attack [33], two or more nodes collaborate and exchange encapsulated messages along existing data routes. For example, if a Route Request packet is encapsulated and sent between two attackers, the packet will not contain the path traveled between the two attackers. This would falsely make the receiver conclude that the path containing the attackers is the shortest path available.

3.4 Security Mechanisms and Solutions

Having seen the various kinds of attacks possible on ad hoc routing, we now look at various techniques employed to overcome these attacks. There can be two types of security mechanisms: preventive and detective. Preventive mechanisms are typically based on message encryption techniques, while detective mechanisms include the application of digital signature and cryptographic hash functions [2].

3.4.1 Message Encryption

Encipherment or message encryption is the science and art of transforming a message into a disguised version which no unauthorized person can read, but which can be recovered in its original form by an intended recipient. In the parlance of cryptography, the original message is called *plaintext* and the secret version of the message is called *ciphertext*. The plaintext is converted into ciphertext by the process of *encryption*, that is, by the use of certain algorithms or functions. The reverse process is called *decryption*. The process of encryption and decryption are governed by *keys*, which are small amount of information used by the cryptographic algorithms.

There are two types of encryption techniques: symmetric key and asymmetric key (or public key). Symmetric key cryptosystem uses the same key (the secret key) for encryption and decryption of a message, where as asymmetric key cryptosystems use one key (the public key) to encrypt a message and another key (the private key) to decrypt it. Public and private keys are related in such a way that only the public key can be used to

encrypt messages and only the corresponding private key can be used for decryption purpose. Even if attacker comprises a public key, it is virtually impossible to deduce the private key. Symmetric key algorithms are usually faster to execute electronically than the asymmetric key algorithms.

3.4.2 Digital signature and Hashing

The process of encryption only ensures the confidentiality of the message being sent. Digital signature is a technique by which one can achieve the other security goals like message integrity, authentication and non-repudiation. In this the sender uses a *signing algorithm* and its *private key to sign* the message. The message and the signature are sent to the receiver. The receiver receives the message and the signature and applies the *verifying algorithm* on the message-signature pair. The verification algorithm requires a *verification key,* which is a public key provided by the signer, to verify the document. After verification if the result is true, the message is accepted; otherwise, it is rejected.

Hashing can be used for the digital signature process especially when the message is long. In this the message is passed through an algorithm called *cryptographic hash function* or *one-way hash function* before signing. It is an algorithm which creates a compressed image of the message in the form of a hash value (or message digest) which is usually much smaller than the message and unique to it. Any change to the message will produce a different hash result even when the same hash function is used.

3.4.3 Key Management Approaches

Both digital signature and encryption mechanisms are key-based approaches. Key distribution and management is therefore at the center of these mechanisms. There are several methods given in [1] that can be employed to perform this operation, all requiring varying amounts of initial configuration, communication and computation. We will however, focus on the method based on public key certificates [30], as we have used this approach in our proposed protocol.

According to this approach, the key management responsibility is shared among a set of trusted certification servers called the certification authorities (CAs). Each CA has a public/private key pair, with its public key known to every node, and signs certificates binding public keys to nodes after verifying their authenticity secretly. The trusted CA has to stay on-line to reflect the current bindings, because the bindings could change over time: a public key should be revoked if the owner node is no longer trusted or is out of the network; a node may refresh its key pair periodically to reduce the chance of a successful brute-force attack on its private key.

3.5 Secure routing

As discussed in Chapter 1, secure routing is highly desirable in ad hoc environment. The previous section pointed out some of the solutions proposed for ensuring security in mobile ad hoc networks. Current research has resulted in a number of secure routing protocols based on these security mechanisms. A survey of the protocols is given in [2, 22, 23] whose details can be found in [33, 34, 35, 36, 37, 38]. In this section we explore the requirements of a secure routing protocol for ad hoc networks and discuss a secure routing protocol called “Authenticated Routing for Ad Hoc Networks (ARAN)”, as we have employed similar security setting as that of ARAN in our proposed protocol.

3.5.1 Requirements for a Secure Routing Protocol

Considering the attacks presented in Section 3.3, we list here the fundamental requisites of a secure routing protocol for mobile ad hoc networks. They are: (1) Routing messages cannot be altered in transit, except according to the normal functionality; (2) Route signaling cannot be spoofed; (3) Fabricated routing messages cannot be injected into the network; (4) Routing loops cannot be formed through malicious action; (5) Routes cannot be redirected from the shortest path by malicious action; (6) Unauthorized nodes should be excluded from route computation and discovery; (7) The network topology must not be exposed by the routing messages either to adversaries or to authorized nodes.

3.5.2 Authenticated Routing for Ad Hoc Networks (ARAN)

Authenticated routing for ad hoc networks (ARAN) [33] is an on-demand protocol designed to provide secure communications in managed open environments. Nodes in a managed-open environment exchange initialization parameters before the start of communication. Session keys are exchanged or distributed through a trusted third party like a certification authority. Each node in ARAN receives a certificate after securely authenticating its identity to a trusted certificate server *T*. Nodes use these certificates to authenticate themselves to other nodes during the exchange of routing messages. The certificate contains the node's IP address, its public key, as well as the time of issuing and expiration. These fields are concatenated and signed by the server *T*. A node *A* receives a certificate as: $T \rightarrow A$: $cert_A = [IP_A, K_{A+,} t, e] K_{T-}$

In the authentication phase, ARAN ensures the existence of a secure path to the destination. Each intermediate node in the network stores the route pair (previous node, the destination node). All the fields are concatenated and signed with source node *I*'s private key. A combination of the nonce number (N_I) and timestamp (t) is used to obtain data freshness and timeliness property. Each time *I* performs a route discovery, it monotonically increases the nonce. The signature prevents spoofing attacks that may alter the route or form loops. Source node *I* broadcasts a route discovery packet (RDP) for a destination *D* as: $I \rightarrow brdcst:$ [RDP, IPD, $cert_I$, N_I, t] K_{I-}.

Each node that receives the RDP for the first time removes any other intermediate node's signature, signs the RDP using its own key, and broadcasts it to all its neighboring nodes. This continues until destination node *D* eventually receives the packet. After receiving the RDP, the destination node *D* sends a reply (REP) packet back along the reverse path to the source node *I*. If *J* is the first node on the reverse path, REP packet is sent as:
$D \rightarrow J:$ [REP, IP_I, $cert_D$, N_I, t] K_{D-}

The source node *I* on receiving the REP packet, verifies the destination's signature K_{D-} and the nonce N_I. When there is no traffic on an existing route for some specific time,

then that route is deactivated in the routing table. Nodes use an ERR message to report links in active routes broken due to node movement.

Using pre-determined cryptographic certificates, ARAN provides network services like authentication and non-repudiation. Simulations show that ARAN is efficient in discovering and maintaining routes but routing packets are larger in size and overall routing load is high. Due to heavy asymmetric cryptographic computation, ARAN has higher cost for route discovery. It is not immune to wormhole attack and if nodes do not have time synchronization, then it is prone to replay attacks as well.

3.6 Conclusion

As research on developing a secure communication system for mobile ad hoc networks had matured, a number of secure routing protocols were proposed [33, 34, 35, 36, 37, 38]. In this chapter, we have described one of these protocols, ARAN. Like ARAN, most of the secure protocols are either proactive or reactive in nature. Studies reveal that, either a purely proactive or a purely reactive protocol performs well in a limited region of network setting. However, in diverse applications of ad hoc networks, the performance of either class of protocols degrades dramatically. For example, the reactive routing protocols are well suited for networks where the call-to-mobility ratio is relatively low, but as they have long route request delay, they are not ideal for an environment where this ratio is relatively high. Proactive routing protocols, on the other hand, are favorable for networks having high node mobility, however in a reverse environment they perform inefficiently, as, they use excess bandwidth in maintaining the routing information.

Researchers advocate that the issue of efficient operation over a wide range of conditions can be addressed by a *hybrid* routing approach, where the proactive and the reactive behavior is mixed in the amounts that best match these operational conditions. In the next chapter we proposed such a hybrid secure routing protocol for ad hoc networks that will address the limitations of both proactive and reactive routing approaches.

Chapter 4

THE SECURE ZONE ROUTING PROTOCOL (SZRP)

Protocol Overview
Certification process
Design of Secure Zone Routing Protocol
The Secure Routing Algorithm
Proactive route computation
Route maintenance
Analysis of Secure Zone Routing Protocol
Conclusion

Chapter 4

The Secure Zone Routing Protocol (SZRP)

Neither a pure proactive nor a pure reactive approach provides a complete solution for secure ad hoc routing that performs efficiency across a wide range of operational conditions and network configuration. So a complete, efficient and applicable solution for secure routing is highly desirable that can operate well on diverse applications of ad hoc networks. This chapter presents the proposed solution, called the Secure Zone Routing Protocol (SZRP), which aims towards addressing this issue. We have detailed the design of the proposed protocol and analyzed its robustness in diverse networking environment, in the presence of multiple possible security attacks.

4.1 Protocol Overview

The Secure Zone Routing Protocol (SZRP) is based on the concept of Zone Routing Protocol (ZRP) [16, 17]. It is a hybrid routing protocol that combines the best features of both proactive and reactive approaches and adds its own security mechanisms to perform secure routing. The reasons for selecting ZRP as the basis of our protocol are as follows: (i) ZRP is based on the concept of routing zones, a restricted area, and it is more feasible to apply the security mechanisms within a restricted area than in a broader area that of the whole network, (ii) Since the concept of zones separate the communicating nodes in terms of interior (nodes within the zone) and exterior (nodes outside the zone) nodes, certain information like network topology and neighbourhood information etc. can be hidden to the exterior nodes, (iii) Incase of a failure, it can be restricted to a zone.

Like ZRP the proposed protocol performs routing in terms of intrazone [18] and interzone [19] routing. It limits the proactive scope within a zone centered on each node and the reactive approach outside the zone. However, it differs from ZRP in security aspects. In ZRP where there is no security consideration, SZRP is designed to address all measure security concerns like end to end authentication, message/packet integrity and data confidentiality during both intra and inter-zone routing. For end to end authentication and message/packet integrity RSA digital signature mechanism [24] is employed, where as data confidentiality is ensured by an integrated approach of both symmetric and asymmetric key encryption [24].

