A systematic security analysis of the AODV protocol

by

Benjamin Robert Jones

A thesis submitted to the graduate faculty

in partial fulfillment of the requirements for the degree of

MASTER OF SCIENCE

Co-majors: Information Assurance; Computer Engineering

Program of Study Committee:
James Davis, Co-major Professor
Doug Jacobson, Co-major Professor
Barbara Licklider

Iowa State University

Ames, Iowa

2003

Graduate College
Iowa State University

This is to certify that the master's thesis of

Benjamin Robert Jones

has met the thesis requirements of Iowa State University

Signatures have been redacted for privacy

Abstract

Wireless networks are becoming a pervasive element of today's computing. With this growth a new type of network is being researched, the ad-hoc network. Wireless ad-hoc networks bring the advantage of flexibility because of the dynamic way that they are formed and left. Looking at these networks, with scenarios such as a tactical battlefield in mind, gives rise to many security issues. Because of the relative newness of wireless ad-hoc networks there is no standard routing protocol for the network layer. Many protocols have been proposed, and this thesis will examine one of the more popular protocols, Ad-hoc On-Demand Distance Vector Routing. With the growing need for security in mind, an analysis of this protocol will be done with relation to the security threats that are prevalent on today's tactical wireless networks. This will then determine the usability of the protocol in such environments were security is a key factor.

Table of Contents

List of Figures

List of Tables

Acknowledgments

The author would like to thank his major professors and all of his committee members for their support and guidance while writing this thesis. The author would also like to thank his entire family, without them this would not have been possible.

Chapter 1: Problem Statement

Ad-hoc Wireless Networks

Wireless networking is quickly becoming pervasive in today's society and commercial market. Because of the potential this creates for industry, the research being done on wireless networks is an increasingly interesting field. The commercial applications currently sold range from voice over wireless data networks to mobile IEEE 802.11 [11] networks being installed in buildings as the network infrastructure. An interesting branch of wireless networking is ad-hoc networking, where the network topology is dynamic based on the location of communicating partners. This creates many advantages for disaster relief situations, for the military, and for commercial applications by providing a network that is easily setup and harder to physically disrupt than wired networks. However, along with the great advantages in flexibility of ad-hoc networks, they by and large lack security (i.e. privacy and confidentiality). Because of the ease in which these networks are formed and the trust relationship inherent in ad-hoc networks, they are easily compromised. This creates a problem when confidential information or even life depending information can be compromised. The focus of this thesis is to examine the background of ad-hoc network routing protocols, and expose security flaws and trust relationships within one particular ad-hoc network routing protocol which is Ad-Hoc on Demand Distance Vector Routing or AODV.

Thesis Goals

The main objective of this thesis is the systematic study of the popular AODV wireless routing protocol in order to evaluate its ability to support secure communications.

In support of that goal this thesis is divided into four parts. Chapter 2 examines current ad-hoc network routing protocols with regard to the underlying algorithms for route discovery and message delivery. The third chapter entails a detailed demonstration of the route discovery of one of the more popular routing protocols, the AODV protocol. From there the security vulnerabilities of the wireless networking environment are described. Finally, the security vulnerabilities of the wireless environment are mapped to the AODV protocol for analysis of AODV's security properties.

Chapter 2: Background on Wireless Network Routing Protocols

2.1 Wireless Ad-hoc Networks

Wireless networks are becoming increasingly prevalent in the world today. With the increased availability and reduced prices, it is now economical for end users to set up and maintain their own wireless networks. The increase in use of wireless networks has not just been limited to the private sector. Many useful applications of wireless are becoming a reality in medicine, automotives, and the military. The main focus of this thesis is on ad-hoc wireless networks. An ad-hoc wireless network is two or more devices with wireless networking capability that create a peer-to-peer network [1]. An example of a simple ad-hoc wireless network is shown in Figure 1. Most ad-hoc networks are comprised of many nodes or mobile hosts (MH) each with their own transmission range.

As is depicted by Figure 1, the circles indicate the transmission ranges that each contain the mobile hosts. The node in the middle (mobile host B, or MH_b) is necessary for communications to take place across the network because it is encompassed in both spheres of communication. If the three nodes on the right-hand side of the network move away such that mobile host B is no longer in the communications circle of mobile hosts A and C, then the ad-hoc network is broken and communications can no longer take place over the entire network. The network these devices form differs from other wireless networks because there is no set infrastructure. Each device is capable of routing, packet forwarding, and connecting to any other device in its range. This flexibility allows ad-hoc wireless networks to operate in multiple situations where a regular wired network would fail. Wireless ad-hoc networks also allow for groups to exchange information quickly in cases such as board meeting presentations. The advantages wireless ad-hoc networks create suggests that further research on the properties and capabilities be done in this field.

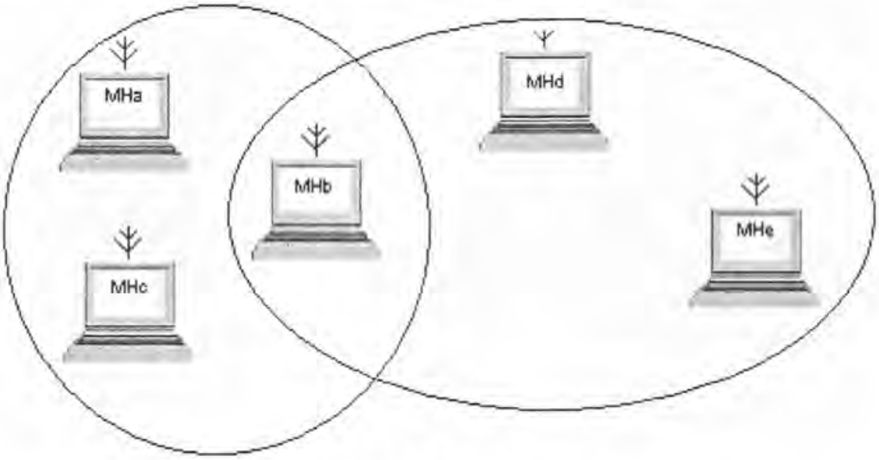

Figure 1 – An Ad-hoc Network

2.2 Introduction to Routing Protocols

Routing protocols play a crucial role in the operation of wireless networks, because the wireless routing protocols are by and large based on reliability and performance assumptions for a wired infrastructure. Depending upon the protocol used, it may create many advantages or disadvantages in the wireless network, which will be discussed in Section 2.5. Many of the protocols developed for wireless networks came from adaptations in wired network routing protocols [1].

These adaptations in wireless routing protocols are categorized into two approaches [1]. The first approach is table driven, where routing information from every node is broadcast to all other nodes. Each node locally maintains routing table information about the entire network. The two common table driven protocols discussed in this thesis are: The Destination Sequenced Distance Vector (DSDV), and The Wireless Routing Protocol (WRP).

