WIRELESS/MOBILE NETWORK SECURITY
Y. Xiao, X. Shen, and D.-Z. Du (Eds.) pp. – - –

Chapter 12

A Survey on Attacks and Countermeasures in Mobile Ad Hoc Networks

Bing Wu, Jianmin Chen, Jie Wu, Mihaela Cardei
Department of Computer Science and Engineering
Florida Atlantic University
E-mail: {bwu, jchen8}@fau.edu, {jie, mihaela}@cse.fau.edu

1 Abstract

Security is an essential service for wired and wireless network communications. The success of mobile ad hoc networks (MANET) strongly depends on people's confidence in its security. However, the characteristics of MANET pose both challenges and opportunities in achieving security goals, such as confidentiality, authentication, integrity, availability, access control, and non-repudiation. We provide a survey on attacks and countermeasures in MANET in this paper. The countermeasures are features or functions that reduce or eliminate security vulnerabilities and attacks. First, we give an overview of attacks according to the protocols stacks, and to security attributes and mechanisms. Then we present preventive approaches following the order of the layered protocol stacks. We also put forward an overview of MANET intrusion detection systems (IDS), which are reactive approaches to thwart attacks and used as a second line of defense.

2 Introduction

In a MANET, a collection of mobile hosts with wireless network interfaces form a temporary network without the aid of any fixed infrastructure or centralized administration. A MANET is referred to as an infrastructureless network because the mobile nodes in the network dynamically set up paths among themselves to transmit

packets temporarily. In a MANET, nodes within each other's wireless transmission ranges can communicate directly; however, nodes outside each other's range have to rely on some other nodes to relay messages [22]. Thus, a multi-hop scenario occurs, where several intermediate hosts relay the packets sent by the source host before they reach the destination host. Every node functions as a router. The success of communication highly depends on other nodes' cooperation.

In 1996, The Internet Engineering Task Force(IETF) set down a MANET workgroup, and its goal is to standardize IP routing protocol functionality suitable for wireless routing applications within both static and dynamic topologies.

A MANET is an autonomous system of mobile nodes. The system may operate in isolation, or may have gateways and interface with a fixed network. Its nodes are equipped with wireless transmitters/receivers using antennas which may be omnidirectional (broadcast), highly-directional (point-to-point), or some combination thereof. At a given time, the system can be viewed as a random graph due to the movement of the nodes, their transmitter/receiver coverage patterns, the transmission power levels, and the co-channel interference levels. The network topology may change with time as the nodes move or adjust their transmission and reception parameters. Thus, a MANET has several salient characteristics [21]: dynamic topologies, resource constraints, limited physical security, and no infrastructure.

Possible applications of MANET include: Soldiers relaying information for situational awareness on the battlefield, business associates sharing information during a meeting; attendees using laptop computers to participate in an interactive conference; and emergency disaster relief personnel coordinating efforts after a fire, hurricane, or earthquake. The other possible applications [22] include personal area and home networking, location-based services, and sensor networks.

There are a wide variety of attacks that target the weakness of MANET. For example, routing messages are an essential component of mobile network communications, as each packet needs to be passed quickly through intermediate nodes, which the packet must traverse from a source to the destination. Malicious routing attacks can target the routing discovery or maintenance phase by not following the specifications of the routing protocols. There are also attacks that target some particular routing protocols, such as DSR, or AODV [10] [20]. More sophisticated and subtle routing attacks have been identified in recent published papers, such as the blackhole (or sinkhole) [35], Byzantine [17], and wormhole [15] [32] attacks. Currently routing security is one of the hottest research areas in MANET.

This paper is structured as follows. In Section 3 we describe the attacks using a hybrid model of OSI and TCP/IP called the Tanenbaum model, which has five layers: application, transport, network, data link, physical. In Section 4, we overview attack countermeasures, including intrusion detection and co-operation enforcement at different network layers. In section 5, we discuss open challenges

Table 1: Security Attacks Classification

Passive Attacks	Eavesdropping, traffic analysis, monitoring
Active Attacks	Jamming, spoofing, modification, replaying, DoS

and future directions briefly.

3 Security attacks

The attacks in MANET can roughly be classified into two major categories, namely passive attacks and active attacks, according to the attack means [9] [23]. A passive attack obtains data exchanged in the network without disrupting the operation of the communications, while an active attack involves information interruption, modification, or fabrication, thereby disrupting the normal functionality of a MANET. Table 1 shows the general taxonomy of security attacks against MANET. Examples of passive attacks are eavesdropping, traffic analysis, and traffic monitoring. Examples of active attacks include jamming, impersonating, modification, denial of service (DoS), and message replay.

The attacks can also be classified into two categories, namely external attacks and internal attacks, according the domain of the attacks. Some papers refer to outsider and insider attacks [39]. External attacks are carried out by nodes that do not belong to the domain of the network. Internal attacks are from compromised nodes, which are actually part of the network. Internal attacks are more severe when compared with outside attacks since the insider knows valuable and secret information, and possesses privileged access rights.

Attacks can also be classified according to network protocol stacks. Table 2 shows an example of a classification of security attacks based on protocol stack; some attacks could be launched at multiple layers.

Some security attacks use stealth [34], whereby the attackers try to hide their actions from either an individual who is monitoring the system or an intrusion detection system (IDS). But other attacks such as DoS cannot be made stealth. Some attacks are non-cryptography related, and others are cryptography primitive attacks. Table 3 shows cryptography primitive attacks and some examples.

For the rest of the section, we present a survey of security attacks in MANET following the order of the protocol stacks. Physical layer attacks are discussed in Section 3.1, followed by link layer attacks in Section 3.2; and network layer attacks in Section 3.3. Transport layer attacks are discussed in Section 3.4, application layer attacks are discussed in Section 3.5, and multi-layer attacks are discussed in

Table 2: Security Attacks on Protocol Stacks

Layer	**Attacks**
Application layer	Repudiation, data corruption
Transport layer	Session hijacking, SYN flooding
Network layer	Wormhole, blackhole, Byzantine, flooding, resource consumption, location disclosure attacks
Data link layer	Traffic analysis, monitoring, disruption MAC (802.11), WEP weakness
Physical layer	Jamming, interceptions, eavesdropping
Multi-layer attacks	DoS, impersonation, replay, man-in-the-middle

Table 3: Cryptography Primitive Attacks

Cryptography Primitive Attacks	**Examples**
Pseudorandom number attack	Nonce, timestamp, initialization vector (IV)
Digital signature attack	RSA signature, ElGamal signature, digital signature standard (DSS)
Hash collision attack	SHA-0, MD4, MD5, HAVAL-128, RIPEMD
Security handshake attacks	Diffie-Hellman key exchange protocol, Needham-Schroeder protocol

Section 3.6. Cryptography primitive attacks are discussed in Section 3.7.

3.1 Physical layer attacks

Eavesdropping is the intercepting and reading of messages and conversations by unintended receivers. The mobile hosts in mobile ad hoc networks share a wireless medium. The majorities of wireless communications use the RF spectrum and broadcast by nature. Signals broadcast over airwaves can be easily intercepted with receivers tuned to the proper frequency [47] [48]. Thus, messages transmitted can be eavesdropped, and fake messages can be injected into network.

Moreover, a radio signal can be jammed or interfered, which causes the message to be corrupted or lost [47] [48]. If the attacker has a powerful transmitter, a signal can be generated that will be strong enough to overwhelm the targeted signals and disrupt communications. The most common types of this form of signal jamming are random noise and pulse. Jamming equipment is readily available. In addition, jamming attacks can be mounted from a location remote to the target networks.

3.2 Link layer attacks

The MANET is an open mulitpoint peer-to-peer network architecture. Sepcifically, one-hop connectivity among neighbors is maintained by the link layer protocols, and the network layer protocols extend the connectivity to other nodes in the network. Attacks may target the link layer by disrupting the cooperation of the layer's protocols.

3.2.1 Wireless MAC protocol

Wireless medium access control (MAC) protocols have to coordinate the transmissions of the nodes on the common transmission medium. Because a token-passing bus MAC protocol is not suitable for controlling a radio channel, IEEE 802.11 protocol is specifically devoted to wireless LANs. The IEEE 802.11 MAC protocol uses distributed contention resolution mechanisms for sharing the wireless channel. The IEEE 802.11 work group proposed two algorithms for contention resolution. One is a fully distributed access protocol called the distributed coordination function (DCF). The other is a centralized access protocol called the point coordination function (PCF).