Packets are signed and/or encrypted (either using symmetric or asymmetric key approach) depending upon their type i.e. whether the packet is a control or a data packet. Most of the control packets are only signed. However, all the data packets and those control packets that contain any secret information like a session key between the source and destination node and are signed as well as encrypted. Since the control packets are small in size they are encrypted using the asymmetric key approach. As the data packets are generally long and symmetric key approach is faster than the asymmetric key encryption we encrypt all the data packets using the symmetric key approach.

Each communicating node has two pairs of private/public keys, one pair for signing and verifying and the other for encrypting and decrypting. For a node X the signing and verifying keys are SK_X and VK_X respectively while, encrypting and decrypting keys are EK_X and DK_X respectively. Among these keys SK_X and DK_X are private keys whereas VK_X and EK_X are public keys. Notations used in our proposed protocol are given in Table 4.1.

The secure zone routing protocol (SZRP) makes the use of public key certificates [30] for key distribution and management. Such certificates are already deployed as part of one-hop 802.11 networks [1]; this is the case on the UMass campus, where an 802.11 VPN is deployed and certificates are carried by nodes. For the process of public key certification, SZRP assumes the presence of trusted certification servers called the certification authorities (CAs) in the network in addition to the communicating nodes which we call

the common nodes (CNs). Each CN before taking part in communication need to be certified by some CA and are granted public keys. The detail of the certification process is described in Section 4.2.

SZRP is a two phase protocol. The first phase is the preliminary certification process where each CN fetches their required keys from their nearest CA. The second phase is secure routing phase which uses these keys to perform secure intra-zone or inter-zone routing by applying the process of digital signature and message encryption.

SK_X	Signature Key of node X (A private key used by X for signing)
VK_X	Signature verification key for node X. (A public key provided by X to verify its signature done with SK_X)
EK_X	Encryption Key for node X (A public key supplied by node X for encrypting any message to be sent to X)
DK_X	Decryption Key of node X (A private key used by X for decrypting any message which is encrypted with EK_X)
[d] SK_X	Packet ‘d’ signed with SK_X, this can be only verified using VK_X
{d}EK_X	Message ‘d’ encrypted with EK_X, this can be only decrypted with DK_X
[d] | b	b is appended to the packet containing d
$cert_X$	Public key certificate of X.
IP_X	IP address of X
t	Time stamp
e	Certificate expiration time
N_X	Nonce issued by node X
SKREQ	Session Key Request packet identifier
SKREP	Session Key Reply packet identifier
SRD	Secure Route Discovery packet identifier
SRR	Secure Route Reply packet identifier
ERR	Error packet identifier

Table 4.1: Notations Used in the Proposed Protocol

4.2 Certification Process

The Secure Zone Routing Protocol (SZRP) requires the presence of trusted certification servers called the certification authorities (CAs) in the network. The CAs are assumed to be safe, whose public keys are known to all valid CNs. Keys are generated apriori and exchanged through an existing, perhaps out of band, relationship between CA and each CN. Before entering the ad hoc network, each node requests a certificate from it's nearest CA. Each node receives exactly one certificate after securely authenticating their identity to the CA. The idea is depicted in Figure 4.1. The methods for secure authentication to the certificate server are numerous and hence it is left to the developers; a significant list is provided by [24, 30].

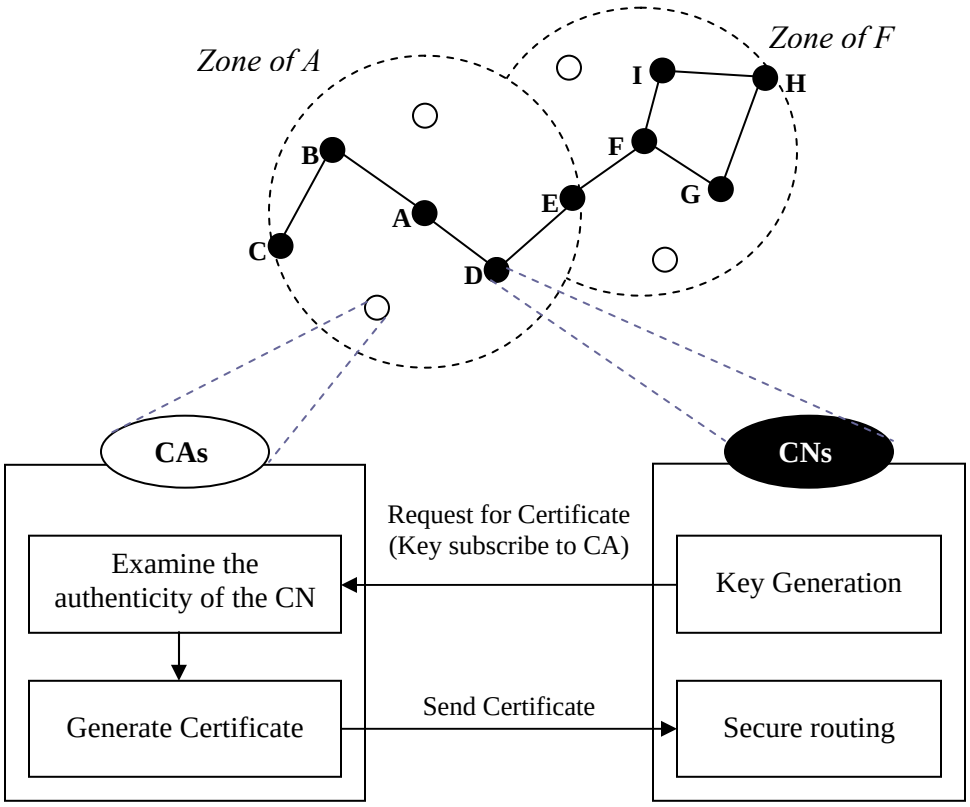


Fig 4.1: Certification Process in SZRP

A common node X receives a certificate from its nearest CA as follows:

$$CA \rightarrow X: \text{cert}_X = [IP_X, VK_X, EK_X, t, e] \mid sign_{CA}$$
$$\text{where, } sign_{CA} = [IP_X, VK_X, EK_X, t, e]\, SK_{CA}$$

The certificate contains the IP address of X, the two public keys VK_X and EK_X of X, one for verifying the signature signed by X and other for encrypting a packet to be send to X, a timestamp 't' of when the certificate was created, and a time 'e' at which the certificate expires, all appended by the signature $sign_{CA}$ of CA. All nodes must maintain fresh certificates with their nearest CA.

4.3 Design of Secure Zone Routing Protocol

This section describes in details the architectural design of the proposed protocol as a whole and its individual components in particular.

4.3.1 Design Assumptions

We have assumed the following things for the design and successful deployment of the proposed protocol.

- The network links are assumed to be bidirectional.
- The resources of different ad hoc network nodes may vary greatly, from nodes with very little computational resources, to resource rich nodes equivalent in functionality to high-performance workstations. To make our results as general as possible, we have designed SZRP to support nodes with moderate resources, such as a Palm Pilot or RIM pager.
- The proposed protocol intends to provide security at IP layer. Hence, for a secure communication across the network protocol stack suitable techniques should be employed to secure MAC and physical layers. A list of such mechanisms is given in [4, 31, 32].

4.3.2 Architecture

The architectural design of SZRP is shown in Figure: 4.2. The proposed architecture is a modification of ZRP [16]. It is designed to support both secure routing (intrazone and interzone) and effective key management. There are dedicated and independent components in SZRP to carry out these tasks. The functionality of each component and their interrelationship is explained below.

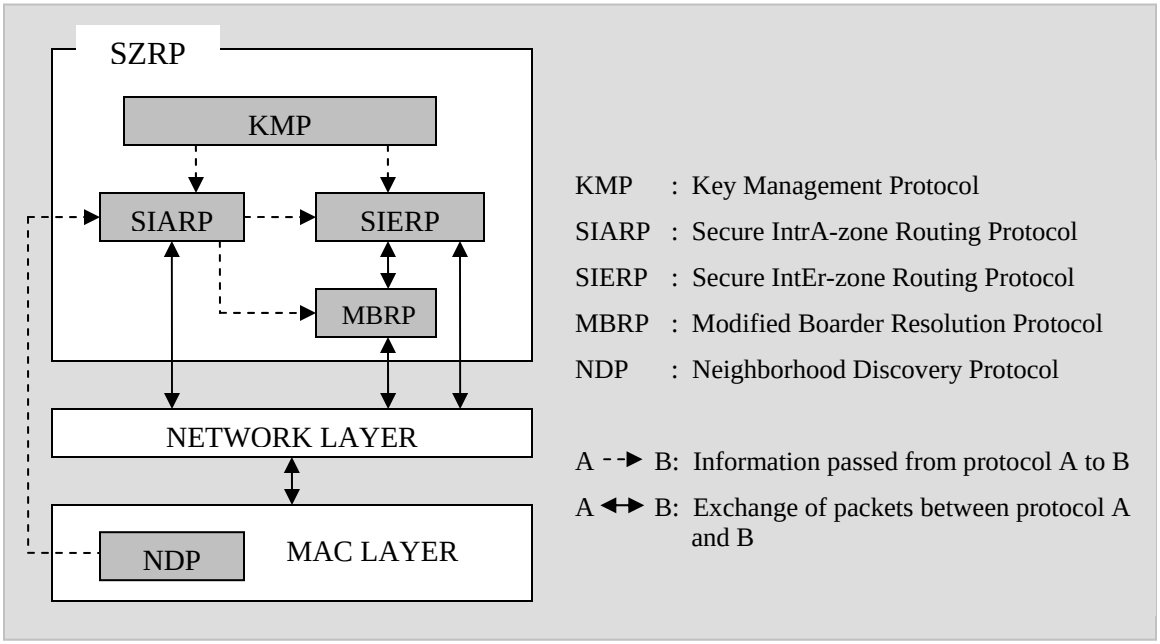


Fig 4.2: Architecture of SZRP

The key management protocol (KMP) is responsible for public key certification process discussed in section 4.2. It fetches the public keys for each CN by certifying them with the nearest CA. The secure intrazone routing protocol (SIARP) and secure interzone routing protocol (SIERP) uses these keys to perform secure intrazone and interzone routing respectively

SIARP is a limited depth proactive link-state routing protocol [5, 6] with inbuilt security features. It periodically computes the route to all intrazone nodes (nodes that are within

the routing zone of a node) and maintains this information in a data structure called SIARP routing table. This process is called *proactive route computation*. The route information to all intrazone nodes collected in proactive route computation phase is used by SIARP to perform secure intrazone routing. A detail discussion on secure intrazone routing and proactive route computation is given in Section 4.4.1 and Section 4.5 respectively.