The second approach to routing is an on-demand method, where a route is created only when a sending node requires it. The on-demand protocols discussed in this thesis are: The Ad Hoc On-Demand Distance Vector (AODV), The Temporally Ordered Routing Algorithm (TORA), and The Dynamic Source Routing (DSR). This analysis will describe each algorithm, demonstrate the operation of two of the algorithms, and then discuss the advantages and disadvantages of the routing protocols. Advantages and disadvantages are discussed to clarify the choice of a particular protocol to route messages because of the significant impact on the robustness and security of the entire communication system the routing protocol creates.

2.3 Ad-hoc Routing Protocol Algorithms

DSDV

The DSDV routing algorithm is essentially the Bellman-Ford algorithm [13] adapted for the dynamics of ad-hoc wireless networks [2]. In DSDV, each node maintains a routing table containing all possible destinations, the number of routing hops, and the next hop node for each destination. Routing updates of two types are sent throughout the network periodically by all nodes.

The first of these types is a full dump routing update. This update carries all routing information each node has available. The second update is incremental, where only information on routes that have changed is exchanged. The updates are controlled by sequence numbering which allow the receiving nodes to determine by comparison of the sequence numbers whether the update is current. Because the route table is always current, requests for routes are satisfied locally with a simple table lookup and no network communication (i.e. no additional network traffic) is needed to determine a path.

WRP

WRP is another table driven routing algorithm. WRP uses a novel approach to avoid the *count-to-infinity problem* (wherein a group of nodes exchange routing information that continually increases the distance by one, until the distance for a certain path "reaches" infinity) by forcing each node to check the consistency of its predecessor [1]. Here the routing table is created by adding entries to the table whenever a node identifies traffic such as acknowledgements and messages from its neighbors. If a node identifies a neighbor, the neighbor is added to the table and the node's tables are broadcast to the new neighbor. All nodes must broadcast at least a 'Hello' message, periodically informing the network that the node exists.

The four tables maintained by each node are: the distance table, the routing table, the link cost table, and the message retransmission table [6}. The distance table maintains the number of hops between the node and the destination. The routing table keeps a list of the next hop nodes. The link cost table is a listing of the delay for each link. The message retransmission list is used to keep track of update messages as they are sent after a node senses a link status change.

AODV

In Ad-Hoc On-Demand Distance Vector Routing, routes are determined at the time of need by requests from the source node [3}. In AODV, when a node sends information it initiates a route request (RREQ) broadcast. Each of these RREQs have a unique sequence number to prevent looping. While forwarding a RREQ, intermediate nodes create entries for their neighbors in a temporary routing table. Any duplicate RREQs as per the unique sequence number are dropped to avoid broadcast storming. Once the destination is reached, a route reply (RREP) is sent back. Each node on the forward path has set up timers as the RREQ was passed on, to facilitate route reply and route expiration. When the source receives confirmation of the route through the RREP, it begins to send information along that path. If the source node moves, it must reinitiate a RREQ to continue transmitting. If a node in the path moves, the upstream node must propagate a link failure. AODV also allows for periodic 'Hello' messages to be broadcast so its neighbors know which nodes are within its communications circle. In AODV the route is chosen according to the newest sequence numbers and the symmetrical status of the links.

TORA

Temporally Ordered Routing Algorithm (TORA) is a Global Position System (GPS) based algorithm that uses position to help determine an optimal route [4]. TORA is also a source initiated routing algorithm, therefore a transmitting node creates a directed acyclic graph (DAG) when initiating a transmission. The DAG is a graph that is rooted at the destination node and is created similarly to the query/reply process of Light Weight Mobile Routing[1]. Routes are created using a height metric that assigns each node in the DAG a downstream or upstream position. If a node moves, a new reference level in the DAG is defined. If a node wishes to be erased from the DAG, it broadcasts a CLR.

DSR

Dynamic Source Routing is the final protocol to be discussed. It is a source initiated on-demand protocol. DSR [5] works as follows: If a source node has a packet to send, it checks its route cache. Ultimately, messages are transmitted along an unexpired route. If no route is available, it broadcasts an RREQ similar to AODV. Each receiving node checks its cache for a route and when failing to locate a route adds its own address to the route record on the RREQ and rebroadcasts. When the destination node or a node containing a path is reached, that node initiates an RREP. This can be an asymmetric protocol because the destination node can initiate its own path discovery back to the source node, and use it upon confirmation of a better return route. The other feature of DSR is route maintenance. To perform route maintenance, a route error packet is utilized between nodes on a broken link to repair the route.

2.4 Routing Algorithm Examples

DSDV

As mobile hosts come online, each will eventually receive a copy of the current routing tables from their neighbors. Using that information and whatever metric is desired for best route choice, the newest mobile host updates its own tables from the received tables. Therefore, after routing tables have been exchanged between MH_a, MH_b. MH_c, MH_d, and MH_e (Figure 2)

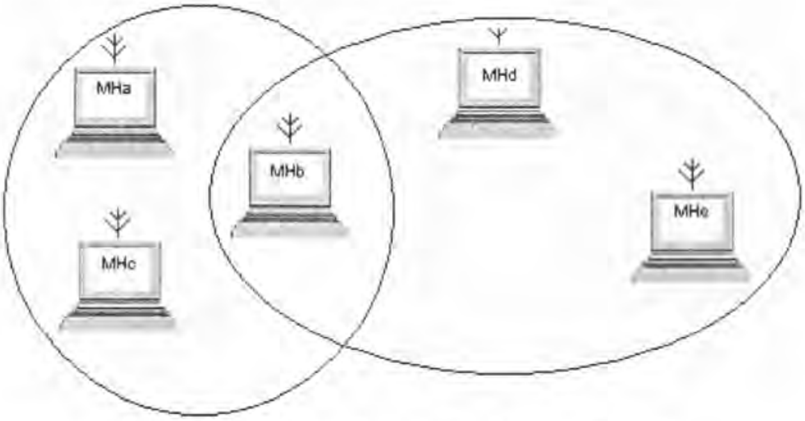

Figure 2 - An Example Wireless Ad-hoc Network

through full routing table dumps, any movement by any of the hosts will require only incremental changes to be sent out. Table 1 is an example of a partial DSDV routing table that corresponds to the mobile hosts in Figure 2. The metric for route selection used in DSDV depends upon the implementation of DSDV. Table 1 is the table that would be broadcast by MH_a on a full dump. As is shown each entry corresponds to one of the nodes in the ad-hoc network. Therefore, whenever MH_a whishes to transmit to another mobile host a lookup of the destination node reveals where the next hop is for that destination. This lookup also reveals a metric that is utilized when comparing routing table updates that have been received from the network. Looking at Figure 2 and Table 1, if mobile host A wishes to transmit a packet to mobile host E, all that is required of MH_a is a simple table query to see that the next hop for MH_e is MH_b. MH_a can then transmit its data to MH_b to be routed to MH_e.