PCF requires a central decision maker such as a base station. DCF uses a carrier sense multiple access/collision avoidance protocol (CSMA/CA) for resolving channel contention among multiple wireless hosts.

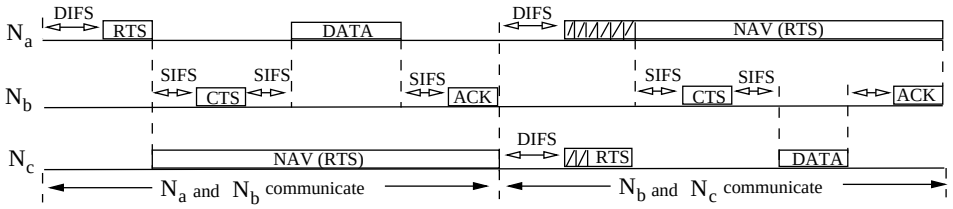

Figure 1: Illustration of Channel Contention in 802.11 MAC

Three values for interframe space (IFS) are defined to provide priority-based access to the radio channel [27]. SIFS is the shortest interframe space and is used for ACK, CTS and poll response frames. DIFS is the longest IFS and is used as the minimum delay for asynchronous frames contending for access. PIFS is the middle IFS and is used for issuing polls by the centralized controller in the PCF scheme. In case there is a collision, the sender waits a random unit of time, based on the binary exponential backoff algorithm, before retransmitting. In Figure 1, node Na and node Nc contend to communicate with node Nb. First node Na gets access and reserves the channel, and then Nc succeeds and reserves the channel while node Na has to back off [30].

3.2.2 Traffic monitoring and analysis

Traffic monitoring and analysis can be deployed to identify the communication parties and functionalities, which could provide information to launch further attacks [28]. Since these attacks are not specific to the MANET, other wireless networks, such as the cellular network, satellite network, and WLAN also suffer from these potential vulnerabilities. We did not focus on attacks on this layer for the security of MANET.

3.2.3 Disruption on MAC DCF and backoff mechanism

As above description, current wireless MAC protocols assume cooperative behaviors among all nodes. Obviously the malicious or selfish nodes are not forced to follow the normal operation of the protocols. In the link layer, a selfish or malicious node could interrupt either contention-based or reservation-based MAC protocols.

A malicious neighbor of either the sender or the receiver could intentionally not follow the protocol specifications. For example, the attacker may corrupt the frames easily by introducing some bits or ignoring the ongoing transmission. It could also just wait SIFS or exploit its binary exponential backoff scheme to launch DoS attacks in IEEE 802.11 MAC. The binary exponential scheme favors the last

winner amongst the contending nodes. This leads to what is called the capture effect [21]. Nodes that are heavily loaded tend to capture the channel by continually transmitting data, thereby causing lightly loaded neighbors to backoff endlessly. Malicious nodes could take advantage of this capture effect vulnerability. Moreover, a backoff at the link layer can incur a chain reaction in any upper layer protocols that use a backoff scheme, like TCP window management.

The network allocation vector (NAV) field carried in RTS/CTS frames exposes another vulnerability to DoS attacks in the link layer [21] [29]. Initially the NAV field was proposed to mitigate the hidden terminal problem in the carrier sense mechanism. During the RTS/CTS handshake the sender first sends a small RTS frame containing the time needed to complete the CTS, data, and ACK frames. Each neighbor of the sender and receiver will update the NAV field and defer their transmission for the duration of the future transaction according to the time that they overheard. An attacker may also overhear the NAV information and then intentionally corrupt the link layer frame through wireless interference to the ongoing transmission.

3.2.4 Weakness of 802.11 WEP

IEEE 802.11 incorporates wired equivalent privacy (WEP) to provide WLAN systems a modest level of privacy by encrypting radio signals. A secondary purpose of WEP is to prevent unauthorized users from accessing WLAN. 802.11 standards support WEP cryptographic keys of 40 bits, though some vendors have implemented 104 bits and even 128 bits. It is well known that WEP has a number of weaknesses and is subject to attacks. Some of the weaknesses are listed below [27] [28] [47],

- WEP protocol does not specify key management.
- The initialization vector (IV) is a 24-bit field sent in clear and is part of the RC4 encryption key. The reuse of IV and the weakness of RC4 lead to analytic attacks.
- The combined use of a non-cryptographic integrity algorithm, CRC 32, with the stream cipher is a security risk.

3.3 Network layer attacks

Network layer protocols extend connectivity from neighboring 1-hops nodes to all other nodes in MANET. The connectivity between mobile hosts over a potentially

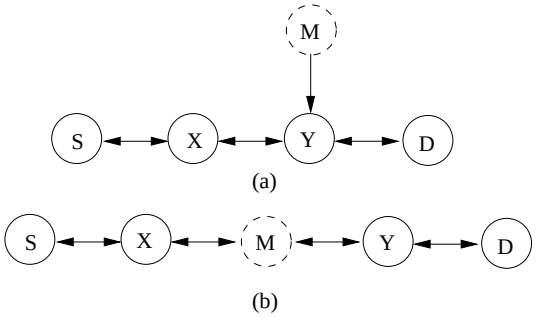

Figure 2: Illustration of Routing Attack

multi-hop wireless link strongly relies on cooperative reactions among all network nodes.

A variety of attacks targeting the network layer have been identified and heavily studied in research papers. By attacking the routing protocols, attackers can absorb network traffic, inject themselves into the path between the source and destination, and thus control the network traffic flow, as shown in Figure 2 (a) and (b), where a malicious node M can inject itself into the routing path between sender S and receiver D.

The traffic packets could be forwarded to a non-optimal path, which could introduce significant delay. In addition, the packets could be forwarded to a nonexistent path and get lost. The attackers can create routing loops, introduce severe network congestion, and channel contention into certain areas. Multiple colluding attackers may even prevent a source node from finding any route to the destination, causing the network to partition, which triggers excessive network control traffic, and further intensifies network congestion and performance degradation.

3.3.1 Attacks at the routing discovery phase

There are malicious routing attacks that target the routing discovery or maintenance phase by not following the specifications of the routing protocols. Routing message flooding attacks, such as hello flooding, RREQ flooding, acknowledgement flooding, routing table overflow, routing cache poisoning, and routing loop are simple examples of routing attacks targeting the route discovery phase [6] [35]. Proactive routing algorithms, such as DSDV [22] and OLSR [45], attempt to discover routing information before it is need, while reactive algorithms, such as DSR [22] and AODV [22], create routes only when they are needed. Thus, proactive algorithms are more vulnerable to routing table overflow attacks. Some of these attacks are listed below.

- **Routing table overflow attack:** A malicious node advertises routes that go to non-existent nodes to the authorized nodes present in the network. It usually happens in proactive routing algorithms, which update routing information periodically. The attacker tries to create enough routes to prevent new routes from being created. The proactive routing algorithms are more vulnerable to table overflow attacks because proactive routing algorithms attempt to discover routing information before it is actually needed. An attacker can simply send excessive route advertisements to overflow the victim's routing table.

- **Routing cache poisoning attack:** In route cache poisoning attacks, attackers take advantage of the promiscuous mode of routing table updating, where a node overhearing any packet may add the routing information contained in that packet header to its own route cache, even if that node is not on the path. Suppose a malicious node M wants to poison routes to node X. M could broadcast spoofed packets with source route to X via M itself; thus, neighboring nodes that overhear the packet may add the route to their route caches.

3.3.2 Attacks at the routing maintenance phase

There are attacks that target the route maintenance phase by broadcasting false control messages, such as link-broken error messages, which cause the invocation of the costly route maintenance or repairing operation. For example, AODV and DSR implement path maintenance procedures to recover broken paths when nodes move. If the destination node or an intermediate node along an active path moves, the upstream node of the broken link broadcasts a route error message to all active upstream neighbors. The node also invalidates the route for this destination in its routing table. Attackers could take advantage of this mechanism to launch attacks by sending false route error messages.

3.3.3 Attacks at data forwarding phase

Some attacks also target data packet forwarding functionality in the network layer. In this scenario the malicious nodes participate cooperatively in the routing protocol routing discovery and maintenance phases, but in the data forwarding phase [18] [33] they do not forward data packets consistently according to the routing table. Malicious nodes simply drop data packets quietly, modify data content, replay, or flood data packets; they can also delay forwarding time-sensitive data packets selectively or inject junk packets.