SIERP is a family of reactive routing protocols [5] with added security features like ARAN [33]. It offers on demand secure route discovery and route maintenance services based on local connectivity information monitored by SIARP. The interzone routing and the route maintenance services offered by SIERP are discussed in Section 4.4.2 and Section 4.6 respectively

In order to detect the neighbor nodes and possible link failures, SZRP relies on the neighborhood discovery protocol (NDP) [18] similar to that of ZRP. NDP does this by periodically transmitting a HELLO beckon (a small packet) to the neighbors at each node and updating the neighbor table [18] on receiving similar HELLO beckons from the neighbors. NDP gives the information about the neighbors to SIARP and also notifies SIARP when the neighbor table updates. We have assumed that NDP is implemented as a MAC layer protocol. A number of security mechanisms suggested in [4, 32, 33] for MAC layer can be employed to secure NDP.

To minimize the delay during interzone route discovery, SIERP uses bordercasting technique [20] similar to ZRP, which is implemented here by the modified border resolution protocol (MBRP). MBRP is a modification of the bordercast technique [20] adopted in ZRP. It not only forwards SIERP's secure route discovery packets to the peripheral nodes of the bordercasting node but also sets up a reverse path back to the neighbour by recording its IP address. MBRP uses the routing table of SIARP to guide these route queries. Since, all security measures are taken by SIERP during interzone routing; no additional security mechanism is adopted by MBRP during bordercasting.

4.4 The Secure Routing Algorithm

This section describes the secure intrazone and interzone routing in details. We consider the network in Figure 4.3 for the illustration.

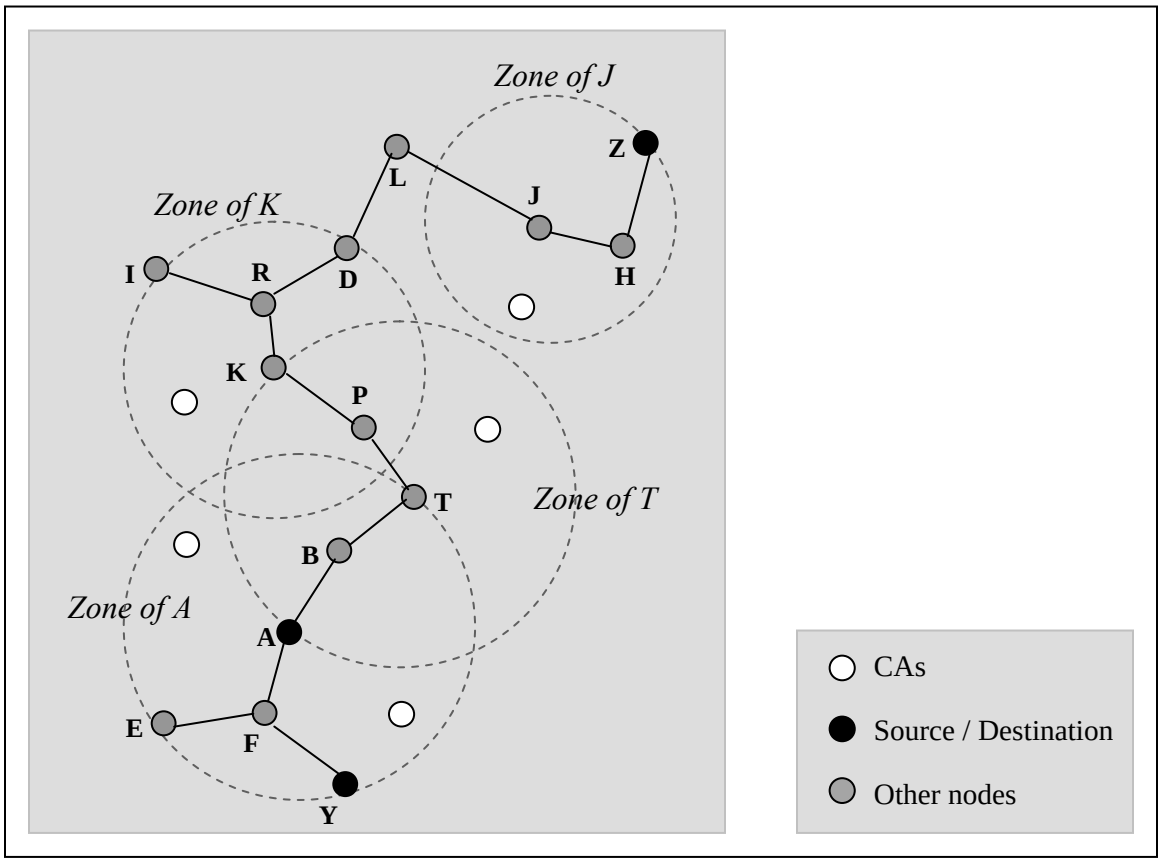


Fig 4.3: Intrazone and Interzone destinations of node A (zone radius $\beta = 2$)

SIARP, at each node, periodically computes the route to all intrazone nodes and maintains this information in SIARP routing table. For example in Figure 4.3, node *A* proactively computes the route to *B*, *T*, *E*, *F* and *Y* and stores this information in its SIARP routing table. This process, called *proactive route computation* is discussed in section 4.5.

When a node has a data packet for another node, it checks its SIARP routing table to determine whether the destination is within its zone or not. If the destination is within the

zone, for example if node *A* has a packet destined for node *Y*, the packet is forwarded to the destination proactively using SIARP. On the other hand if the destination is outside the zone, for example if node *A* wants to transmit a packet to *Z*, then interzone routing is performed using SIERP

4.4.1 Secure Intra-Zone routing

For intrazone routing we consider *A* as the source and *Y* as the destination. The following steps are taken by SIARP (at node *A*) to route a data packet from *A* to *Y.*

Step 1: *A* looks for the route to *Y* in its SIARP routing table and finds it to be A-F-Y.

Step 2: *A* sends a SKREQ packet to *Y* along this route requesting a session key K_{AY} between *A* and *Y*.

$$A \rightarrow Y : [SKREQ, IP_Y, cert_A] \mid sign_A$$
$$\text{where, } sign_A = [SKREQ, IP_Y, cert_A]\ SK_A$$

The SKREQ packet contains a packet type identifier "SKREQ", the IP address of the destination *Y*, and *A*'s certificate, all appended by the signature $sign_A$ of *A* signed using SK_A.

Step 3: *Y* on receiving this request, verifies the signature using VK_A, which it extracts from *A*'s certificate, creates the session key K_{AY}, encrypts it using EK_A and sends it to *A* as SKREP packet along the reverse route Y-F-A.

$$Y \rightarrow A : [SKREP, IP_A, cert_Y, \{K_{AY}\}EK_A] \mid sign_Y$$
$$\text{where, } sign_Y = [SKREP, IP_A, cert_Y, \{K_{AY}\}EK_A]\ SK_Y$$

The packet contains a packet type identifier "SKREP", the IP address of *A*, the certificate of *Y* and the session key encrypted using EK_A, all appended by the signature $sign_Y$ of *Y* signed using SK_Y.

Step 4: *A* on receiving the SKREP packet, verifies the packet using VK_Y, conforms its authenticity, decrypts it using DK_A and extracts the session key K_{AY}.

Once *A* gets the session key K_{AY}, it can encrypt the data packet using K_{AY} and send it to *Y* along the same route A-F-Y. All further communication between *A* and *Y* takes place similarly, using this session key.

4.4.2 Secure Inter-Zone routing

Secure interzone routing is done using SIERP. The interzone routing is initiated with an on demand *secure route discovery* phase in which the source finds the route to the desired interzone destination. The source then sends the data packet along this route. In our case when *A* wants to send a packet to *Z*, *A* looks in its SIARP routing table for a valid route to *Z*. Since *Z* is not within the zone of *A*, *A* fails to find the route. In this case, *A* begins the secure route discovery process to *Z*. The *secure route discovery process* gives *A* the authentic route to *Z* after which *A* forwards the data packet to *Z* along this route. In addition to secure route discovery, SIERP also performs route maintenance services based on the local connectivity information monitored by SIARP. Route maintenance is discussed in Section 4.6.

The following steps are taken by SIERP to route the data packet from *A* to *Z*:

Step 1: SIERP at *A* begins the *secure route discovery* process to *Z* by bordercasting to its peripheral nodes *T, E* and *Y,* a *SRD packet* with the help of MBRP.

$$A \rightarrow bordercast : [SRD, IP_Z, cert_A, \beta, N_A, t] \mid sign_A$$
$$\text{where, } sign_A = [SRD, IP_Z, cert_A, \beta, N_A, t]\, SK_A$$

The packet contains a packet type identifier “SRD”, the IP address of the destination *Z,* A’s certificate, the zone radius ‘β’, a nonce N_A created by *A* and the current time *t*, all appended by the signature $sign_A$ of *A*. The nonce N_A is monotonically increased every time *A* performs route discovery. N_A and *t*

together with the IP address of A (IP_A) uniquely identify the SRD which prevents the replay attack. N_A is made large enough such that, it will not need to be recycled within the probable clock skew between receivers. If a nonce later reappears in a valid packet that has a later timestamp, the nonce is assumed to have wrapped around, and is therefore accepted. Note that a hop count is not included with the message.

Step 2: When a peripheral node of A (T, E *or* Y), receives the SRD, it checks the (IP_A, N_A, t) tuple to verify that it has not already processed this SRD. Nodes do process packets for which they have already seen this tuple. The receiving node uses A's public key, which it extracts from A's certificate, to validate the signature and verify that A's certificate has not expired. If the packet is found to be authentic, it sets up a reverse path back to the source A by recording the neighbor from which it received the SRD, for example when the peripheral node T receives the SRD it sets up a reverse path back to A by recording the neighbor B from which it received the SRD (B sets up a reverse path to A during bordercasting. Now, T sets up the reverse path to B. So a reverse path from T to A is set).