Table 1 – Sample DSDV Routing Table

Destination	NextHop	Metric
A	A	0
B	B	1
C	C	1.5
D	B	2
E	B	3

AODV

The steps are significantly different for AODV route discovery than DSDV. If mobile host A wishes to transmit data to mobile host E (Figure 2) using the AODV routing protocol, first MH_a broadcasts a route request. MH_b and MH_c both receive the route request and check their cache to see if a route has recently been opened to the destination node E. If not then

MH_b and MH_c rebroadcast the route request. MH_a, MH_b, and MH_c will all receive the rebroadcasted route requests, but will discard the RREQs based on a comparison of sequence numbers that are part of the routing tables. MH_d will see the route request and again check the cache for a route to MH_e. When no route is again found, it will broadcast the route request again. Any previous hosts receiving this request will ignore it again based on the sequence number. MH_e finally receives the route request then sends a route reply back to MH_a. This route reply is forwarded back through the route to MH_a using the forward path set up from the route request as long as the timers on the forward path have not expired. MH_a finally receives the route reply and the path for data is established between MH_a and MH_e.

2.5 Advantages and Disadvantages of the Routing Protocols

Table Driven Routing Protocols

The foremost advantage a table driven algorithm has is that all of its routes are precomputed. This is an advantage in a tactical network because nodes do not need to exchange potentially large routing tables before each message. Another advantage to these routing schemes is that important equipment such as weapons systems can decline to forward traffic; therefore traffic would not be routed through critical systems.

A disadvantage of the table driven approaches of DSDV and WRP is that both table driven protocols require frequent communication between nodes to keep the contents of the routing tables current. In a mobile environment where most nodes are operating with limited battery life, energy requirements can cause problems. These protocols consume a significant amount of power simply passing updates of entire routing tables. If a mobile host is in sleep mode, its entire routing table will be expired when it awakens. Another major disadvantage is that when the movement of the nodes is great, most of the bandwidth will be taken up by route updates, which are tables, as opposed to actual data.

On-Demand Routing Protocols

The foremost advantage of on-demand routing is that all the routes are computed only when needed. Therefore, the overhead is very small when the nodes are not exchanging large amounts of data or moving during transmissions such that new routes must be created. Also, if transmissions between nodes increase, previous routes are cached for a period of time to help ease route discovery time. Another advantage of on-demand routing occurs in DSR where the links can be asymmetric. This is ideal for nodes that need to save power, because of the potential for a better one-way route back to the source node that DSR would be able to utilize.

One point that is both an advantage and a disadvantage in a tactical communications network is that TORA utilizes GPS. This could be a benefit to keeping track of nodes, but also a hazard by disclosing their location. As with table driven approaches, critical systems can also be protected from having traffic routed through them in on-demand routing.

Chapter 3: The AODV Protocol in Detail

To more clearly understand the security issues of AODV, it is helpful to study an example route discovery. This section of the thesis is devoted to looking at the AODV protocol in detail. To accomplish the task, this section will examine the packets used in route creation, the routing tables that are created during a route discovery process, and finally demonstrate the process using the data structures defined to explain the route discovery process in detail.

AODV uses two types of messages to find and establish a route employing the two packets shown in Figures 3 and 4. These are the packets being generated between each node as the route discovery is taking place. The packet formats for the Route Request (RREQ) and the Route Reply (RREP) are additionally wrapped in an IP packet for exchange between the nodes[3].

For this example it is assumed when a message is broadcast the IP layer will use the broadcasting node's IP as the source in the IP packet, and 255.255.255.0 as the broadcast destination in the IP layer. This packet is then wrapped around the RREQ and RREP packets. In addition to these two packet types, soft route tables (because the tables are associated with a timer so that they are able to expire) are generated at each node (Figure 5) for route retention and quicker recovery from link breakage. This section will further detail the packet types, routing tables, and then show a detailed example of what is occurring when a route is being found.

3.1 The Route Request Packet (RREQ)

In Figure 3, the fields are defined as follows by [3]:

Type - is equal to 1 for an RREQ
Flags - are used in multicasting operations as well as control functions
Rsv - Sent as 0 and ignored on reception
Hop Count - The number of hops from the Originator IP Address to the node handling the request
RREQ ID - A sequence number uniquely identifying the particular RREQ when taken in conjunction with the originating node's IP address
Destination IP Address - The IP address of the destination for which a route is desired
Destination Sequence Number - The greatest sequence number received in the past by the originator for any route towards the destination
Originator IP Address - The IP address of the node which originated the Route Request
Originator Sequence Number - The current sequence number to be used for route entries pointing to (and generated by) the originator of the route request.

For the purpose of this example demonstration it is possible to disregard all but the important information in the RREQ packet while doing the demonstration. Therefore a sample packet will look like this: <RREQ, Hop Count, RREQ ID, Dest. IP, Dest. Seq. #, Orig. IP, Orig. Seq. #>.

<table>
<tr><td>Type</td><td>Flags</td><td>Rsv</td><td>Hop Count</td></tr>
<tr><td colspan="4">RREQ ID</td></tr>
<tr><td colspan="4">Destination IP Address</td></tr>
<tr><td colspan="4">Destination Sequence Number</td></tr>
<tr><td colspan="4">Originator IP Address</td></tr>
<tr><td colspan="4">Originator Sequence Number</td></tr>
</table>

Figure 3 – Route Request Packet Format RREQ

3.2 The Route Reply Packet (RREP)

Looking at the RREP packet (Figure 4), the fields are defined as follows by [3]:

Type - is equal to 2 for an RREP

Flags - used for multicasting and acknowledgment requests
Rsv – same as in RREQ
Prefix Size - If nonzero, the 5-bit Prefix Size specifies that the indicated next hop may be used for any nodes with the same routing prefix (as defined by the Prefix Size) as the requested destination
Hop Count - The number of hops from the Originator IP Address to the Destination IP Address
Destination IP Address - The IP address of the destination for which a route is supplied
Destination Sequence Number - The destination sequence number associated to the route
Originator IP Address - The IP address of the node which originated the RREQ for which the route is supplied
Lifetime - The time in milliseconds for which nodes receiving the RREP consider the route to be valid.

For the purpose of this example demonstration it is also possible to disregard all but the important information in the RREP packet while doing the walkthrough. Therefore a sample packet will look like this: <RREP, Hop Count, RREQ ID, Dest. IP, Dest. Seq. #, Orig. IP, Lifetime Orig. Seq. #>.

Type	Flags	Rsv	Prefix Sz	Hop Count
RREQ ID				
Destination IP Address				
Destination Sequence Number				
Originator IP Address				
Lifetime				

Figure 4 – Route Reply Packet Format RREP

3.3 The Soft State Routing Table

The routing tables in AODV are termed soft state because there is a timer associated with each table. This timer enables the soft state tables to expire so that routes are fresh. The soft state routing table (Figure 5) is filled with the above information that is provided in the RREQ and RREP packets. Therefore, the only new fields seen in the routing table are:

Interface – which interface the packet came in on, Next Hop – the next hop to the destination, and List of Precursors – used for upstream notification in case of link failure.