3.3.4 Attacks on particular routing protocols

There are attacks that target some particular routing protocols. In DSR, the attacker may modify the source route listed in the RREQ or RREP packets. It can delete a node from the list, switch the order, or append a new node into the list. In AODV, the attacker may advertise a route with a smaller distance metric than the actual distance, or advertise a routing update with a large sequence number and invalidate all routing updates from other nodes.

3.3.5 Other advanced attacks

More sophisticated and subtle routing attacks have been identified in recent research papers. The blackhole (or sinkhole), Byzantine, and wormhole attacks are the typical examples, which are described in detail below.

- **Wormhole attack:** An attacker records packets at one location in the network and tunnels them to another location. Routing can be disrupted when routing control messages are tunneled. This tunnel between two colluding attackers is referred as a wormhole [8] [32]. Wormhole attacks are severe threats to MANET routing protocols. For example, when a wormhole attack is used against an on-demand routing protocol such as DSR or AODV, the attack could prevent the discovery of any routes other than through the wormhole.

- **Blackhole attack:** The blackhole attack has two properties. First, the node exploits the mobile ad hoc routing protocol, such as AODV, to advertise itself as having a valid route to a destination node, even though the route is spurious, with the intention of intercepting packets. Second, the attacker consumes the intercepted packets without any forwarding. However, the attacker runs the risk that neighboring nodes will monitor and expose the ongoing attacks. There is a more subtle form of these attacks when an attacker selectively forwards packets. An attacker suppresses or modifies packets originating from some nodes, while leaving the data from the other nodes unaffected, which limits the suspicion of its wrongdoing.

- **Byzantine attack:** A compromised intermediate node works alone, or a set of compromised intermediate nodes works in collusion and carry out attacks such as creating routing loops, forwarding packets through non-optimal paths, or selectively dropping packets, which results in disruption or degradation of the routing services [17].

- **Rushing attack:** Two colluded attackers use the tunnel procedure to form a wormhole. If a fast transmission path (e.g. a dedicated channel shared by attackers) exists between the two ends of the wormhole, the tunneled packets can propagate faster than those through a normal multi-hop route. This forms the rushing attack [19]. The rushing attack can act as an effective denial-of-service attack against all currently proposed on-demand MANET routing protocols, including protocols that were designed to be secure, such as ARAN and Ariadne [20].

- **Resource consumption attack:** This is also known as the sleep deprivation attack. An attacker or a compromised node can attempt to consume battery life by requesting excessive route discovery, or by forwarding unnecessary packets to the victim node.

- **Location disclosure attack:** An attacker reveals information regarding the location of nodes or the structure of the network. It gathers the node location information, such as a route map, and then plans further attack scenarios. Traffic analysis, one of the subtlest security attacks against MANET, is unsolved. Adversaries try to figure out the identities of communication parties and analyze traffic to learn the network traffic pattern and track changes in the traffic pattern. The leakage of such information is devastating in security-sensitive scenarios.

3.4 Transport layer attacks

The objectives of TCP-like Transport layer protocols in MANET include setting up of end-to-end connection, end-to-end reliable delivery of packets, flow control, congestion control, clearing of end-to-end connection. Similar to TCP protocols in the Internet, the mobile node is vulnerable to the classic SYN flooding attack or session hijacking attacks. However, a MANET has a higher channel error rate when compared with wired networks. Because TCP does not have any mechanism to distinguish between whether a loss was caused by congestion, random error, or malicious attacks, TCP multiplicatively decreases its congestion window upon experiencing losses, which degrades network performance significantly [49].

- **SYN flooding attack:** The SYN flooding attack is a denial-of-service attack. The attacker creates a large number of half-opened TCP connections with a victim node, but never completes the handshake to fully open the connection.

  For two nodes to communicate using TCP, they must first establish a TCP connection using a three-way handshake. The three messages exchanged during the handshake, illustrated in Figure 3, allow both nodes to learn that

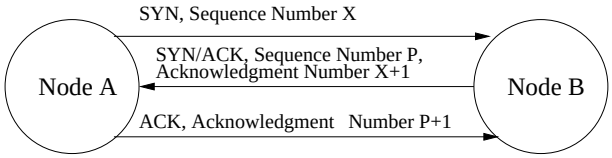

Figure 3: TCP Three-way Handshake

the other is ready to communicate and to agree on initial sequence numbers for the conversation.

During the attack, a malicious node sends a large amount of SYN packets to a victim node, spoofing the return addresses of the SYN packets. The SYN-ACK packets are sent out from the victim right after it receives the SYN packets from the attacker and then the victim waits for the response of ACK packet. Without any response of ACK packets, the half-open data structure remains in the victim node. If the victim node stores these half-opened connections in a fixed-size table while it awaits the acknowledgement of the three-way handshake, all of these pending connections could overflow the buffer, and the victim node would not be able to accept any other legitimate attempts to open a connection. Normally there is a time-out associated with a pending connection, so the half-open connections will eventually expire and the victim node will recover. However, malicious nodes can simply continue sending packets that request new connections faster than the expiration of pending connections.

- **Session hijacking:** Session hijacking takes advantage of the fact that most communications are protected (by providing credentials) at session setup, but not thereafter. In the TCP session hijacking attack, the attacker spoofs the victim's IP address, determines the correct sequence number that is expected by the target, and then performs a DoS attack on the victim. Thus the attacker impersonates the victim node and continues the session with the target.

  The TCP ACK storm problem, illustrated in Figure 4, could be created when an attacker launches a TCP session hijacking attack. The attacker sends injected session data, and node A will acknowledge the receipt of the data by sending an ACK packet to node B. This packet will not contain a sequence number that node B is expecting, so when node B receives this packet, it will try to resynchronize the TCP session with node A by sending it an ACK packet with the sequence number that it is expecting. The cycle goes on and on, and the ACK packets passing back and forth create an ACK storm. Hijacking a session over UDP is the same as over TCP, except that UDP

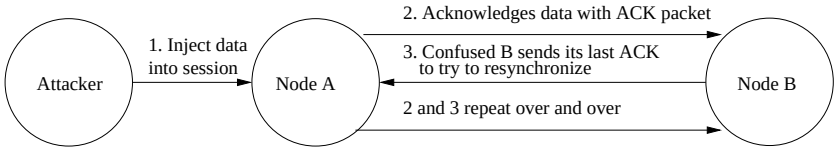

Figure 4: TCP ACK Storm

attackers do not have to worry about the overhead of managing sequence numbers and other TCP mechanisms. Since UDP is connectionless, edging into a session without being detected is much easier than the TCP session attacks.

3.5 Application layer attacks

Application layer attacks can be mobile viruses, worm attacks, and repudiation attacks.

- **Mobile virus and worm attacks:** The application layer contains user data, and it normally supports many protocols such as HTTP, SMTP, FTP. Malicious code, which includes viruses and worms, is applicable across operating systems and applications.

  As we know, malicious programs are widely spread in networks. There are a number of techniques by which a worm can discover new machines to exploit. One example is IP address scanning used by Internet worms. That technique consists of generating probe packets to a vulnerable UDP/TCP port at many different IP addresses. Hosts that are hit by the scan respond, receive a copy of the worm, and hence get infected. The Code Red worm [50] is one of the scanning worms.

  Some worms use a loophole of the system. For example, Worm.Blaster and Worm.Sasser [50] each use a different loophole: Worm.Blaster uses a system RPC DCOM loophole, and Worm.Sasser uses the system LSASS (local security authentication subsystem service). In MANET, an attacker can also produce a worm attack using any loophole of the system of the mobile ad hoc network.

- **Repudiation attack:** In the network layer, firewalls can be installed to keep packets in or keep packets out. In the transport layer, entire connections can be encrypted, end-to-end. But these solutions do not solve the authentication or non-repudiation problems in general. Repudiation refers to a denial of

participation in all or part of the communications. For example, a selfish person could deny conducting an operation on a credit card purchase, or deny any on-line bank transaction, which is the prototypical repudiation attack on a commercial system.

3.6 Multi-layer attacks

Some security attacks can be launched from multiple layers instead of a particular layer. Examples of multi-layer attacks are denial of service (DoS), man-in-the-middle, and impersonation attacks.