The peripheral node then signs the contents of the message originally bordercast by A and appends this signature and its own certificate to the SRD. It checks in its SIARP routing table whether it has a valid path to the destination Z. If it has (Z is within the zone of the node), it forwards the SRD directly to Z along this route, otherwise it reborderscasts the packet to its peripheral nodes. In the present case since none of the peripheral nodes T, E and Y has the route to Z (Z is not within the zone of T, E or Y), all reborderscasts the SRD to their peripheral nodes, for example, T reborderscasts the SRD to K.

$$T \rightarrow bordercast : [[SRD, IP_Z, cert_A, \beta, N_A, t] \mid sign_A] \mid sign_T, cert_T$$

$$\text{where, } sign_T = [[SRD, IP_Z, cert_A, \beta, N_A, t] \mid sign_A]]SK_T$$

Step 3: Upon receiving the SRD, *T'*s peripheral node *K* checks the *(IP_A, N_A, t)* tuple, validates *T'*s signature and sets up the reverse path to *T* (if the signature is authentic). *K* then removes *T'*s certificate and signature, signs the contents of the message originally bordercast by *A* and appends this sign along with its own certificate to the SRD. It checks in its SIARP routing table whether it has a valid path to *Z.* Since it doesn't, it again rebordercasts the packet to its peripheral nodes *I* and *D.*

$$K \rightarrow bordercast : [[SRD, IP_Z, cert_A, \beta, N_A, t] \mid sign_A] \mid sign_K, cert_K$$
$$\text{where, } sign_K = [[SRD, IP_Z, cert_A, \beta, N_A, t] \mid sign_A]]SK_K$$

Each node along the path repeats these steps of validating the previous node's signature, recording the previous node's IP address for setting up the reverse path, removing the previous node's certificate and signature, signing the original contents of the message, appending its own certificate and rebordercasting the message, until the SRD reaches a node, that has a valid route to the destination *Z* (*Z* is within the zone of the node). In this case the node instead of rebordercasting the SRD, directly forwards it to *Z.* For example, when the SDR reaches *J*, it validates the packet, sets up the reverse path to the bordercasting node *D*, removes *D*'s certificate and signature, signs the contents of the message originally bordercast by *A,* appends this signature and its certificate and forwards the SRD to *Z.*

$$J \rightarrow Z : [[SRD, IP_Z, cert_A, \beta, N_A, t] \mid sign_A] \mid sign_J, cert_J$$
$$\text{where, } sign_H = [[SRD, IP_Z, cert_A, \beta, N_A, t] \mid sign_A]]SK_J$$

Step 4: Finally, the SRD arrives at destination *Z*, which replies to the first SRD that it receives for a source and a given nonce. There is no guarantee that the first SRD received traveled along the shortest path from the source. A SRD that travels along the shortest path may be prevented from reaching the destination first if it encounters congestion or network delay, either legitimately or maliciously

manifested. In this case, however, a non-congested, non-shortest path is likely to be preferred to a congested shortest path because of the reduction in delay. Because SRDs do not contain a hop count or specific recorded source route, and because messages are signed at each hop, malicious nodes have no opportunity to redirect traffic.

Z on getting this SRD packet verifies it using both VK_J and VK_A, confirms its authenticity and extracts EK_A. *Z* creates a *secure route reply* (SRR) packet and *unicasts* it back to the source along the reverse path. The first node that receives the SRR sent by *Z* is *H.*

$$Z \rightarrow H : [SRR, IP_A, cert_Z, N_A, t, \{K_{AZ}\}EK_A] \mid sign_Z$$
$$\text{where, } sign_Z = [SRR, IP_A, cert_Z, N_A, t, \{K_{AZ}\}EK_A]SK_Z$$

The SRR includes a packet type identifier "SRR", the IP address of *A*, the certificate of *Z*, the nonce N_A, the associated time stamp *t* sent by *A* and a session key K_{AZ} between *A* and *Z* encrypted with EK_A, all appended by the signature $sign_Z$ of *Z*. Nodes that receive the SRR forward the packet back to the predecessor from which they received the original SRD. Each node along the reverse path back to the source signs the SRR and appends its own certificate before forwarding the SRR to the next hop. Since, *J* is the next hop node to the source *A* after *H*:

$$H \rightarrow J : [[SRR, IP_A, cert_Z, N_A, t, \{K_{AZ}\}EK_A] \mid sign_Z] \mid sign_H, cert_H$$
$$\text{Where, } sign_H = [[SRR, IP_A, cert_Z, N_A, t, \{K_{AZ}\}EK_A] \mid sign_Z]\, SK_H$$

J on getting the SRR validates *H*'s signature on it, removes *H*'s signature and certificate, signs the contents of the message and appends this signature and its own certificate before unicasting the SRR to its neighbour *L*.

$$J \rightarrow L : [[SRR, IP_A, cert_Z, N_A, t, \{K_{AZ}\}EK_A] \mid sign_Z] \mid sign_J, cert_J$$
$$\text{Where, } sign_J = [[SRR, IP_A, cert_Z, N_A, t, \{K_{AZ}\}EK_A] \mid sign_Z]\, SK_J$$

Each node checks the nonce and signature of the previous hop as the SRR is returned to the source. This avoids attacks involving impersonation and replay of the message. Eventually the source *A* receives the SRR.

Step 5: On getting the SRR, *A* verifies *Z*'s signature and the nonce returned by *Z* to conform its authenticity. It then extracts the session key K_{AZ}. *A* now encrypt the data packet using K_{AZ} and send it to *Z* along the same route.

4.5 Proactive route computation

For *proactive route computation* each node within its routing zone periodically advertises a *link state packet (LSP)*. For example, node *A* advertises the LSP within the zone of *A*.

$$A \rightarrow brdcast : [LSP, IP_A, cert_A, \beta, TTL, SNo, neighbour[n], link\_metric[n]] \mid sign_A$$

Where, $sign_A = [LSP, IP_A, cert_A, \beta, TTL, SNo, neighbour[n], link\_metric[n]]\ SK_A$

The packet contains a packet type identifier "LSP", the IP address of the broadcasting node *A*, the certificate of *A*, the zone radius 'β', a time-to-live (TTL) value, the sequence number *SNo* of the packet which is used to track the link state history of the source node *A,* the list of neighbours of *A*, and link metrics, all appended by the signature $sign_A$ of *A*. The TTL field is used to control the scope of the packet which is initialized to β*-1* hops by *A*. Upon receipt the packet, the TTL value is decremented and as long as the value is greater than 0, the LSP is rebroadcasted.

When a neighbour of *A*, receives the LSP, it verifies the authenticity of the packet using VK_A which it extract from *A'*s certificate in the LSP, add LSP's information to its link-state table [6], decrement the value of TTL field and again forwards this LSP as long as the value of TTL field is greater than 0 else the LSP is dropped. Because every node within the zone of *A* receives the same LSPs, all the nodes build the same link state table. A typical link state table is shown in Figure 4.4.

Source Address	Zone radius	Neighbour ID	Insert time	Route metrics

Fig 4.4: Link state routing table maintained at each node

Once the link-state table is built, each node computes the route to every other node within its zone by applying the Dijkstra algorithm [6] to its link state table and stores this information in its SIARP routing table. A typical SIARP routing table at maintained at node *A* is shown in Figure 4.5.

Destination Address	Routes	Route metrics
Y	A-F-Y	.....
T	A-B-T	.....
E	A-F-E	.....
F	A-F	.....
B	A-B	.....

Fig 4.5: SIARP routing table maintained at node *A*

4.6 Route maintenance

The secure zone routing protocol (SZRP) is a hybrid routing protocol. SIARP is proactive and SIERP is reactive in nature. SIARP doesn't mandate for route maintenance, as the node mobility within a zone is periodically updated. However, route maintenance is required in SIERP for interzone routing.

For route maintenance, SIERP at each node keeps track of routes whether they are active or not. When there is no flow of traffic on an existing route for that route's lifetime, the route is deactivated by the node. Data received on an inactive route causes nodes to generate an Error (ERR) message. A node generates an ERR message in either of the following cases: (i) if data is received on an inactive route, or (ii) the link of an active route is broken due to node mobility or some other reasons. The node send the ERR message to the source along the reverse path. All ERR messages must be signed to check the authenticity of the sender as well as the message. For a route between source *A* and destination *X*, a node *M* generates the ERR message for its neighbor *N* as follows:

$$M \rightarrow N : [ERR, IP_A, IP_X, cert_M, N_M, t] \mid sign_M$$

Where, $sign_M = [ERR, IP_A, IP_X, cert_M, N_M, t]\ SK_M$

This message is forwarded along the path to the source without modification. A nonce and timestamp ensure that the ERR message is a fresh. Since the ERR messages are signed, malicious nodes cannot generate ERR messages for other nodes. The non-repudiation provided by the signed ERR message allows a node to be verified as the source of each ERR message that it sends. The source node drops the duplicate ERR message with same nonce and time stamp.

4.7 Analysis of Secure Zone Routing Protocol

In this section, we analyze the security aspects of SZRP by evaluating its robustness in the presence of attacks mentioned in Section 3.3. SZRP can prevent against all types of attacks that include information disclosure, impersonation, modification, fabrication and replay of packets caused by both an external advisory and an internal compromised node.

Prevention from Information Disclosure: No hop count information is present in the SRD or SRR packets. This prevents an external advisory or an internal compromised node from getting any kind of information about the network topology. Topology information is restricted to nodes within a zone. This is harmless as nodes accept packets

only after verifying the sender's signature. Further all the data packets and the control packets that contain the session key are encrypted which ensures the confidentiality.

Attacks involving impersonation: SZRP participants, accept only those packets that have been signed with a certified key issued by a CA. In intrazone routing since the SKREQs and SKREPs can only be signed by an authenticated source with its own private signature key, nodes can't impersonate (spoof) other nodes. Interzone routing follows hop-by-hop authentication during route discovery and end-to-end authentication during the route reply phase. So it is impossible for an external node or an internal compromised node to impersonate an intermediate node during interzone routing. Further since the SRD packet is signed by the source node using its private key, it guarantees that only the source can initiate a route discovery process. Similarly, the SRR packets include the destination's certificate and signature, ensuring that only the destination can respond to the route discovery. This prevents attacks where the source, the destination or any intermediate nodes are spoofed e.g. man-in-the-middle attack and sybil attack.