Destination IP Address
Destination Sequence Number
Valid Destination Sequence Number
Interface
Hop Count (number of hops needed to reach destination)
Next Hop
List of Precursors (described in Section 5.2)
Lifetime (expiration or deletion time of the route)
Routing Flags
State

Figure 5 – Example Routing Table

3.4 The Example Route Discovery Walkthrough

Figure 6 is the five node network that will be utilized as the example in this demonstration. The ranges of communication are designated by the circles, and are bi-directional.

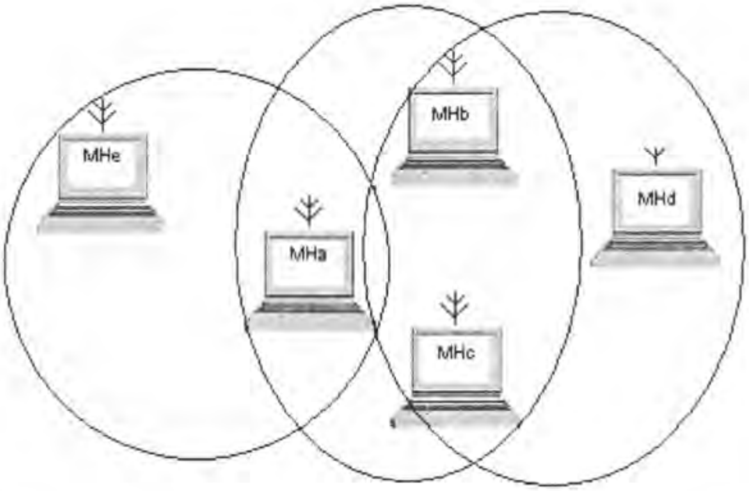

Figure 6 – Example Wireless Network For AODV Walkthrough

Mobile host A (MH_a) wishes to send a message to mobile host D (MH_d). This will be accomplished using the Ad-Hoc On Demand Distance Vector Routing protocol. The first step in this process is that MH_a generates a RREQ with its IP address in the originator field and MH_d's IP address in the destination field of the RREQ. MH_a then broadcasts the RREQ as shown in Figure 6a.

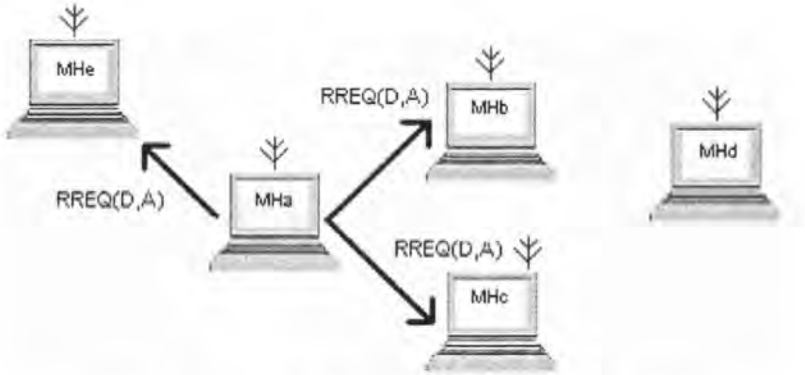

Figure 6a – The First Broadcasted RREQ

Mobile hosts E, B, and C (MH_e, MH_b, MH_c) receive the RREQ from MH_a. The three hosts that received the broadcasted RREQ will first check their routing tables to see if they currently have a route to MH_d. If they did not previously have a route they then create two new entries in their routing tables. One of the entries will be for MH_a, and the other will be for the forward path to MH_d. All three mobile hosts will then rebroadcast the RREQ as shown in Figure 6b.

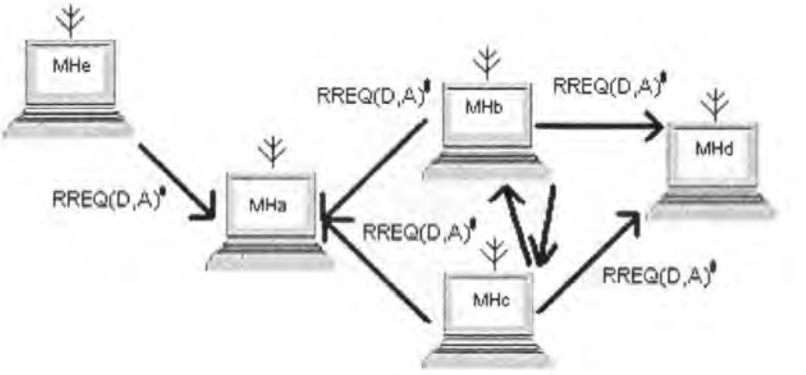

Figure 6b – The Next Broadcast of RREQ

MH$_a$ will discard the three RREQs it receives back, because of the RREQ ID and the originator address it knows this is the same RREQ it has already sent. MH$_b$ and MH$_c$ will also discard the RREQs that they receive from each other, respectively, because they know that they have also received them. However, they do add an entry for each other in their routing tables. MH$_d$ receives both RREQs and sees that MH$_a$ is trying to establish a route to it. MH$_d$ then updates its routing tables with the RREQs it received. Then, depending on which RREQ arrived first and the sequence number in that RREQ, MH$_d$ will unicast a RREP back along the forward path it has chosen, as shown in Figure 6c.

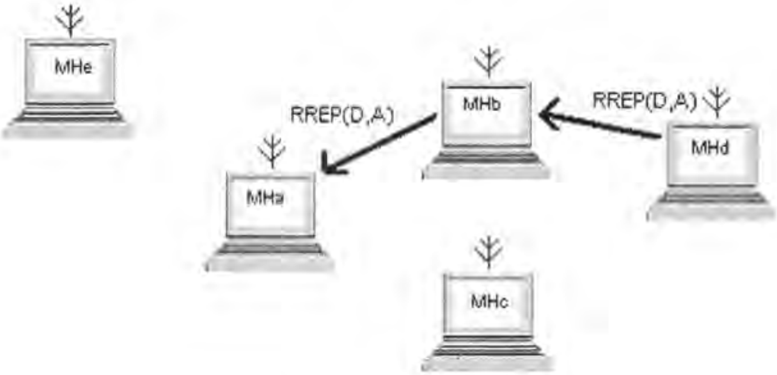

Figure 6c – The Route Reply Unicasts

As shown in Figure 6c, MH$_d$ has chosen the route through MH$_b$. Therefore it unicasts a RREP back to MH$_b$. MH$_b$ then queries its routing table and utilizes the forward path back to MH$_a$ and unicasts the RREP back to MH$_a$. Upon receiving the RREP, MH$_a$ may then begin to communicate along this route with MH$_d$. The other entries for the forward paths that were created by the RREQ in MH$_e$ and MH$_c$ will eventually timeout. However, it is useful for those entries to be present for a certain amount of time in case of a link breakage so that a new route may be quickly established.

Chapter 4: Wireless Ad-Hoc Network Security Issues

The goal of computer security is most commonly defined to be the protection of information as addressed by confidentiality, integrity, and availability [7]. Using these characteristics, the threats to wireless ad-hoc networks are classified. By developing a clearer understanding of the threats posed to wireless ad-hoc networks, the security is further improved. The following sections define the three goals of computer security and examples of threats to each characteristic.