- **Denial of service:** Denial of service (DoS) attacks could be launched from several layers. An attacker can employ signal jamming at the physical layer, which disrupts normal communications. At the link layer, malicious nodes can occupy channels through the capture effect, which takes advantage of the binary exponential scheme in MAC protocols and prevents other nodes from channel access. At the network layer, the routing process can be interrupted through routing control packet modification, selective dropping, table overflow, or poisoning. At the transport and application layers, SYN flooding, session hijacking, and malicious programs can cause DoS attacks.

- **Impersonation attacks:** Impersonation attacks are just the first step for most attacks, and are used to launch further sophisticated attacks. For example, a malicious node can precede an attack by altering its MAC or IP address.

- **Man-in-the-middle attacks:** An attacker sits between the sender and the receiver and sniffs any information being sent between two ends. In some cases the attacker may impersonate the sender to communicate with the receiver, or impersonate the receiver to reply to the sender.

3.7 Cryptographic primitive attacks

Cryptography is an important and powerful security tool. It provides security services, such as authentication, confidentiality, integrity, and non-repudiation. In all likelihood, there exist attacks on many cryptographic primitives that have not yet been discovered. There could be new attacks designed and developed for hash functions, digital signatures, both block and stream ciphers, and fingerprinting schemes. Most security leaks are not because of the weakness of cryptographic algorithms, but because of the design of security protocols. For example, authentication protocols and cryptography key management protocols are more often the targets of malicious attacks. Some examples of cryptographic primitive attacks are

pseudorandom number attacks [51], digital signature attacks [14], and hash collision attacks [46].

3.7.1 Pseudorandom number attack

To make packets fresh, a timestamp or random number (nonce) can be used to prevent a replay attack [51]. The session key can be generated from a random number. In the public key infrastructure the shared secret key can be generated from a random number too. The conventional random number generators in most programming languages are designed for statistical randomness, not to resist prediction by cryptanalysts. In the optimal case, random numbers are generated based on physical sources of randomness that cannot be predicted. The noise from an electronic device or the position of a pointer device is a source of such randomness. However, true random numbers are difficult to generate. When true physical randomness is not available, pseudorandom numbers must be used. Cryptographic pseudorandom generators typically have a large pool (seed value) containing randomness. New environmental noise should be mixed into the pool to prevent others from determining previous or future values. The design and implementation of cryptographic pseudorandom generators could easily become the weakest point of the system.

3.7.2 Digital signature attacks

The RSA public key algorithm can be used to generate a digital signature. The signature scheme has one problem: it could suffer the blind signature attack. The user can get the signature of a message and use the signature and the message to fake another message's signature. The ElGamal signature is based on the difficulty in breaking the discrete log problem. Digital Signature Arithmetic (DSA) is an updated version of the ElGamal digital signature scheme published in 1994 by FIPS, and was chosen as the digital signature standard (DSS) [14]. The attack models for digital signature can be classified into known-message, chosen-message, and key-only attacks. In the known-message attack, the attacker knows a list of messages previously signed by the victim. In the chosen-message attack, the attacker can choose a specific message that it wants the victim to sign. But in the key-only attack, the adversary only knows the verification algorithm, which is public. Very often the digital signature algorithm is used in combination with a hash function. The hash function needs to be collision resistant.

3.7.3 Hash collision attacks

A collision attack is to find two messages with the same hash, but the attacker cannot pick what the hash will be. Collision attacks were announced in SHA-0, MD4,

MD5, HAVAL-128, and RIPEMD. The collisions against MD4, MD5, HAVAL-128, and RIPEMD were found by the Chinese researcher Wang with other co-authors. In February 2005, an attack against SHA-1 was reported by Wang. Wang found the collisions in SHA-1 with an estimated effort of 269 hash computations [46].

Normally all major digital signature techniques (including DSA and RSA) involve first hashing the data and then signing the hash value. The original message data is not signed directly by the digital signature algorithm for both performance and security reasons. Collision attacks could be used to tamper with existing certificates. An adversary might be able to construct a valid certificate corresponding to the hash collision.

3.7.4 Key management vulnerability

Key management protocols deal with the key generation, storage, distribution, updating, revocation, and certificate service. Attackers could launch attacks to disclose the cryptographic key at the local host or during the key distribution procedure. The lack of a central trusted entity in MANET makes it more vulnerable to key management attacks [5] [7] [9] [24]. For example, the man-in-the-middle attack is a design pitfall of the Diffie-Hellman (DH) key exchange protocol. For key management protocols that rely on a trusted key distribution center or certificate authority, the trusted central entity becomes the focus of attacks.

4 Security attacks countermeasures

Security is an essential service for wired and wireless network communications. The success of MANET strongly depends on whether its security can be trusted. However, the characteristics of MANET pose both challenges and opportunities in achieving the security goals, such as confidentiality, authentication, integrity, availability, access control, and non-repudiation.

The mobile hosts forming a MANET are normally mobile devices with limited physical protection and resources. Security modules, such as tokens and smart cards, can be used to protect against physical attacks. Cryptographic tools are widely used to provide powerful security services, such as confidentiality, authentication, integrity, and non-repudiation. Unfortunately, cryptography cannot guarantee availability; for example, it cannot prevent radio jamming. Meanwhile, strong cryptography often demands a heavy computation overhead and requires the auxiliary complicated key distribution and trust management services, which mostly are restricted by the capabilities of physical devices (e.g. CPU or battery).

The characteristics and nature of MANET require the strict cooperation of participating mobile hosts. A number of security techniques have been invented and a list of security protocols has been proposed to enforce cooperation and prevent misbehavior, such as 802.11 WEP, IPsec, SEAD, SAODV, SRP, ARAN, SSL, and so on. However, none of those preventive approaches is perfect or capable to defend against all attacks. A second line of defense called intrusion detection systems (IDS) is proposed and applied in MANET. IDS are some of the latest security tools in the battle against attacks. Distributed IDS were introduced in MANET to monitor either the misbehavior or selfishness of mobile hosts. Subsequent actions can be taken based on the information collected by IDS.

The attacks countermeasures presentation is as follows. An overview of security attributes and security mechanisms is presented in Sections 4.1 and 4.2, respectively. We describe the attack countermeasures by different network layers. Physical layer defense is discussed in Section 4.3, link layer defense is discussed in Section 4.4, and network layer defense is discussed in Section 4.5. Transport layer defense and application layer defense are discussed in Section 4.6 and Section 4.7 respectively. Multi-layer defense is in Section 4.8. Defense against key management attacks is in Section 4.9, and MANET intrusion detection systems are discussed in 4.10.

4.1 Security attributes

Security is the combination of processes, procedures, and systems used to ensure confidentiality, authentication, integrity, availability, access control, and non-repudiation.

- **Confidentiality** is to keep the information sent unreadable to unauthorized users or nodes. MANET uses an open medium, so usually all nodes within the direct transmission range can obtain the data. One way to keep information confidential is to encrypt the data, and another technique is to use directional antennas.

- **Authentication** is to be able to identify a node or a user, and to be able to prevent impersonation. In wired networks and infrastructure-based wireless networks, it is possible to implement a central authority at a point such as a router, base station, or access point. But there is no central authority in MANET, and it is much more difficult to authenticate an entity.

- **Integrity** is to be able to keep the message sent from being illegally altered or destroyed in the transmission. When the data is sent through the wireless

medium, the data can be modified or deleted by malicious attackers. The malicious attackers can also resend it, which is called a replay attack.

- **Non-repudiation** is related to a fact that if an entity sends a message, the entity cannot deny that the message was sent by it. By producing a signature for the message, the entity cannot later deny the message. In public key cryptography, a node A signs the message using its private key. All other nodes can verify the signed message by using A's public key, and A cannot deny that its signature is attached to the message.

- **Availability** is to keep the network service or resources available to legitimate users. It ensures the survivability of the network despite malicious incidents.

- **Access control** is to prevent unauthorized use of network services and system resources. Obviously, access control is tied to authentication attributes. In general, access control is the most commonly thought of service in both network communications and individual computer systems.

4.2 Security mechanisms

A variety of security mechanisms have been invented to counter malicious attacks. The conventional approaches such as authentication, access control, encryption, and digital signature provide a first line of defense. As a second line of defense, intrusion detection systems and cooperation enforcement mechanisms implemented in MANET can also help to defend against attacks or enforce cooperation, reducing selfish node behavior.