Routing message Modification: SZRP specifies that all fields of LSPs, SRD and SRR packets remain unchanged between the source and the destination. Since all packets are signed by the initiating node, any alterations in transit would be immediately detected by intermediate nodes along the path, and the altered packets would be subsequently discarded. Repeated instances of altering packets could cause other nodes to exclude the errant node from routing. Thus, modification attacks like redirection of routing messages and DoS attacks are prevented.

Fabrication of messages: Messages can be fabricated only by the internal compromised nodes with certificates. In that case, SZRP does not prevent fabrication of routing messages, but it does offer a deterrent by ensuring non-repudiation. A node that continues to inject false messages into the network may be excluded from future route computation.

Replay Attacks: Replay attacks like tunneling and wormhole attacks are prevented by including a nonce and a timestamp with routing messages.

4.8 Conclusion

In this chapter, we presented the design of a new secure ad hoc routing protocol called the Secure Zone Routing Protocol (SZRP) which is based upon the concept of hybrid routing. We analyzed the robust of the protocol against multiple attacks in the network and found that, the proposed protocol gives a better solution towards achieving the security goals like message integrity, data confidentiality and authentication, by taking an integrated approach of digital signature and both the symmetric and asymmetric key encryption technique. In the next chapter, we present the possible implementation and performance evaluation of the proposed protocol through simulation work.

Chapter 5

SIMULATION OF SECURE ZONE ROUTING PROTOCOL

Simulation setup

Performance metrics

Simulation results and analysis

Remarks

Chapter 5

Simulation of Secure Zone Routing Protocol

The performance of Secure Zone Routing Protocol (SZRP) was evaluated using Network Simulator-2 version 2.1b6a (NS-allinone-2.1b6a) [40]. NS-2 provides a framework for simulation of wired and wireless networks, including some facility for emulation. The NS-2 simulator is written in C++ with a Tcl shell in the front-end that uses oTcl (object-oriented Tcl) libraries. Scenarios are run by feeding an oTcl script to the NS-2 executable. The output can be read directly or post-processed by an interactive graphics viewer called NAM. Generally NS-2 has a different architecture for wireless and wired simulation. Current version of NS-2 does not support any sort of security architecture. So for that purpose special classes were designed.

This chapter describes how the proposed protocol has been simulated, the technology and hardware used for simulation, the network scenario and the analysis of simulation results.

5.1 Simulation Setup

The simulation of Secure Zone Routing Protocol (SZRP) was conducted in NS-allinone-2.1b6a, on an Intel Pentium IV processor (2.4 GHz) and 512 MB of RAM running Ubuntu 7.2. To make our results as general as possible, we have simulated SZRP to support nodes with moderate resources. The proposed protocol has been implemented over the ZRP protocol specification document for NS-2, contributed by Robin Poss in [41], with required modifications to support the adopted security mechanisms.

5.1.1 Network Scenario

In the studied scenario, we simulated two types of field configurations: 10 nodes distributed over a 700m x 700m terrain and 20 nodes over a 1200m x 1200m terrain. Node transmission range was taken to be 250m. The initial positions of the nodes were random. Node mobility was simulated according to the random waypoint mobility model [39], in which each node travels to a randomly selected location at a configured speed and then pauses for a configured pause time, before choosing another random location and repeating the same steps. We ran simulations for a constant node speeds of 0, 1, 5 and 10 m/s, with pause time fixed at 30 seconds.

The implementation used 802.11 MAC layer and CBR traffic over UDP. We simulated five CBR sessions in each run, with random source and destination pairs. Each session generated 500 data packets of 512 bytes each at the rate of 4 packets per second. All simulations were run for 150 seconds of simulated time. Figure 5.1 and Figure 5.2 shows the screenshots of the simulation scenario with 10 and 20 nodes respectively.

5.1.2 Security Model

SZRP was simulated using a 512 bit key and 16 byte RSA digital signature. These values are reasonable to prevent compromises during the routing process.

For the proposed protocol, we assumed a routing packet processing delay of 2ms. This value was obtained through field testing of the ZRP protocol implementation [41] under identical conditions as that of SZRP. Additionally, a digital signature generation delay of 6.5ms and verification delay of 0.5ms was simulated for SZRP. These values were obtained by measuring the multiple running times of the RSA digital signature and verification algorithm implemented in J2SE 1.4 on a desktop computer with a Pentium IV processor (2.4 GHz) and 512 MB of RAM running Ubuntu Linux 7.2. Additionally, a random delay between 0 and 10ms was introduced before a packet is transmitted in order to minimize collisions.

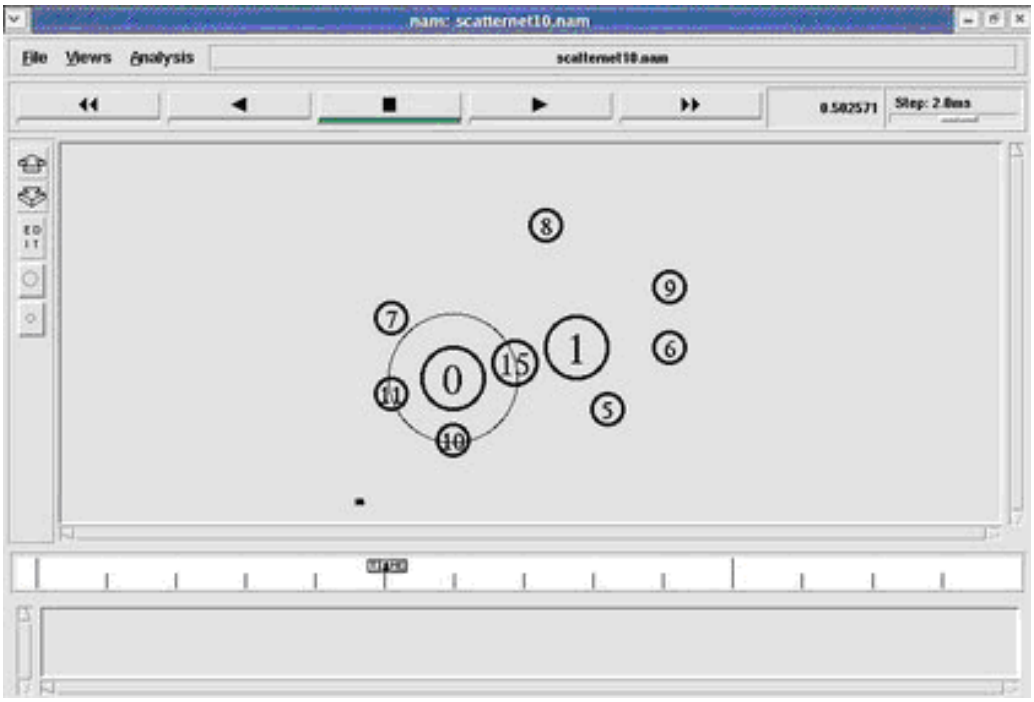


Fig 5.1: Transmission between 10 Nodes distributed over a 700m x 700m terrain

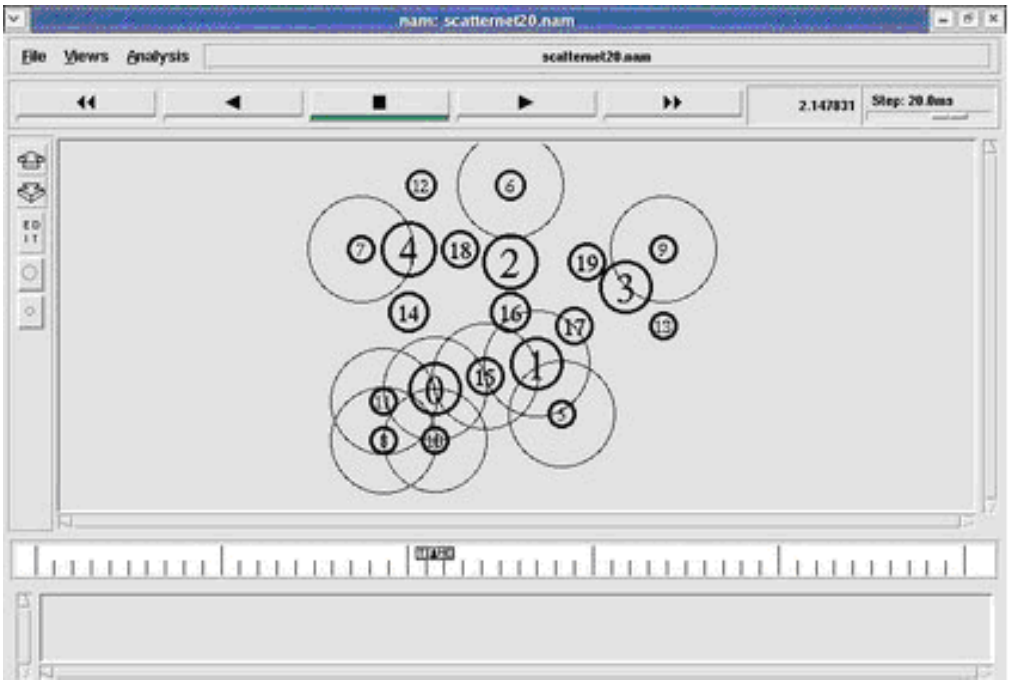


Fig 5.2: Transmission between 20 Nodes distributed over a 1200m x 1200m terrain

5.2 Performance Metrics

In order to evaluate the performance of Secure Zone Routing Protocol (SZRP), both ZRP and SZRP were run and compared under identical mobility patterns and traffic scenarios. A basic version of ZRP was used, which did not include optimizations. This enables a consistent comparison of results. The version of ZRP implemented here is contributed by Robin Poss in [41]. We used two classes of metrics to compare the performance of ZRP and SZRP. The first class of metrics evaluates both the protocols under a non-adversarial network setting, assuming all the nodes in the network to be well-behaved and benign. The second class of metrics was used to compare their performances under a hostile environment where malicious nodes are present in the network.

5.2.1 Metrics used for a Non-Adversarial Environment

We evaluated four performance metrics to compare the proposed protocol with ZRP under a trusted environment where all the nodes in the network are assumed to be benign. They are discussed below:

- **Average packet delivery fraction:** This is the fraction of the data packets generated by the CBR sources that are delivered to the destination. This metric is important as it evaluates the ability of the protocol to discover routes.