4.1 Threats to Confidentiality

Confidentiality is a prime aspect of computer security. Only authorized parties should be granted access to assets of a computer system. The leakage of information or location of sensitive computer assets could have devastating effects on companies, countries, or militaries [8]. Even routing information in ad-hoc networks must at times be kept confidential to protect the assets. Some of the threats to confidentiality are: eavesdropping, failures in authentication, failures in key management, and compromised mobile nodes.

Eavesdropping

Wireless networks are broadcast in nature. This means all transmissions can be received by anyone within range of a transmitting node. This factor poses a significant problem to confidentiality. If mobile host A wishes to transmit a message to mobile host B, it broadcasts that message with mobile host B as the receiving address. Because most ad-hoc routing protocols do not employ encryption, mobile host C will be able to intercept the message. Even if the message contents are encrypted, mobile host C will still be able to determine that there are two hosts A and B talking. Therefore the tactic of eavesdropping in wireless ad-hoc networks is easy to perform and difficult to prevent.

The confidentiality of routing information follows very closely with eavesdropping. Because most routing protocols do not employ some sort of encryption, the routing information transmitted is clearly visible within the transmitting range. If routes are being exchanged, in some routing protocols this might give away sensitive information about location, or even disclose major activities about to take place. This information is given away by routing protocols that utilize on demand route discovery. Therefore the routing information must be kept highly confidential.

Failures in Authentication

The inherit property of an ad-hoc network is that people can join and leave a group at anytime. This poses many problems for authentication of a mobile host. The proposal to correct authentication in wireless ad-hoc networks is for a key to be given to all parties that will be forming the network before they connect [10]. This is done by letting all of the users know a certain pass phrase that will be used to authenticate all communications either through digital signatures or encryption.

The main security problem of authentication is solved by using this method. However, it leaves many security risks. The first is that the key phrase is compromised, i.e. someone leaves it written down somewhere. The second vulnerability is that a node may be compromised through theft of the mobile device. This creates a problem for confidentiality when the wrong people are authenticated or gain authentication through the aforementioned ways. They then have access to the information being shared by the ad-hoc network. This compromises confidentiality.

Failures in Key Management

Even in wired networks, key management or key distribution poses a significant problem. The overhead generated even on a wired network by key distribution makes it difficult to propose an algorithm that will efficiently distribute the keys. Even when keys are distributed, the problem remains of updating the keys and switching to the new key at the proper time such that all the traffic is being encrypted with the correct key [7]. The problem of key management is also closely tied to authentication. In wireless ad-hoc networks there is no distribution center readily available to control key exchanges, and the scalability of exchanging a symmetric key with a changing number of mobile nodes makes that scheme also difficult. These security problems are compounded by the fact that mobile nodes are much more easily compromised through theft, or if a member of the ad-hoc network leaves they still possess the ability to communicate with the group even though they no longer should, because they retain the key.

These key management problems are at the very heart of confidentiality because if key management did work easily, efficiently and correctly, only the proper parties would be able to view the messages being transmitted, whereas with the security issues of key management unauthorized users might be eavesdropping.

Compromised Mobile Nodes

Looking at two scenarios compromised mobile nodes breach confidentiality. The first scenario is the theft of a mobile device. This device might still be authorized to engage in communications with the ad-hoc group. This allows an attacker to view confidential information about other nodes present such as keys, passwords, or configurations. The second scenario is closely tied to theft, but is more military in nature. If a node is discovered by the enemy and captured, it compromises the confidentiality of the system by allowing the enemy to gain access to location and other highly sensitive information.

4.2 Threats to Integrity

The integrity of data is maintained when only authorized parties are allowed to modify the information in authorized ways. This includes routing information, files, computers, and broadcast traffic. Integrity also depends on assets not being corrupted during transmission. Threats to integrity include *man-in-the-middle attacks*, corrupted nodes, impersonation, and spoofing.

Man-in-the-Middle Attack

Man-in-the-middle attacks are a security risk in ad-hoc networks because the majority of routing protocols rely on the integrity of nodes when searching for a route. A man-in-the-middle attack occurs when a malicious user inserts itself into a route or flow of data and then gains control of the flow of traffic between two communicating nodes [7]. For example, if mobile host A wishes to talk to mobile host B, it uses route discovery or looks up in the routing table directing it how to get there. If mobile host C has maliciously inserted him or her self into the route then mobile host C will be able to see the data being passed between A and B, and either hijack the session, change or corrupt the data. Even if the payload is encrypted, mobile host C can still corrupt the payload and disrupt the integrity of the communication.

Corrupted Nodes

Corrupted nodes are either nodes being purposefully malicious by using man-in-the-middle tactics to change the payload data of a transmission, or they might be distributing false routing information. In each case the corrupted node is compromising the integrity of not only the communications, but of the entire ad-hoc network. If false routing information is being distributed by a corrupted node by stating that some nodes are unreachable or that all traffic should come through the corrupted node, then the integrity of the entire ad-hoc network is compromised.

Impersonation

When a node actively pretends to be another node on the network such that all the nodes on the network believe that it belongs to the group this is defined as impersonation. By taking the identity as a member of an ad-hoc network, a malicious user becomes a trusted part of the network. This affects the integrity of the network because this user is allowed to transmit any false information they wish, or launch a man-in-the-middle attack as a trusted user and actively corrupt data. Impersonation as defined by [8] is "...concerns all critical operations in ad-hoc network...by accessing or destroying data that is stored or being exchanged...and causing permanent damage to other nodes or services." This greatly affects the integrity of the wireless ad-hoc network because if the nodes, services, or data provided by the network are damaged, then the trust relationship that the network depends upon for its integrity is obsolete.

Spoofing

The integrity of schemes that rely on the MAC or IP address for authentication can be compromised through address spoofing. By changing the MAC or IP address a person can either impersonate a valid user and damage the integrity of the ad-hoc network, or they can again change routing information to go to the invalid or spoofed address again compromising communications.

4.3 Threats to Availability

The availability of wireless ad-hoc networks is a precious commodity. Wireless networks are advantageous over wired networks because natural disasters or man-made interruptions make it easier to take down a wired system than a wireless. However, there are still many factors affecting the availability of a wireless system. Denial of service attacks, improper key management, sleep deprivation, and frequency jamming can all affect the availability of an ad-hoc network.

Denial of Service Attacks

The availability of a wireless ad-hoc network is greatly compromised by denial of service attacks. Because of the distributed nature of an ad-hoc network, there is no main central point of failure for a malicious user to target. However, the vulnerability to ad-hoc networks lies in the routing protocols themselves. Enough traffic can be generated such that routing updates or route finding cannot take place, which prevents new nodes from joining the ad-hoc network. Existing nodes will no longer be able to communicate, effectively shutting down the availability of the wireless ad-hoc network. These attacks can also take place on a smaller scale in the form of route spoofing, impeding communication from taking place between a few nodes in the system. A malicious attacker effectively partitions an ad-hoc network in this manner.