- **Preventive mechanism:** The conventional authentication and encryption schemes are based on cryptography, which includes asymmetric and symmetric cryptography. Cryptographic primitives such as hash functions (message digests) can be used to enhance data integrity in transmission as well. Threshold cryptography can be used to hide data by dividing it into a number of shares. Digital signatures can be used to achieve data integrity and authentication services as well.

  It is also necessary to consider the physical safety of mobile devices, since the hosts are normally small devices, which are physically vulnerable. For example, a device could easily be stolen, lost, or damaged. In the battlefield they are at risk of being hijacked. The protection of the sensitive data on a physical device can be enforced by some security modules, such as tokens or a smart card that is accessible through PIN, passphrases, or biometrics.

Although all of these cryptographic primitives combined can prevent most attacks in theory, in reality, due to the design, implementation, or selection of protocols and physical device restrictions, there are still a number of malicious attacks bypassing prevention mechanisms.

- **Reactive mechanism:** An intrusion detection system is a second line of defense. There are widely used to detect misuse and anomalies. A misuse detection system attempts to define improper behavior based on the patterns of well-known attacks, but it lacks the ability to detect any attacks that were not considered during the creation of the patterns; Anomaly detection attempts to define normal or expected behavior statistically. It collects data from legitimate user behavior over a period of time, and then statistical tests are applied to determine anomalous behavior with a high level of confidence. In practice, both approaches can be combined to be more effective against attacks. Some intrusion detection systems for MANET have been proposed in recent research papers.

4.3 Physical layer defense

Wireless communication is broadcast by nature. A common radio signal is easy to jam or intercept. Spread spectrum technology, such as frequency hopping (FHSS) [27] or direct sequence (DSSS) [27], can make it difficult to detect or jam signals. It changes frequency in a random fashion to make signal capture difficult or spreads the energy to a wider spectrum so the transmission power is hidden behind the noise level. Directional antennas can also be deployed due to the fact that the communication techniques can be designed to spread the signal energy in space.

- **FHSS:** The signal is modulated with a seemingly random series of radio frequencies, which hops from frequency to frequency at fixed intervals. The receiver uses the same spreading code, which is synchronized with the transmitter, to recombine the spread signals into their original form. Figure 5 shows an example of a frequency-hopping signal.

  With the transmitter and the receiver synchronized properly, data is transmitted over a single channel. However, the signal appears to be unintelligible duration impulse noise for the eavesdroppers. Meanwhile, interference is minimized as the signal is spread across multiple frequencies.

- **DSSS:** Each data bit in the original signal is represented by multiple bits in the transmitted signal, using a spreading code. The spreading code spreads the signal across a wider frequency band in direct proportion to the number of bits used. The receiver can use the spreading code with the signal to

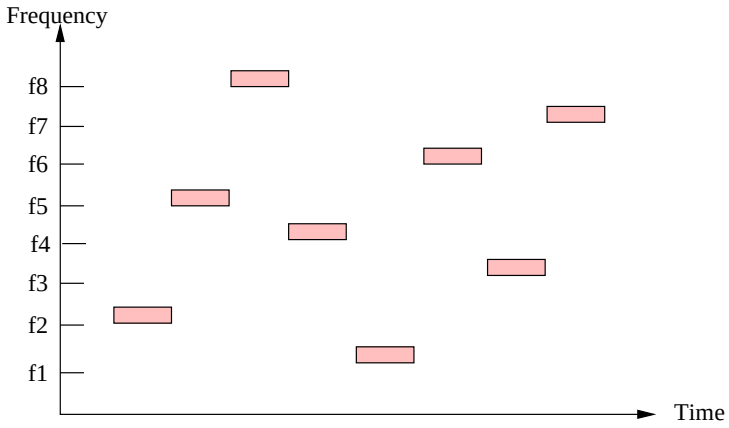

Figure 5: Illustration of Frequency Hopping Spread Spectrum

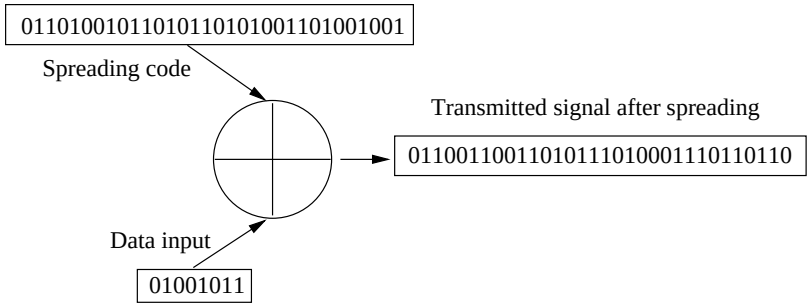

Figure 6: Illustration of Direct Sequence Spread Spectrum

recover the original data. Figure 6 illustrates that each original bit of data is represented by 4 bits in the transmitted signal. The first bit of data, a 0 is transmitted as 0110 which is first 4 bits of spreading code. The second bit, 1, is transmitted as 0110 which is bit-wise complement of the second 4 bits of spreading code. In turn, each input bit is combined, using exclusive-or, with four bits of the spreading code.

Both FHSS and DSSS pose difficulties for outsiders attempting to intercept the radio signals. The eavesdropper must know the frequency band, spreading code, and modulation techniques in order to accurately read the transmitted signals. The property that spread spectrum technologies do not interoperate with each other further adds difficulties to the eavesdropper. Spread spectrum technology also minimizes the potential for interference from other radios and electromagnetic devices. Despite the capability of spread spectrum technology, it is secure only when the

hopping pattern or spreading code is unknown to the eavesdroppers.

4.4 Link layer defense

There are malicious attacks that target the link layer by disrupting the cooperative nature of link layer protocols. Link layer protocols help to discover 1-hop neighbors, handle fair channel access, frame error control, and maintain neighbor connections. Selfish nodes could disobey the channel access rule, manipulate the NAV field, cheat backoff values, and so on in order to maximize their own throughput. Neighbors should monitor these misbehaviors. Although it is still an open challenge to prevent selfishness, some schemes are proposed, such as ERA-802.11 [12], where detection algorithms are proposed. Traffic analysis is prevented by encryption at data link layer.

The wired equivalent privacy (WEP) encryption scheme defined in the IEEE 802.11 wireless LAN standard uses link encryption to hide the end-to-end traffic flow information. However, WEP has been widely criticized for its weaknesses [47]. Some secure link layer protocols have been proposed in recent research, such as LLSP.

In MANET, some papers propose to create a security cloud, construct a traffic cover mode or dynamic mix method, or use traditional traffic padding and traffic rerouting techniques to prevent traffic analysis. A security cloud means that each node under the security cloud is identical in terms of traffic generation. A traffic cover mode hides the changes of an end-to-end flow traffic pattern, because certain tactical information might be inferred from the unusual changes in the traffic pattern. A dynamic mix method is used to hide the source and destination information during message delivery via a cryptography method and to ”mix” nodes in the network.

4.5 Network layer defense

The passive attack on routing information can be countered with the same methods that protect data traffic. Some active attacks, such as illegal modification of routing messages, can be prevented by mechanisms source authentication and message integrity. DoS attacks on a routing protocol could take many forms. DoS attacks can be limited by preventing the attacker from inserting routing loops, enforcing the maximum route length that a packet should travel, or using some other active approaches. The wormhole attack can be detected by an unalterable and independent physical metric, such as time delay or geographical location. For example, packet leashes are used to combat wormhole attacks [15].

In general, some kind of authentication and integrity mechanism, either the hop-by-hop or the end-to-end approach, is used to ensure the correctness of routing information. For instance, digital signature, one-way hash function, hash chain, message authentication code (MAC), and hashed message authentication code (HMAC) are widely used for this purpose. IPsec and ESP are standards of security protocols on the network layer used in the Internet that could also be used in MANET, in certain circumstances, to provide network layer data packet authentication, and a certain level of confidentiality; in addition, some protocols are designed to defend against selfish nodes, which intend to save resources and avoid network cooperation. Some secure routing protocols have been proposed in MANET in recent papers. We outline those defense techniques at below sections.