- **Average routing load in bytes:** This is the ratio of overhead control bytes to delivered data bytes. Secure Zone Routing Protocol (SZRP) has larger control overhead due to the certificate and signature embedded in the packets. For the calculation of this metric, the transmission at each hop along the route was counted as one transmission.

- **Average routing load in terms of packets:** This metric is similar to the above, but here the ratio of control packet overhead to data packet overhead is calculated.

- **Average route acquisition latency:** This is the average delay between the sending of a secure route discovery packet by a source for discovering a route to a destination and the receipt of the first corresponding route reply. This includes all the delays caused during the route discovery and route reply phases for signature verification and their replacement, in addition to the normal processing of the packets. If a route request timed out and needed to be retransmitted, the sending time of the first transmission was used for calculating the latency.

5.2.2 Metrics used for a Hostile Network Setting

The metrics described in the previous section compare the performance of SZRP and ZRP when all the nodes in the network are well-behaved. We conducted additional experiments to determine the effect of malicious node behavior on the two protocols. For this, we used the field configuration of 20 nodes distributed over a 1200m x 1200m area and ran simulations with 20% and 30% malicious nodes for each protocol. The malicious nodes were selected randomly. We measured the following metric:

- **Percentage of Packets Dropped that passed through Malicious Nodes:** This metric indicates the percentage of total packets dropped that traverse malicious nodes when using each routing protocol, in the presence of different percentages of malicious nodes. Assuming that all the packets that pass through a malicious or compromised node were altered, this metric can be calculated as follows:

$$\textit{\% of Packets Dropped that passed through Malicious Nodes} = \left(\frac{\textit{No. of packets dropped by the benign nodes that are previously generated by or passed through any malicious node in the network}}{\textit{Total number of packets communicated}} \right) \text{X } 100$$

  The metric evaluates the degree to which the communication is secure, as packets passing through malicious nodes may possibly disrupt secure communication.

5.3 Simulation Results and Analysis

In this section we present and analyze the observed results for each of the performance metric discussed in the previous section under the network and security setup given in Section 5.1. The resulting data were plotted using Gnuplot 4.2.5 [42]. Each data point in the resulting graphs is an average of 5 simulation runs with identical configuration but different randomly generated mobility patterns.

5.3.1 Average Packet Delivery Fraction

Figure 5.3 shows the observed results for average packet delivery fraction for both the 10 and 20 node networks. As shown in the figure, the packet delivery fraction obtained using SZRP is above 96% in all scenarios and almost identical to that obtained using ZRP. This suggests that SZRP is highly effective in discovering and maintaining routes for delivery of data packets, even with relatively high node mobility.

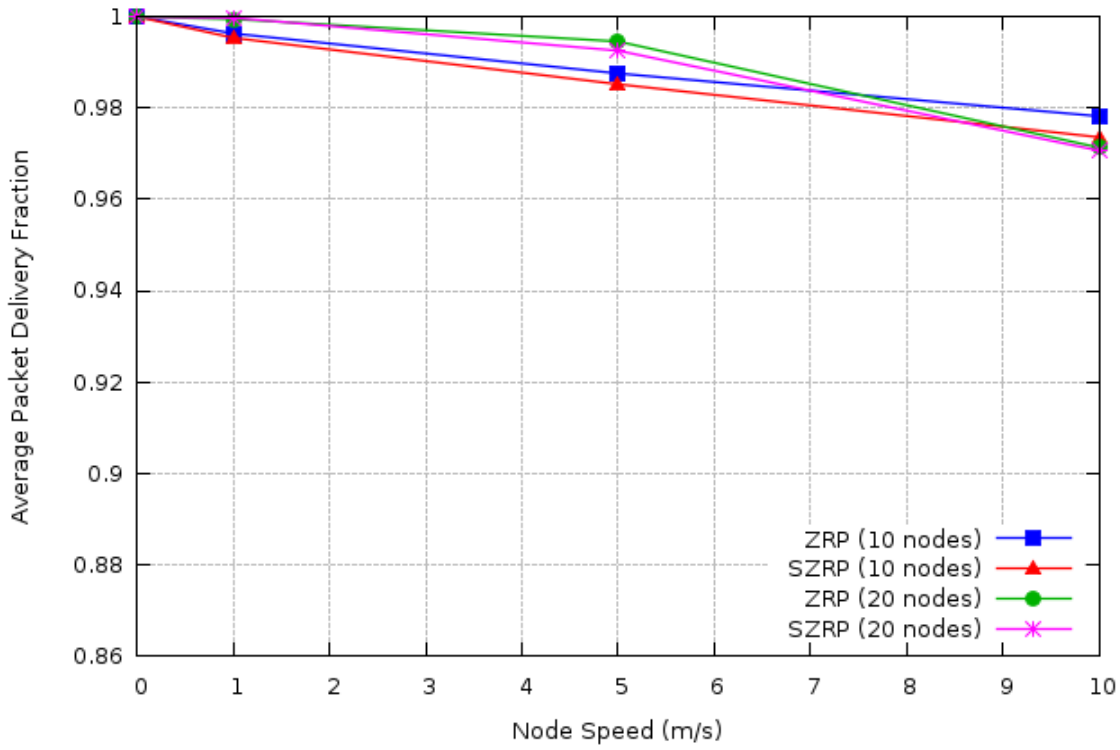


Fig 5.3: Simulation Results – Average Packet Delivery Fraction

5.3.2 Average Routing Load in Bytes

Figure 5.4 shows the routing load measurements for both the protocols in terms of number of control bytes per data bytes delivered. As shown in the figure, the byte routing load of Secure Zone Routing Protocol (SZRP) is higher compared to that of ZRP. For example, it is nearly 40% for 20 nodes moving at 5 m/s, as compared to 22% for ZRP with identical topology and mobility pattern. With further increase in node mobility to 10 m/s, it increases to 75%, compared 45% for ZRP.

This overhead is due to the certificate and signature embedded in the packets. The RSA digital signature is of 16 bytes and the certificate is 512 bytes long. Though these extra bytes are pure overhead they are necessary for security provisioning. Additionally, since ZRP has the advantage of smaller sized packets, the packet size of SZRP is not that much larger compared to other secure routing protocols even after inserting the security data.

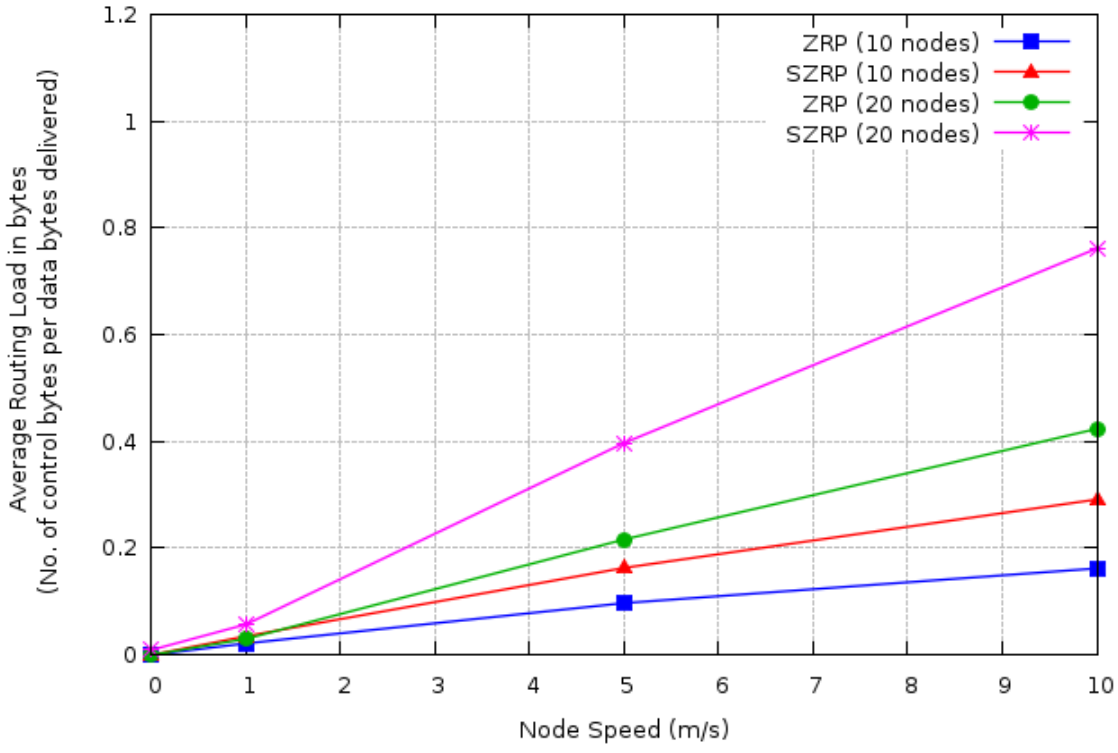


Fig 5.4: Simulation Results – Average Routing Load in bytes

5.3.3 Average Routing Load in Terms of Packets

While the number of control bytes transmitted by SZRP is larger than that of ZRP, the number of control packets transmitted by the two protocols is roughly equivalent. Figure 5.5 shows the average number of control packet transmitted per delivered data packet. Except for the scenario of 20 nodes moving at 1 m/s, where they exhibit some difference, the packet routing load for both the protocols are nearly the same for other scenarios.

This is due to the fact that SZRP did not employ any extra control packets compared to ZRP for secure routing, except for the case of intrazone routing, which requires two additional control packets SKREQ and SKREP. However, with high node mobility, for example, when the nodes move with the speed of 5 m/s or 10 m/s, the number of times interzone routing carried out was significantly higher than intrazone routing. In this respect, the two protocols demonstrate nearly the same amount of packet overhead.