In a tactical situation, if the ad-hoc network becomes partitioned, then smaller groups will no longer receive commands or status updates and could be easily targeted because the smaller group will have been broken away from communicating with the main force.

Improper Key Management

Key management falls into all three categories of risks to security. The availability of a network can be damaged by improper key management during times of key updates. If all the parties of the ad-hoc network do not receive the new key, or they change to the new key at the improper time, then they will no longer be able to communicate with the rest of the network. Key management might also hamper the availability of a wireless ad-hoc network if too much traffic is being generated by the key management protocol. If all network resources are being utilized to update and process key changes, then no actual data will be communicated between the parties of the ad-hoc network.

Sleep Deprivation

The main goal of sleep deprivation is to cause power drain on the battery of a mobile system. Sleep deprivation as described in [9] can occur if a host is repeatedly informed that it must stay awake. This takes place if another mobile host has been holding data for the host that wishes to sleep, and keeps sending it instructions to stay awake so it can receive this data, but then never transmitting any data. Certain protocols also force hosts to stay awake at beacon intervals. By beaconing in a manner the host never has the ability to enter sleep mode, it is deprived of its ability to conserve power. In wireless ad-hoc networks, power is a key issue. If hosts are not allowed to conserve power, they quickly drain their batteries reducing the longevity of the wireless ad-hoc network and compromising its availability.

Frequency Jamming

Frequency jamming is primarily a physical layer issue. By overpowering the frequency on which the wireless ad-hoc network is operating, because of the broadcast nature of the network, this creates a situation where no hosts are able to transmit. This can be overcome in at least two ways at the physical layer. The first is through the use of frequency hopping such that no single frequency is used for an extended period, effectively disabling jamming. The second is through digital sequence spread spectrum. This allows a transmitting node to distribute the transmissions almost making them look like the background noise.

4.4 Summary

In summary the major threats are eavesdropping, failures in key management, failures in authentication, compromised mobile nodes, man-in-the-middle attacks, corrupted nodes, impersonation, spoofing, denial-of-service attacks, sleep deprivation, and frequency jamming. As is shown there are many security threats to ad-hoc wireless networks. The next step is to examine the AODV protocol against these threats and ascertain whether or not the protocol as defined is suitable for secure communications.

Chapter 5: Analyzing the Security Weaknesses of AODV

Now that the underlying algorithm of the AODV protocol has been examined, and security issues of ad-hoc wireless networks have been explored, the next step is to take a look at how the AODV protocol measures up against security threats. This will be accomplished by looking at each security threat and seeing if AODV allows these threats to confidentiality, threats to integrity, and threats to availability to occur. Thereby, uncovering an important trust relationship in AODV.

5.1 Threats to Confidentiality

Eavesdropping

Eavesdropping can occur easily to the AODV protocol which greatly compromises the confidentiality of routing information. Because there is no requirement for encryption in the routing protocol itself, all of the messages can be seen by any party. This ensures that any 'Hello' beacons broadcast by the AODV protocol will reveal which nodes are in the network. Also, when hosts are trying to find a route to a destination, the routing information broadcast in the RREQ and RREP will be visible to a malicious eavesdropper. This routing information can not only give an attacker information as to what data is being transferred and between whom, but it also lets the attackers know when an increase in communication is taking place.

Another key point is that AODV provides upstream notification when a link is broken. Figure 7 shows an example network where MH_e moves from behind MH_c to behind MH_d. When this occurs, MH_c notifies MH_a that MH_e is no longer in the same position. MH_b

will overhear this notification and see the new route created through MH_d. Therefore, an eavesdropper will know when a node is moving, and can determine where it resides when the new path is discovered, making the mobile host vulnerable to physical attacks.

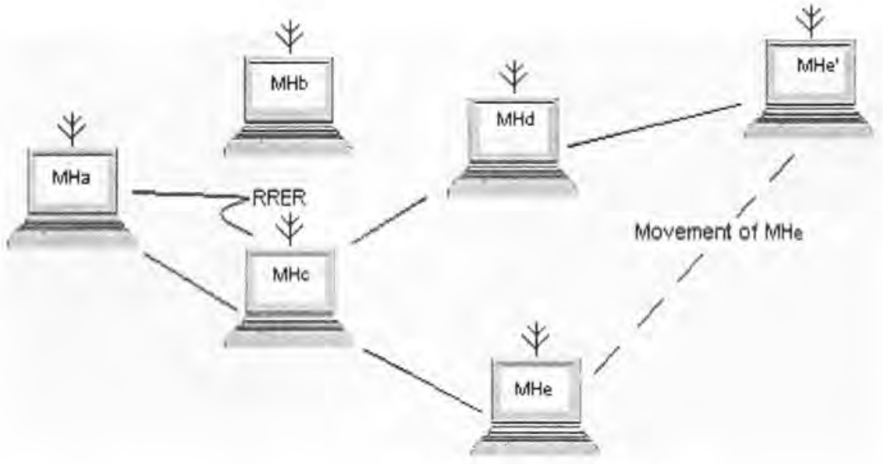

Figure 7 – Example of Movement Notification

This gives an eavesdropper the ability to discern movement within the network and uncover directions, locations, or missions the group is moving towards. A possible solution for this is to utilize IPSEC (IP Security) [7] or some other form of payload encryption, because the AODV protocol is wrapped inside of the IP protocol.

Authentication

The AODV protocol does not address authentication of mobile hosts. However, it does provide, in a section at the end of the protocol, a description that if there are pre-established security associations, then AODV should be able to use the same authentication mechanisms based on their IP address [3]. This does not allow for strong authentication. Authentication based on IP has proved time and again to be a weak and insecure form of authentication, because of the ease of spoofing an IP address. Therefore authentication is left up to other layers by AODV. This leaves the confidentiality of the network in jeopardy if mobile hosts are not authenticated. However, if AODV is running on top of IEEE 802.11 [11] there is a scheme present for authentication in that protocol, but it still contains many security holes. The authentication provided by IEEE 802.11 is merely a pass/challenge scheme that authenticates in one direction and for only one hop.

Key Management

The issue of key management is closely tied to the authentication issue in AODV. Because there is no specific way to authenticate, key management cannot be trusted, because none of the mobile hosts are authenticated including the key manager. This creates a chasm in the security of AODV. If keys cannot be distributed, then no traffic in the network can be encrypted, and therefore the confidentiality of the network is compromised. However, this might be circumvented using strong application layer end-to-end encryption. The problem with this approach is that it does not scale well and utilizes a large amount of bandwidth.

Compromised Mobile Nodes

The confidentiality exposed when a mobile node is compromised by theft is near impossible to contain in any protocol. AODV does take one step that allows for the notification of route failures in a blacklist set. This could be adapted to allow for mobile nodes to blacklist any rogue nodes and no longer utilize them in data transmissions or path findings.