Section 4.5.1 describes the proposed defense against wormhole attacks. Section 4.5.2 outlines the defense against blackhole attacks. Section 4.5.3 presents the defense against impersonation and repudiation attacks. Section 4.5.4 talks about the defense against modification attacks.

4.5.1 Defense against wormhole attacks

A packet leash protocol [15] is designed as a countermeasure to the wormhole attack. The SECTOR mechanism [52] is proposed to detect wormholes without the need of clock synchronization. Directional antennas [42] are also proposed to prevent wormhole attacks.

In the wormhole attack, an attacker receives packets at one point in the network, tunnels them to another point in the network, and then replays them into the network from that point. To defend against wormhole attacks, some efforts have been put into hardware design and signal processing techniques. If data bits are transferred in some special modulating method known only to the neighbor nodes, they are resistant to closed wormholes. Another potential solution is to integrate the prevention methods into intrusion detection systems. However, it is difficult to isolate the attacker with a software-only approach, since the packets sent by the wormhole are identical to the packets sent by legitimate nodes.

Packet leashes [15] are proposed to detect wormhole attacks. A leash is the information added into a packet to restrict its transmission distance. A temporal packet leash sets a bound on the lifetime of a packet, which adds a constraint to its travel distance. A sender includes the transmission time and location in the message. The receiver checks whether the packet has traveled the distance between the sender and itself within the time frame between its reception and transmission. Temporal packet leashes require tightly synchronized clocks and precise location knowledge. In geographical leashes, location information and loosely synchronized clocks together verify the neighbor relation.

The SECTOR [52] mechanism is based primarily on distance-bounding techniques, one-way hash chains, and the Merkle hash tree. SECTOR can be used to prevent wormhole attacks in MANET without requiring any clock synchronization or location information. SECTOR can also be used to help secure routing protocols in MANET using last encounters, and to help detect cheating by means of topology tracking.

Directional antennas [42] are also proposed as a countermeasure against wormhole attacks. This approach does not require either location information or clock synchronization, and is more efficient with energy.

4.5.2 Defense against blackhole attacks

Some secure routing protocols, such as the security-aware ad hoc routing protocol (SAR) [54], can be used to defend against blackhole attacks. The security-aware ad hoc routing protocol is based on on-demand protocols, such as AODV or DSR. In SAR, a security metric is added into the RREQ packet, and a different route discovery procedure is used. Intermediate nodes receive an RREQ packet with a particular security metric or trust level. At intermediate nodes, if the security metric or trust level is satisfied, the node will process the RREQ packet, and it will propagate to its neighbors using controlled flooding. Otherwise, the RREQ is dropped. If an end-to-end path with the required security attributes can be found, the destination will generate a RREP packet with the specific security metric. If the destination node fails to find a route with the required security metric or trust level, it sends a notification to the sender and allows the sender to adjust the security level in order to find a route.

To implement SAR, it is necessary to bind the identity of a user with an associated trust level. To prevent identity theft, stronger access control mechanisms such as authentication and authorization are required. In SAR, a simple shared secret is used to generate a symmetric encryption/decryption key per trust level. Packets are encrypted using the key associated with the trust level; nodes belonging to different levels cannot read the RREQ or RREP packets. It is assumed that an outsider cannot obtain the key.

In SAR, a malicious node that interrupts the flow of packets by altering the security metric to a higher or lower level cannot cause serious damage because the legitimate intermediate or destination node is supposed to drop the packet, and the attacker is not able to decrypt the packet. SAR provides a suite of cryptographic techniques, such as digital signature and encryption, which can be incorporated on a need-to-use basis to prevent modification.

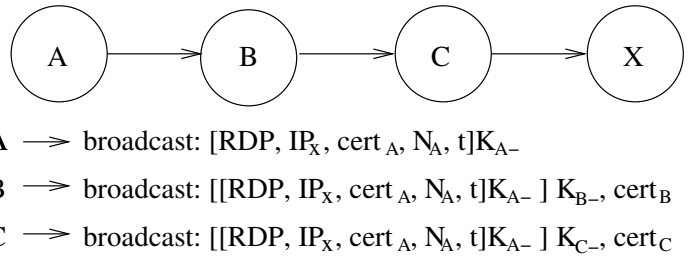

Figure 7: Illustration of ARAN Routing Discovery Authentication at Each Hop

4.5.3 Defense against impersonation and repudiation attack

ARAN [32] can be used to defend against impersonation and repudiation attacks. ARAN provides authentication and non-repudiation services using predetermined cryptographic certificates for end-to-end authentication. In ARAN, each node requests a certificate from a trusted certificate server. Route discovery is accomplished by broadcasting a route discovery message RDP from the source node. The reply message REP is unicast from the destination to the source. The routing messages are authenticated at each intermediate hop in both directions.

Routing discovery authentication at each hop is illustrated in Figure 7. The RDP packet includes [RDP, IP_X, $Cert_A$, N_A, t]K_{A-} , where RDP is a packet identifier, A is the source node, IP_X is the destination node X's IP address, N_A is a nonce, $Cert_A$ is A's certificate, t is the current time, and K_{A-} after the packet $RDP, IP_X, Cert_A, N_A, t$ means the packet was signed with A's private key. If the intermediate node B is the first hop from node A, after validating A's signature and checking its certificate for expiration, it will decide to sign the packet by adding its own signature and certificate, and then it will forward [[RDP, IP_x, $Cert_A$, N_A, t]K_{A-}]K_{B-}, Cert$_B$ to all its neighbors. Each hop verifies the signature of the previous hop and replaces it with its own. The destination node X unicasts a REP packet [REP, IP_A, $Cert_X$, N_A, t]K_{X-} back to source A.

Because RDPs do not contain a hop count or specific recorded source route, and because messages are signed at each hop, malicious nodes have no chance to form a routing loop by redirecting traffic or using impersonation to instantiate routes. The disadvantage of ARAN is that it uses hop-by-hop authentication, which incurs a large computation overhead. Meanwhile, each node needs to maintain one table entry per source-destination pair that is currently active.

Table 4: SEAD Example: Hash Function used for Message Authentication, i is sequence number, j is metric, the network diameter (m) is 5, the length of hash chain (n) is 20

	j=0	1	2	3	4
i=1	h_{15}	h_{16}	h_{17}	h_{18}	h_{19}
2	h_{10}	h_{11}	h_{12}	h_{13}	h_{14}
3	h_5	h_6	h_7	h_8	h_9
4	h_0	h_1	h_2	h_3	h_4

4.5.4 Defense against modification attacks

The security protocol SEAD [11] is used here as an example of a defense against modification attacks. Similar to a packet leash [15], the SEAD protocol utilizes a one-way hash chain to prevent malicious nodes from increasing the sequence number or decreasing the hop count in routing advertisement packets. In SEAD, nodes need to authenticate neighbors by using TESLA [12] broadcast authentication or a symmetric cryptographic mechanism. Specifically, in SEAD, a node generates a hash chain and organizes the chain into segments of m elements as (h_0, h_1, ..., h_{m-1}), ..., (h_{km}, h_{km+1}, ..., h_{km+m-1}), ..., h_n, where k = $\frac{n}{m}$ - i, m is the maximum network diameter, and i is the sequence number.

Illustrated in table 4, the network diameter is 5, the length of hash chain n's value is 20, i is the sequence number, and j is the metric, which is number of hops to destination. Because h_i=H(h_{i-1}), given h_i it is easy to verify the authenticity of h_j, as long as $j<i$. Given h_i, h_j cannot be derived for $j<i$, but h_j can be derived for $j>i$. Because different hash function is used for different i and j and used by the order showed in table 4, the attacker can never forge lower metric value, or greater sequence value. Because, after receiving a routing update in routing protocol DSDV, a node updates its advertised routing table when the sequence number is greater or when the sequence number is the same but the metric is lower, SEAD prevents malicious nodes from decreasing the hop count value or increasing the sequence number based on the design of DSDV.

4.6 Transport layer defense

In MANET, like TCP protocols in the Internet, nodes are vulnerable to the classic SYN flooding attack, or session hijacking attack.

Point-to-point or end-to-end encryption provides message confidentiality at or above the transport layer in two end systems. TCP is a connection-oriented reliable transport layer protocol. Because TCP does not perform well in MANET, TCP

feedback (TCP-F) [49], TCP explicit failure notification (TCP-ELFN) [49], ad hoc transmission control protocol (ATCP) [49], and ad hoc transport protocol (ATP) [49] have been invented, but none of these protocols are designed with security in mind.