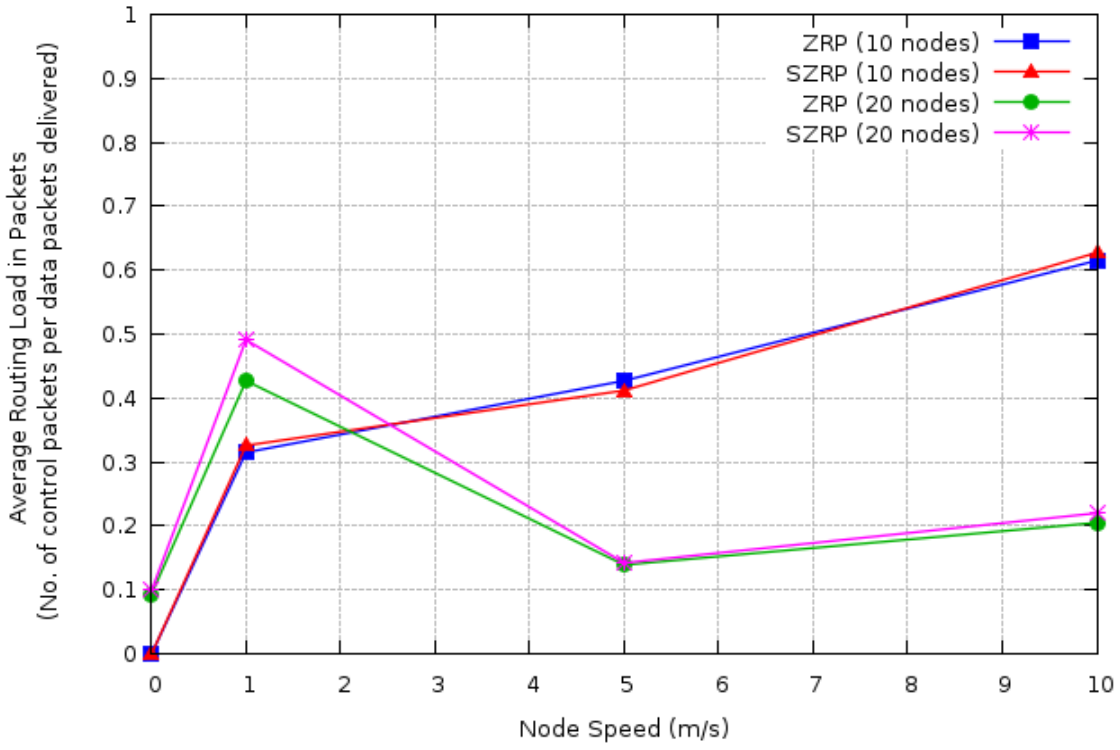


Fig 5.5: Simulation Results – Average Routing Load in Packets

5.3.4 Average Route Acquisition Latency

Figure 5.6 shows that the average route acquisition latency for Secure Zone Routing Protocol (SZRP) is approximately 1.7 times as that of ZRP. For example, for 10 nodes moving at 5 m/s, it is 60ms as compared to 100ms for ZRP, while for 20 nodes moving at 10 m/s, it is nearly 135ms as compared to 75ms as in the case of ZRP.

While processing SZRP routing control packets, each node has to verify the digital signature of the previous node, and then replace this with its own digital signature, in addition to the normal processing of the packet as done by ZRP. This signature generation and verification causes additional delays at each hop, and so the route acquisition latency increases.

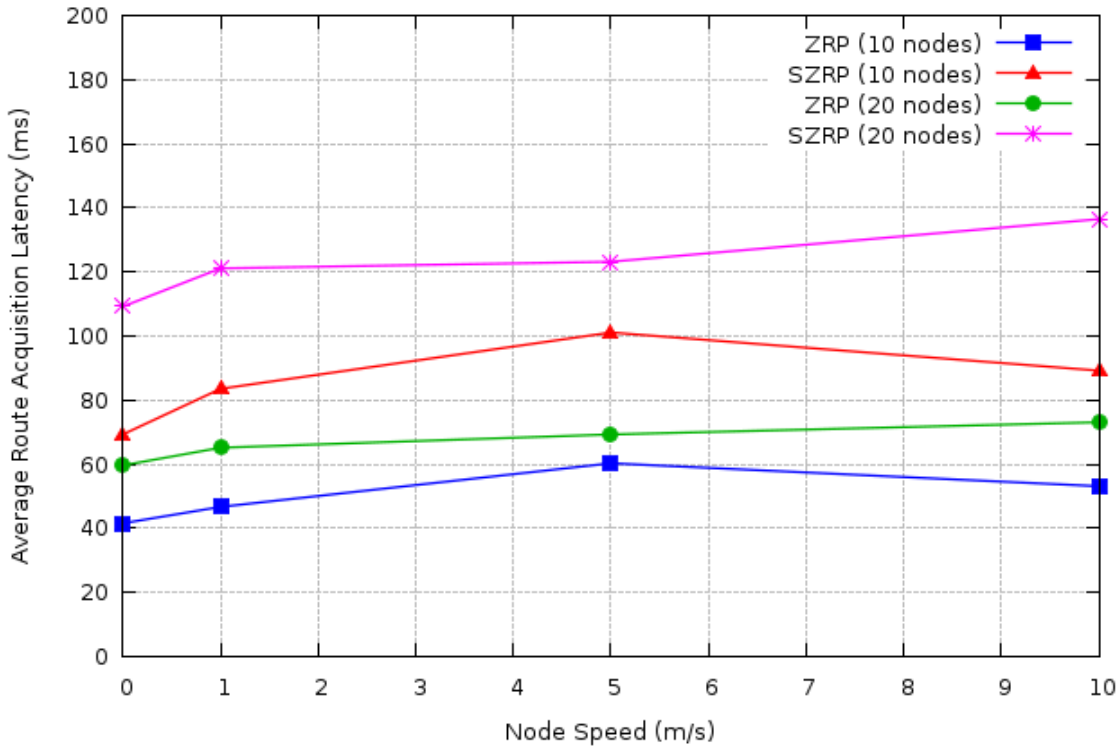


Fig 5.6: Simulation Results – Average Route Acquisition Latency

5.3.5 Percentage of Packets Dropped that Passed Through Malicious Nodes

Figure 5.7 illustrates the results of the experiments. As shown in the figure, when using SZRP, a much larger fraction of packets that passed through malicious nodes were dropped, as compared to that of ZRP. For instance, in the presence of 30% malicious nodes with no node mobility, only 26% of packets that pass through malicious nodes were dropped when using ZRP, as compared to almost 47% when using SZRP.

These results show that about 50% of packets that were possibly altered by malicious nodes in the network remained undetected and could potentially make their way through authentic nodes when using ZRP, as compared to the proposed protocol. This is a significant increase in the degree of security level.

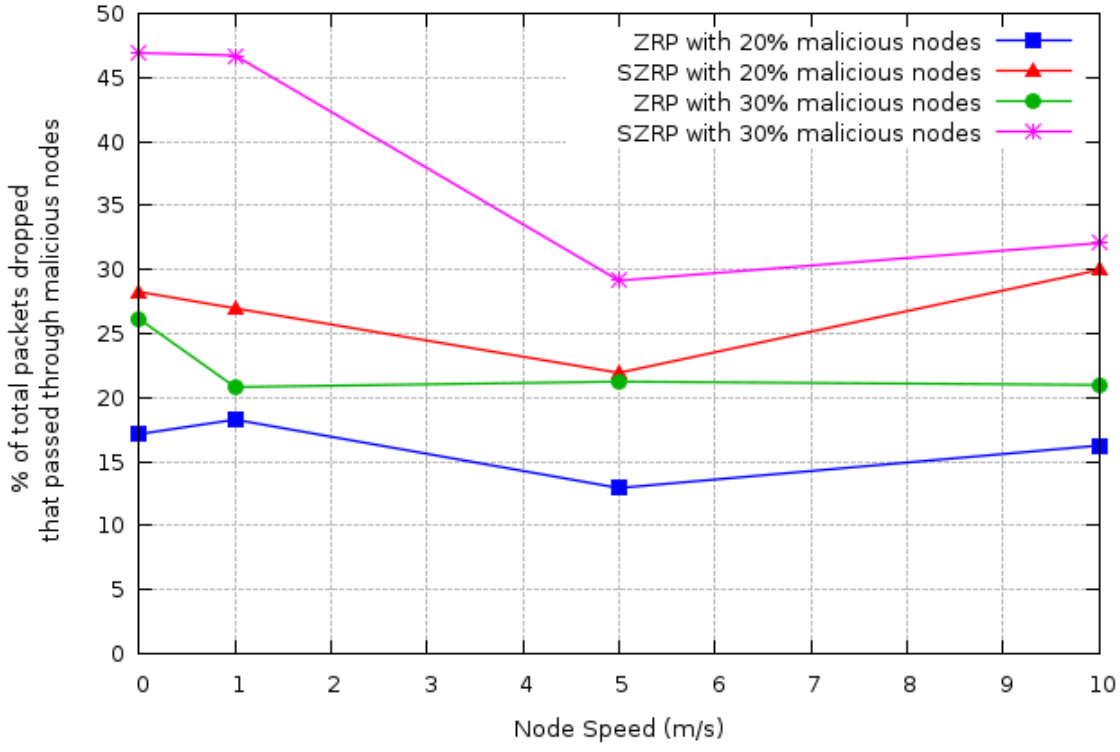


Fig 5.7: Simulation Results – Percentage of Packets Dropped that passes through Malicious Nodes

5.4 Conclusion

The simulation results for Secure Zone Routing Protocol under different mobility patterns and traffic scenarios show that the proposed protocol is as efficient as ZRP in discovering and maintaining routes, at the cost of using larger routing packets which result in a higher overall routing load, and at the cost of higher latency in route discovery because of the cryptographic computation that must occur. However, the impact of the overhead caused is almost insignificant and negligible as compared to the proposed degree of security, which SZRP provides compared to its other counterparts.

Chapter 6

CONCLUSION

Future Works

Chapter 6

Conclusion

In this thesis, we have considered the routing approaches in mobile ad hoc networks from the security viewpoint. We have analyzed the threats against ad hoc routing protocols and presented the requirements that need to be addressed for secure routing. Existing secure routing protocols for mobile ad hoc networks are either proactive or reactive in nature; hence are limited in their approach in terms of providing security across diverse networking applications. We explored the advantages of hybrid routing in dealing with these limitations, where the proactive and the reactive behavior is mixed in the amounts that best match these operational conditions.