5.2 Threats to Integrity

Man-in-the-Middle Attacks

The AODV protocol is vulnerable to man-in-the-middle attacks, and in fact, these attacks can be carried out with ease. Because the traffic is not encrypted and little or no authentication of the nodes takes place, it is easy for a node to insert itself into a route between two hosts. This can be accomplished in AODV because there is little check on the sequence numbers used to establish which route is the best. If the attacker can determine the other metrics as well as the sequence number used to choose the best route, then there is nothing in the protocol to prevent a malicious host from asserting itself as having the best route. This is shown in Figure 8.

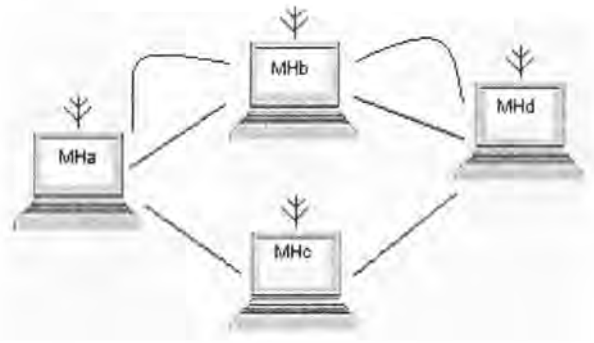

Figure 8 – Example of Man-in-the-Middle Attack

The nodes of this network are MH_a, MH_d, and MH_c. MH_b is the malicious attacker. When the RREQ goes out from MH_a looking for a route to MH_d, MH_b forwards the RREQ to MH_d, but gives the best metric for the route. Therefore, when MH_d sends the RREP back to MH_a the route it chooses goes through MH_b. This compromises not only the integrity of the route, but the data as well if the attacker changes any data as it passes through its hop on the route.

Corrupted Nodes

A corrupted node can not only damage the integrity of the route, but the integrity of the data as well. Little information is provided in the protocol to protect in the integrity of the data except for what is being used by other layers. The integrity of the route can be compromised by a node stating that it has the best route, and then never establishing that route for the source node. It could also pass along corrupted information as part of the route.

Impersonation and Spoofing

Impersonation and spoofing greatly affect the integrity of the ad-hoc network. In AODV there is a large trust relationship that a mobile host is providing the correct IP address. This trust relationship will be explored later, but if a malicious attacker spoofs its IP or impersonates another mobile host, there is very little way to know this. Therefore a malicious user can easily exploit the IP trust relationship and collect and disseminate false data that other users believe to be correct. Looking at Figure 9, we see an example network. Mobile host A (MH_a) is the mobile host that is being spoofed. Mobile host A' is the mobile host that is spoofing A.

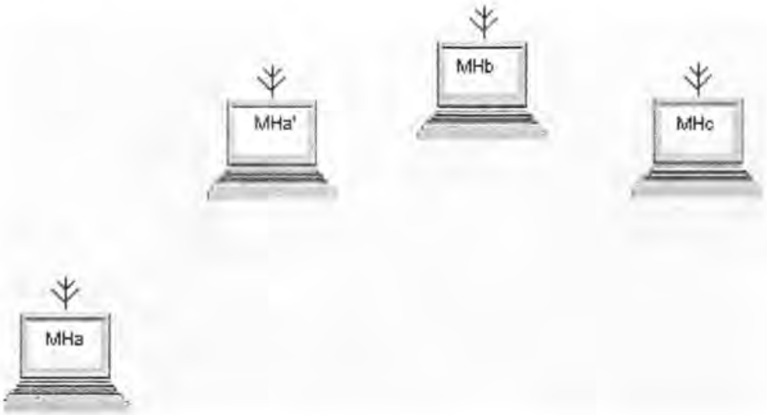

Figure 9 – Example Network for Spoofing

In this example for MH_a' to spoof MH_a it need only to initiate a RREQ with MH_c as the destination. In the packets MH_a' will provide the IP of MH_a and the MAC of MH_a. When the RREQ goes out from MH_a' it is seen by MH_b and eventually passed on to MH_c, and the RREP is sent to establish the route. MH_c has no reason not to trust MH_a because of the trust relationship built on IP addresses. Even when MH_a sees the RREQ coming from MH_a' it will not suspect a security breach as long as the sequence number is large enough. MH_a will simply discard the RREQ because the originator's IP address is itself. The feature of the protocol that helps prevent multiple RREQs from storming the network is therefore used against itself during the impersonation.

5.3 Threats to Availability

Denial of Service Attacks

Analysis of AODV's susceptibility to denial of service reveals two major schemes for attacking AODV. The first attack is perpetrated if enough traffic is generated by a malicious user or users, then routing information such as RREQs and RREPs cannot be exchanged and no new routes can be established. Also, there may be key nodes through which all information is flowing, and if this node cannot communicate because a malicious hacker queries false routes from it, then the ad-hoc network is effectively partitioned. The second way that a denial of service can take place is by a malicious hacker responding to a RREQ with the best metric route, and then never forwarding the traffic once data is sent. This also takes place if the attacker spoofs itself to be the destination and then discards all information headed for that destination.

Improper Key Management

The AODV protocol neither helps nor hinders the possibility of improper key management, but it does lead to availability issues. Because the protocol does not specify a key management strategy then it would have to lie either below or above AODV in the protocol stack. If the key management strategy lies below AODV in the protocol stack, then AODV will not affect the key management. However, if the key management lies above AODV in the protocol stack, then steps would have to be taken to assure that proper authentication is taking place in the distribution of keys because of the IP trust relationship that AODV relies upon.

Sleep Deprivation and Frequency Jamming

Sleep deprivation and frequency jamming, while great threats to network security, lie more in the physical and data link layer realms of the ad-hoc network. However, because AODV utilizes the mobile hosts to route information, the mobile hosts must stay awake longer, and therefore consume more power than in a single hop network. There is no mechanism in the protocol for a mobile host to only be able to receive data and not help pass it along other than always discarding RREQs so that it will not be used in the route, but its neighbors are still able to reach it because of the 'Hello' beacon that it broadcasts periodically.

5.4 Trust Relationships

Exposing a trust relationship is one of the greatest threats to the security of a protocol. In AODV there is a core trust relationship believing the correctness of the IP addresses of the mobile nodes. As is evidenced in the previous threats, almost all of the attacks occur because the mobile hosts in the network believe the mobile host's authenticity based on the IP address given in a packet. The features of the protocol that prevent RREQ flooding are also used against themselves when a mobile host sees its own IP address on an incoming RREQ and discards the packet. It is because of this trust relationship that AODV is so open to spoofing attacks and any of the above attacks that take advantage of the IP of another mobile host. A method to prevent this is described in Secure AODV (SADOV) [12]. However, the approach taking in SADOV only utilizes a signature on messages. Because of the nature of wireless ad-hoc networks, this is not a very viable solution. How can a corrupted mobile host be prevented from assigning or verifying the signatures? The error prone transmission nature of wireless networks makes it difficult to distribute signatures to all the mobile hosts in a large scale. One way to resolve the problem of the IP trust relationship in AODV is to rely on the data link layer to provide authentication for all mobile hosts. The problem with this approach, however is that all mobile hosts must be within one hop of a controlling station to be authenticated, but then the network would no longer be truly ad-hoc. There is no easy solution to the trust relationship that AODV is built upon. Therefore as it stands AODV is not an ideal protocol for high security applications.