Secure Socket Layer (SSL) [51], Transport Layer Security (TLS) [51], and Private Communications Transport (PCT) [51] protocols were designed for secure communications and are based on public key cryptography. TLS/SSL can help secure data transmission. It can also help to protect against masquerade attacks, man-in-the-middle (or bucket brigade) attacks, rollback attacks, and replay attacks. TLS/SSL is based on public key cryptography, which is CPU-intensive and requires comprehensive administrative configuration. Therefore, the application of these schemes in MANET is restricted. TLS/SSL has to be modified in order to address the special needs of MANET. Some firewall at a higher level can be configured to defend against SYN flooding attacks.

4.7 Application layer defense

Like the other protocol layers, the application layer also needs to be secured. In a network with a firewall installed, the firewall can provide access control, user authentication, packet filtering, and a logging and accounting service. Application layer firewalls can effectively prevent many attacks, and application-specific modules, for example, spyware detection software, have also been developed to guard mission-critical services. However, a firewall is mostly restricted to basic access control and is not able to solve all security problems. For example, it is not effective against attacks from insiders. Because of MANET's lack of infrastructure, a firewall is not particularly useful.

In MANET, an Intrusion Detection System (IDS) can be used as a second line of defense. Intrusion detection can be installed at the network layer, but in the application layer it is not only feasible, but also necessary. Certain attacks, such as an attack that tries to gain unauthorized access to a service, may seem legitimate to the lower layers, such as the MAC protocols. Also some attacks may be more obvious in the application layer. For instance, the application layer can detect a DoS attack more quickly than the lower layers when a large number of incoming service connections have no actual operations, since low layers need more time to recognize it.

4.8 Defense against multi-layer attacks

The DoS attacks, impersonation attacks, man-in-the-middle attacks, and many other attacks can target multiple layers. The countermeasures for these attacks

need to be implemented at different layers. For example, directional antennas [52] are used at the media access layer to defend against wormhole attacks, and packet leashes [15] are used as a network layer defense against wormhole attacks. The countermeasures for multi-layer attacks can also be implemented in an integrated scheme. For example, if a node detects a local intrusion at a higher layer, lower layers are notified to do further investigation.

As an example, we give a detailed description about the defense against DoS attacks.

- **Defense against DoS attacks:** In MANET, two types of DoS attacks [55] are quite common. One is at the routing layer, and another is at the MAC layer. Attacks at the routing layer could consist of but is not limited to the following misbehaviors:

  1. The malicious node participates in a route but simply drops some of the data packets.
  2. The malicious node transmits falsified route updates.
  3. The malicious node could potentially replay stale updates.
  4. The malicious node reduces the TTL (time-to-live) field in the IP header so that the packet never reaches the destination.

  If end-to-end authentication is enforced, attacks by independent malicious node of types (2) and (3) may be thwarted. An attack of type (1) may be handled by assigning confidence levels to nodes and using routes that provide the highest level of confidence. An attack of type (4) may be countered by making it mandatory that a relay node ensures that the TTL field is set to a value greater than the hop count to the intended destinations.

  If nodes collude, the authentication mechanisms fail and it is an open problem to provide protection against such routing attacks.

  At the MAC layer DoS attacks could include, among others, the following misbehaviors:

  1. Keeping the channel busy in the vicinity of a node leads to a denial of service attack at that node.
  2. By using a particular node to continually relay spurious data, the battery life of that node may be drained.

  End-to-end authentication may prevent the above two cases from succeeding. If the node does not have a certificate of authentication, it may be prevented

from accessing the channel. Usually the nodes are outsiders. However, if nodes collude, and the colluding nodes include the sending node and the destination, MAC layer attacks are very feasible.

4.9 Defense against key management attacks

Cryptography algorithms are security primitives, which are widely used for the purposes of authentication, confidentiality, integrity, and non-repudiation. Most cryptographic systems rely on the underlining secure, robust, and efficient key management system. Key management is in the central part of any secure communication, and is the weak point of system security and protocol design. A key is a piece of input information for cryptography algorithms. If the key were released, the encrypted information would be disclosed. The secrecy of the symmetric key and private key must be assured locally. The Key Encryption Key (KEK) approach could be used at local hosts to build a line of defense.

Key distribution and key agreement over an insecure channel are at high risk and suffer from potential attacks. In the traditional digital envelop approach, a session key is generated at one side and is encrypted by the public-key algorithm. Then it is delivered and recovered at the other end. In the Diffie-Hellman (DH) scheme, the communication parties at both sides exchange some public information and generate a session key on both ends. Several enhanced DH schemes have been invented to counter man-in-the-middle attacks. In the symmetric approach, the sequence number or a nonce could be included to prevent the replay attack on setting up a session key. In addition, a multi-way challenge response protocol, such as Needham-Schroeder, can also be used. Kerberos, which is based on a variant of Needham-Schroeder, is an authentication protocol used in many real systems including Windows.

Key integrity and ownership should be protected from advanced key attacks. Digital signature, message digest, and hashed message authentication code (HMAC) are techniques used for data authentication or integrity purposes. Similarly, public key is protected by the public-key certificate, in which a trusted entity called the certification authority (CA) in PKI vouches for the binding of the public key with the owner's identity. In systems lacking a trusted third party (TTP), the public-key certificate is vouched for by peer nodes in a distributed manner, such as pretty good privacy (PGP). In some distributed approaches, the system secret is distributed to a subset or all of the network hosts based on threshold cryptography. Obviously, a certificate cannot prove whether an entity is "good" or "bad", but can prove ownership of a key. Mainly it is for key authentication.

A cryptographic key could be compromised or disclosed after a certain period of usage. Since the key should no longer be useable after its disclosure, some

mechanism is required to enforce this rule. In PKI, this can be done implicitly or explicitly. The certificate contains the lifetime of validity-it is not useful after expiration. But in some cases, the private key could be disclosed during the valid period, in which case certification authority (CA) needs to revoke a certificate explicitly and notify the network by posting it onto the certificate revocation list (CRL) to prevent its usage.

Currently there are three types of key management on MANET: the first one is virtual CA approach [3], the second one is certificate chaining [57], and the third one is composite key management, which combines the first two [9].

4.10 MANET intrusion detection systems (IDS)

Because MANET has features such as an open medium, dynamic changing topology, and the lack of a centralized monitoring and management point, many of the intrusion detection techniques developed for a fixed wired network are not applicable in MANET. Zhang [37] gives a specific design of intrusion detection and response mechanisms for MANET. Marti [36] proposes two mechanisms: watchdog and pathrater, which improve throughput in MANET in the presence of nodes that agree to forward packets but fail to do so. In MANET, cooperation is very important to support the basic functions of the network so the token-based mechanism, the credit-based mechanism, and the reputation-based mechanism were developed to enforce cooperation. Each mechanism is discussed in this paper.

4.10.1 MANET IDS agent conceptual architecture

The basic approach in MANET [36] is that each mobile node runs an IDS agent independently. It has to observe the behavior of neighboring nodes, detect local intrusion, cooperate with neighboring nodes, and, if needed, make decisions and take actions. An IDS agent has data collection, a local detection engine, local response, a cooperative detection engine, global response, and secure communication with neighboring IDS agents. Figure 8 is a conceptual model of an IDS agent.

4.10.2 Approaches to detect routing misbehavior

Watchdog and pathrater [36] are proposed for the DSR routing protocol. It is assumed that wireless links are bi-directional; wireless interfaces support promiscuous mode operation, which means that if a node A is within the transmission range of B, it can overhear communications to and from B even if those communications do not directly involve A.