We have presented the design and analysis of a new secure routing protocol for mobile ad hoc networks, called the Secure Zone Routing Protocol (SZRP). The proposed protocol is hybrid in nature and based on the concept of zone routing protocol (ZRP). It provides a solution for secure routing in an open and managed-open environment. In designing SZRP, we carefully fit the inexpensive cryptographic primitives to each part of the protocol functionality to create an efficient protocol that is robust against multiple attacks in the network. The proposed protocol gives a better solution towards achieving the security goals like message integrity, data confidentiality and message authentication, by taking an integrated approach of digital signature and both the symmetric and asymmetric key encryption techniques.

Secure Zone Routing Protocol (SZRP) is a simple protocol that supports nodes with moderate resources and does not require significant additional work from the nodes within the group. Our simulations show that SZRP is as efficient as ZRP in discovering

and maintaining routes, at the cost of using larger routing packets which result in a higher overall routing load, and at the cost of higher latency in route discovery because of the cryptographic computation that must occur. However, the impact of the overhead caused would be almost insignificant and negligible as compared to the proposed degree of security, which SZRP will provide to any network system, if adopted.

The proposed protocol intends to provide security at IP layer. Together with existing approaches for securing the physical layer and MAC layer within the network protocol stack, the Secure Zone Routing Protocol (SZRP) provides a foundation for governing a secure communication system for mobile ad hoc networks.

6.1 Future Works

The proposed protocol presented in this thesis considers that, the certification authorities (CAs) are safe within the network and are free from any kind of attacks caused either by external advisory or internal compromised nodes. A possible extension of the work may include employing additional feature to SZRP so that it can handle a scenario where the trusted certification authorities are compromised or attacked.

Additionally we have assumed that, the Neighborhood Discovery Protocol (NDP) is implemented as a MAC layer protocol. But in some special cases the MAC layer does not include an implementation of NDP. In such situations the proposed protocol may be modified to provide the functionality of NDP at IP layer.

BIBLIOGRAPHY

[1] C. Siva Ram Murthy and B. S Manoj, “Ad Hoc Wireless Networks, Architecture and Protocols”, Prentice Hall PTR, 2004.

[2] Stefano Basagni, Macro Conti, Silvia Giordano and Ivan Stojmenovic, “Mobile Ad Hoc Networks” , IEEE press, A john Wily & Sons, INC. publication, 2003

[3] George Aggelou, “Mobile Ad Hoc Networks”, 2nd edition, Mc GRAW Hill professional engineering, 2004

[4] Imrich Chlamtac, Marco Conti, Jenifer J.-N. Liu, “Mobile Ad Hoc Networking: Imperatives and Challanges”, Elsevier Network Magazine, vol. 13, pages 13-64, 2003

[5] E.M. Belding-Royer and C. K. Toh, “A review of current routing protocols for ad-hoc mobile wireless networks”, IEEE Personal Communications Magazine, pages 46–55, April 1999.

[6] Behrouz A. Forouzan, “Data communication and Networking,” 2nd edition, Tata McHill publication, 2001

[7] C. E. Perkins and P. Bhagwat, “Highly Dynamic Destination-Sequenced Distance-Vector Routing (DSDV) for Mobile Computers,” Comp. Commun. Rev., Oct. 1994, pp. 234–44.

[8] C.-C. Chiang, “Routing in Clustered Multihop, Mobile Wireless Networks with Fading Channel,” Proc. IEEE SICON ’97, Apr. 1997, pp. 197–211.

[9] S. Murthy and J. J. Garcia-Luna-Aceves, “An Efficient Routing Protocol for Wireless Networks,” ACM Mobile Networks and App. J., Special Issue on Routing in Mobile Communication Networks, Oct. 1996, pp. 183–97.

[10] P. Jacquet, P. Muhlethaler, A. Qayyum, “Optimized Link State Routing Protocol”, Internet Draft, draft-ietf-manetolsr- 00.txt, November 1998.

[11] A. Laouiti, A.Qayyum, and L.Viennot, “Multipoint Relaying; An Efficient Technique for Flooding in Mobile Wireless Networks,” in Proceeedings of the 35th Annual Hawaii International Conference on System Science (HICSS’ 2002), Waikoloa, HI, January 2002.

[12] D.B. Johnson, D.A. Maltz, “Dynamic source routing in adhoc wireless networks”, in: T. Imielinski, H. Korth (Eds.), Mobile Computing, Kluwer Academic Publishers, Dordrecht, 1996, pp. 153–181.

[13] C. E. Perkins and E. M. Royer, “Ad hoc on-demand distance vector routing”, In IEEE Workshop on Mobile Computing Systems and Applications, pages 90–100, Feb. 1999.

[14] V. D. Park and M. S. Corson, “A Highly Adaptive Distributed Routing Algorithm for Mobile Wireless Networks,” Proc. INFOCOM ’97, Apr. 1997.

[15] C-K. Toh, “A Novel Distributed Routing Protocol to Support Ad-Hoc Mobile Computing”, Proc. 1996 IEEE 15th Annual Int’l. Phoenix Conf. Comp. and Commun., Mar. 1996, pp. 480–86.

[16] Haas Z. J., Pearlman M. R., and Samar P., “The Zone Routing Protocol (ZRP)”, IETF Internet Draft, draft-ietf-manet-zone-zrp-04.txt, July 2002.

[17] Jan Schaumann, “Analysis of Zone Routing Protocol”, Course CS765, Stevens Institute of Technology Hoboken, New Jersey, USA, 8th December 2002

[18] Haas, Zygmunt J., Pearlman, Marc R., Samar, P.: “Intrazone Routing Protocol (IARP)”, IETF Internet Draft, draft-ietf-manet-iarp-01.txt, June 2001

[19] Haas, Zygmunt J., Pearlman, Marc R., Samar, P.: “Interzone Routing Protocol (IERP)”, IETF Internet Draft, draft-ietf-manet-ierp-01.txt, June 2001

[20] Haas, Zygmunt J., Pearlman, Marc R., Samar, P.: "The Bordercast Resolution Protocol (BRP) for Ad Hoc Networks", IETF Internet Draft, draft-ietf-manet-brp-01.txt, June 2001

[21] M.Joa-Ng and I. T. Lu, "A Peer-to-Peer Zone-Based Two-Level Link State Routing for Mobile Ad Hoc Networks," IEEE journal on Selected areas in Communications, vol. 17, no. 8, pp. 1415- 1425, August 1999

[22] L. Zhou and Z. J Haas, "Securing Ad Hoc networks," IEEE Network Magazine, vol. 13, no. 6, December 1999

[23] M. O. Pervaiz, M. Cardei, and J. Wu, "Routing Security in Ad Hoc Wireless Networks," Network Security, S. Huang, D. MacCallum, and D. -Z. Du (eds.), Springer, 2008.

[24] Behrouz A. Forouzan, "Cryptography and Network Security", Special Indian Edition, Tata McHill publication, 2007

[25] J. Newsome, E. Shi, D. Song, and A. Perrig, The Sybil Attack in Sensor Networks: Analysis & Defenses, Proc. of the 3rd Intl. Symp. on Information Processing in Sensor Networks, 2004

[26] B. Awerbuch, D. Holmer, C. Nita-Rotaru, and H. Rubens, "An On-Demand Secure routing protocol Resilient to Byzantine failures", Proceedings of ACM workshop on wireless security 2003, September 2003

[27] Y. Hu, A. Perrig and D. B Johnson, "Rushing attacks and defense in Wireless Ad HocNetwork Routing Protocol", Proceedings of ACM workshop on wireless security 2003, September 2003

[28] Jaydip Sen, M. Girish Chandra, Harihara S.G., Harish Reddy and P. Balamuralidhar, "A Mechanism for Detection of Gray Hole Attack in Mobile Ad Hoc Networks", Proceedings of 6th International Conference on Information, Communications and Signal Processing 2007, December 2007

[29] Y. Hu, A. Perrig and D. B Johnson, “Packet Leashes: A defense against Wormhole attacks in Wireless Ad Hoc networks”, Proceedings IEEE INFOCOM 2003, vol. 3, April 2003

[30] J. J. Tardo and K. Algappan, “SPX: Global authentication using public key certificates”, In Proceedings of the 1991 IEEE Symposium on Security and Privacy, pages 232–244, Oakland, CA USA, May 1991. IEEE Computer Society Press.

[31] R. Hauser, T. Przygienda, and G. Tsudik, “Lowering security overhead in link state routing”, Computer Networks, 31(8):885–894, April 1999.

[32] P. Michiardi, R. Molva, “Ad hoc networks security”, in: S. Basagni, M. Conti, S. Giordano, I. Stojmenovic (Eds.), Ad Hoc Networking, IEEE Press Wiley, New York, 2003.

[33] K.Sanzgir, and B.Dahill, “A secure routing protocol for ad hoc networks”, Proceeding of the 10th IEEE International Conference on Network Protocols, 2002, pp.1-10.

[34] Y. -C. Hu, D. B. Johnson and A. Perrig, “SEAD: Secure Efficient Distance Vector Routing for Mobile Wireless Ad Hoc Networks”, Fourth IEEE Workshop on Mobile Computing Systems and Applications (WMCSA’02), Jun. 2002.

[35] Y. -C. Hu, D. B. Johnson, and A. Perrig, Ariadne: A Secure On-Demand Routing Protocol for Ad Hoc Networks, Mobicom’02, 2002.

[36] R. Kravets, S. Yi, and P. Naldurg, “A Security-Aware Routing Protocol for Wireless Ad Hoc Networks”, In ACM Symp. on Mobile Ad Hoc Networking and Computing, 2001.

[37] P. Papadimitratos and Z. J. Haas, “Secure Routing for Mobile Ad hoc Networks”, In Proc. of the SCS Communication Networks and Distributed Systems Modeling and Simulation Conference (CNDS 2002), Jan. 2002.

[38] S. Buchegger and J. L. Boudec, “Performance Analysis of the CONFIDANT Protocol Cooperation Of Nodes Fairness In Dynamic Ad-hoc NeTworks”, In Proc. of IEEE/ACM Symposium on Mobile Ad Hoc Networking and Computing (MobiHOC), Jun. 2002.

[39] J. Broch, D. A. Maltz, D. B. Johnson, Y-C. Hu and J. Jetcheva, “A performance comparison of multi-hop wireless ad hoc network routing protocols”, In Proc. ACM MOBICOM, pages 85–97, Oct. 1998.

[40] http://www.isi.edu/nsam/ns

[41] http://www.geocities.com/robin_poss/zrp.html

[42] http://www.gnuplot.info