Chapter 6: Conclusion

6.1 Summary

The main objective of this thesis was the systematic study of a popular wireless network routing protocol in order to evaluate its ability to support secure communications. In support of that goal this thesis accomplished four things. It first examined current ad-hoc network routing protocols with regard to the underlying algorithms for route discovery and message delivery. To accomplish this goal this thesis examines five network routing algorithms for wireless ad-hoc networks: DSDV, WRP, AODV, TORA, and DSR. This portion of the thesis summarized the operation of each protocol in order to give the reader a foundation of wireless routing protocols to further the understanding the state of research that is being conducted in the wireless environment.

The next section of this thesis entailed a detailed demonstration of the route discovery of one of the more popular routing protocols, the AODV protocol. This section gave the reader a deeper understanding of the AODV protocol and the underlying operations that occur in a route discovery of AODV. The thesis detailed the types of messages that are exchanged, and how these exchanges occur.

From there the security vulnerabilities of the wireless networking environment were described, and finally the security vulnerabilities of the wireless environment were mapped to the AODV protocol for analysis of AODV's security properties. The next step that this thesis took was to examine the wireless medium, and the security challenges that are faced in the wireless medium. This creates a foundation of understanding that allows for the AODV protocol to be compared to for security weakness. This thesis demonstrated that for the three main goals of security confidentiality, integrity, and availability the AODV protocol is not designed to support secure communications.

AODV as it stands by itself is unable to handle the problems of eavesdropping, failures in authentication, failures in key management, the compromising of mobile nodes, man-in-the-middle attacks, corrupted nodes, impersonation and spoofing, and denial of service attacks. As is shown in the previous chapter there are some measures that can be taken to increase the security for each problem, but the underlying AODV protocol does nothing for secure communications. It is therefore concluded that the AODV protocol is not a viable protocol for communicating in a wireless ad-hoc network.

6.2 Future Work

The fact that wireless ad-hoc networks is a relatively new area of research creates an opportunity for plethora of further work in this area. Three directions that future work can be taken from this thesis are: the simulation of the security vulnerabilities, enhancing the AODV protocol for security, and a comparison of AODV against other ad-hoc protocols for security.

Simulating the security vulnerabilities of AODV can take two different paths. The first way that the simulations can take place is through the use of a network simulator tool such as NS. This would allow a researcher to easily implement the protocol and run controlled experiments to determine the ability of AODV to carry on secure communications. By running the simulations through the NS tool each scenario can be easily carried out and quickly adapted to new situations. The other way in which these vulnerabilities can be simulated is through actual implementation of the AODV protocol on a number of machines and carrying out each attack on the protocol. The relative merits of both ways are fairly similar because research between using a simulator tool and actual implementation has shown them to be fairly close approximations.

Another vein for future work is enhancing the AODV protocol for security. Steps can be taken to correct some of the oversights in security that AODV contains. If IPSEC is utilized during communications this can dramatically decrease the ability to eavesdrop data, routing information, and man-in-the-middle attacks. By implementing upper layer security measures other security vulnerabilities can also be circumvented. Another direction that this future work can take is to actually alter the protocol with security in mind trying to fix some of the inherent trust relationships that are built into AODV.

The final area for future work that this thesis will discuss, comparing AODV against other protocols for security can take on many different directions as well. Not only can the comparisons be made by analyzing the algorithms of the other routing protocols in the same way that this thesis did and creating a matrix of vulnerabilities that each protocol can be rated against, but each protocol can be simulated against the security vulnerabilities to determine whether or not they can be utilized for secure communications.

Given the nature of ad-hoc networks and the inherent trust that they are built upon, securing these networks will be a difficult task. As was detailed in this paper there are many vulnerabilities in the wireless environment that make secure computing difficult, but this is a necessary task in current age that we are in given the worth of information in today's businesses, governments, and militaries.

Bibliography

[1] Toh, C.K. Ad Hoc Mobile Wireless Networks: Protocols and Systems. Upper Saddle River: Prentice Hall PTR, 2002.

[2] C. Perkins and P. Bhagwat. Highly Dynamic Destination-Sequenced Distance-Vector Routing (DSDV) for Mobile Computers. In Proc. of the ACM SIGCOMM, Oct. 1994. http://citeseer.nj.nec.com/perkins94highly.htm. (Accessed: Aug. 2002).

[3] C. Perkins and Elizabeth Belding. Ad hoc On-Demand Distance Vector (AODV) Routing INTERNET DRAFT, Nokia Research Center, 19 June 2002 http://www.cs.ucsb.edu/~ebelding/txt/aodvid.txt. (Accessed: Aug. 2002).

[4] V. D. Park and M. S. Corson, "A highly adaptive distributed routing algorithm for mobile wireless networks," in IEEE Infocom, 1997. http://citeseer.nj.nec.com/park97highly.html. (Accessed: Sept. 2002).

[5] D. Johnson and D. Maltz, "Dynamic source routing in ad hoc wireless networks," in Mobile Computing (ed. T. Imielinski and H. Korth), Kluwer Academic Publishers, Dordrecht, The Netherlands, 1996. http://citeseer.nj.nec.com/johnson96dynamic.html. (Accessed: Aug. 2002).

[6] Shree Murthy and J.J. Garcia-Luna-Aveces, "A Routing Protocol for Packet Radio Networks," Proc. ACM International Conference on Mobile Computing and Networking, pp. 86-95, November, 1995. http://citeseer.nj.nec.com/murthy95routing.html. (Accessed: May 2002).

[7] Pfleeger, Charles P. Security in Computing. Upper Saddle River, NJ: Prentice Hall PTR, 2000.

[8] L. Zhou and Z. J. Haas. Securing Ad Hoc Networks. IEEE Network Magazine, vol. 13, no.6, November/December 1999. http://citeseer.nj.nec.com/zhou99securing.html. (Accessed: Oct. 2002).

[9] F. Stajano and R. Anderson. The Resurrecting Duckling: Security Issues for Ad-hoc Wireless Networks. B. Christianson, B. Crispo, and M. Roe (Eds.)., Security Protocols, 7th International Workshop Proceedings, Lecture Notes in Computer Science, 1999. http://citeseer.nj.nec.com/stajano99resurrecting.html. (Accessed: May 2002).

[10] Schiller, Jochen. Mobile Communications. Great Britain: Addison-Wesley, 2000.

[11] IEEE 802.11 Protocol. http://standards.ieee.org/wireless/overview.html#802.11. (Accessed: Sept. 2002).

[12] Manel Guerrero Zapata. "Secure Ad hoc On-Demand Distance Vector (SAODV) Routing". IETF Internet Draft, http://www.cs.ucsb.edu/~ebelding/txt/saodv.txt. (Accessed: Jan. 2003).

[13] Paul Black. NIST Dictionary of Algorithms and Data Structures. Bellman Ford. http://www.nist.gov/dads/HTML/bellmanford.html. (Accessed: Jan. 2003).