The watchdog methods detect misbehaving nodes. A node may measure a neighboring node's frequency of dropping or misrouting packets, or its frequency

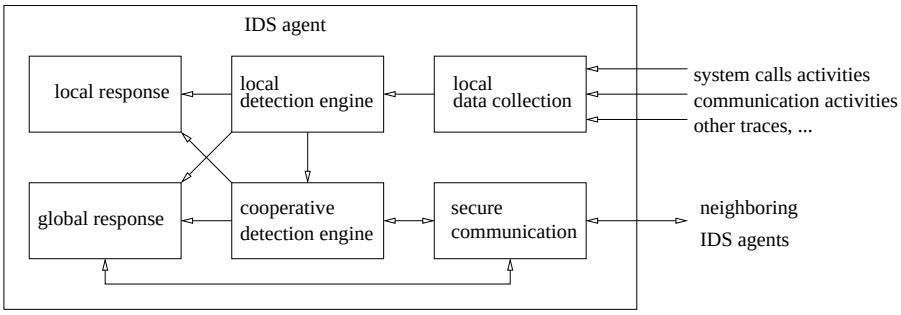

Figure 8: A Conceptual Model for an IDS Agent in MANET

of invalid routing information advertisements. The implementation of a watchdog is maintains a buffer of recently sent packets and compares each overheard packet with the packets in the buffer to see if there is a match. If there is a match, the node removes the packet from the buffer; otherwise if a packet has remained in the buffer for longer than a certain timeout, the watchdog increments a failure tally for the neighboring node. If the tally exceeds a certain threshold bandwidth, it sends a message to the source notifying it of the misbehaving node. The weaknesses of watchdog are that it might not detect a misbehaving node because of ambiguous collisions, receiver collisions, limited transmission power, false behavior, collusion, and partial dropping.

In another scheme, pathrater is run by each node. Each node keeps track of the trustworthiness rating of every known node, including calculating path metrics by averaging the node ratings in the path to each known node. If there are multiple paths to the same destination, then according to standard DSR routing protocol the shortest path in the route cache is chosen, but when using pathrater the path with the highest metric is chosen.

4.10.3 Cooperation enforcement

Generally, there are two kinds of misbehaving nodes: one is the selfish node, and the other is the malicious node. Selfish nodes don't cooperate for selfish reasons, such as saving power. Even though the selfish nodes do not intend to damage other nodes, the main threat from selfish nodes is the dropping of packets, which may affect the performance of the network severely. Malicious nodes have the intention to damage other nodes, and battery saving is not a priority. Without any incentive for cooperating, network performance can be severely degraded. The mechanisms to enforce cooperating are currently split into three research areas: token-based, micro-payment, and reputation-based. Yang [58] proposed a token-based scheme.

Buttyan [59] proposed the nuglets scheme. The nuglets scheme is micro-payment scheme. Buchegger's CONFIDANT [41], Michiard's CORE [60], and S.Bansel's OCEAN [61] are reputation-based schemes.

- **Token-based mechanism:** The token-based scheme [58] is a unified network-layer security solution in MANET based on the AODV protocol. In this scheme, each node carries a token in order to participate network operations, and its local neighbors collaboratively monitor any misbehavior in routing or packet forwarding services. The approach is different from a watchdog, which monitors neighbors alone, not collaboratively.

  Nodes without a valid token are isolated in the network, and all of their legitimate neighbors will not interact with them in routing and forwarding services. Upon expiration of the token, each node renews its token via its neighbors. The lifetime of a token is related to the node's behavior. A well-behaving node with a good record needs to renew its token less often.

  This approach uses asymmetric cryptographic primitives such as RSA. There is a global secret key and public key pair. Each legitimate node carries a token stamped with an expiration time and marked with a signature. The design is based on several assumptions to simplify the mechanism:

  1. Any two nodes within wireless transmission range may monitor each other.
  2. The approach is only based on network-layer security, not physical-layer or link-layer issues.
  3. Only the secure route for data forwarding between the source and destination is discussed, not data packet confidentiality and integrity.
  4. Each node has a unique ID.
  5. Multiple attackers are possible, but there is a limit to attackers in any neighborhood.
  6. Every legitimate node has a token signed with the system secret key, which can be verified by its neighbors.

- **Credit-based mechanism:** The nuglets scheme [59] is an approach analogous to virtual currency. A node that consumes a service must pay the nodes that provide the service in nuglets. The combination of watchdog and pathrater cannot hold any misbehaving nodes accountable, and misbehaving nodes are still able to send and receive packets. However, in the nuglets scheme, a misbehaving node will be locked out by its neighbors. That is much better in fairness.

Nuglets are designed to simulate packet forwarding. The nuglets are related to the counters in the nodes. The counter is maintained by a trusted and tamper-resistant hardware module at each node. A packet purse holds nuglets, which are contained in the packet. The packet purse is protected from unauthorized modification and detachment from the original packet by cryptographic mechanisms. The packet forward protocol is designed on fixed per hop charges.

- **Reputation-based mechanism:** CONFIDANT [41] presents an extension to the routing protocol in order to detect and isolate misbehaving nodes. The protocol is designed to be able to make cooperation fair. With CONFIDANT, each node has four components: a monitor, a reputation system, a trust manager, and a path manager.

  The CONFIDANT approach copes with MANET security, robustness, and fairness by retaliating for malicious behavior and warning affiliated nodes to avoid bad experiences. Nodes learn not only from their own experience, but also from observing the neighborhood and from the experience of their friends.

5 Open challenges and future directions

Security is such an important feature that it could determine the success and wide deployment of MANET. A variety of attacks have been identified. Security countermeasures either currently used in wired or wireless networking or newly designed specifically for MANET are presented in the above sections. Security must be ensured in the entire system including the security primitives, such as key management protocols, since overall security level is determined by the system's weakest point.

The research on MANET is still in an early stage. Existing proposals are typically based on one specific attack. They could work well in the presence of designated attacks, but there are many unanticipated or combined attacks that remain undiscovered. A lot of research is still on the way to identify new threats and create secure mechanisms to counter those threats. More research can be done on the robust key management system, trust-based protocols, integrated approaches to routing security, and data security at different layers. Here are some research topics and future work in the area:

- Cryptography is the fundamental security technique used in almost all aspects of security. The strength of any cryptographic system depends on proper key management. The public-key cryptography approach relies on

the centralized CA entity, which is a security weak point in MANET. Some papers propose to distribute CA functionality to multiple or all network entities based on a secret sharing scheme, while some suggest a fully distributed trust model, in the style of PGP. Symmetric cryptography has computation efficiency, yet it suffers from potential attacks on key agreement or key distribution. For example, the Diffie-Hellman (DH) scheme is vulnerable to the man-in-the-middle attack. Many complicated key exchange or distribution protocols have been designed, but for MANET, they are restricted by a node's available resources, dynamic network topology, and limited bandwidth. Efficient key agreement and distribution in MANET is an ongoing research area.

- Most of the current work is on preventive methods with intrusion detection as the second line of defense. One interesting research issue is to build a trust-based system so that the level of security enforcement is dependant on the trust level. Building a sound trust-based system and integrating it into the current preventive methods can be done in future research.

- Since most attacks are unpredictable, a resiliency-oriented security solution will be more useful, which depends on a multi-fence security solution. Cryptography-based methods offer a subset of solutions. Other solutions will be in future research.

6 Acknowledgement

This work was supported in part by NSF grants CCR 0329741, CNS 0422762, CNS 0434533, ANI 0073736, EIA 0130806, and by a federal earmark project on Secure Telecommunication Networks.

References

[1] A. Salomaa, *Public-Key Cryptography*, Springer-Verlag, 1996.

[2] A. Tanenbaum, *Computer Networks*, PH PTR, 2003.

[3] L. Zhou and Z. Haas, Securing Ad Hoc Networks, *IEEE Network Magazine* Vol.13 No.6 (1999) pp. 24-30.

[4] S. Yi, P. Naldurg, and R. Kravets, Security Aware Ad hoc Routing for Wireless Networks. Report No.UIUCDCS-R-2002-2290, UIUC, 2002.

[5] H. Luo and S. Lu, URSA: Ubiquitous and Robust Access Control for Mobile Ad-Hoc Networks, *IEEE/ACM Transactions on Networking* Vol.12 No.6 (2004) pp. 1049-1063.

[6] W. Lou and Y. Fang, A Survey of Wireless Security in Mobile Ad Hoc Networks: Challenges and Available Solutions. *Ad Hoc Wireless Networks*, edited by X. Chen, X. Huang and D. Du. Kluwer Academic Publishers, pp. 319-364, 2003.

[7] S. Burnett and S. Paine, *RSA Security's Official Guide to Cryptography*, RSA Press, 2001.

[8] M. Ilyas, *The Handbook of Ad Hoc Wireless Networks*, CRC Press, 2003.

[9] S. Yi and R. Kravets, Composite Key Management for Ad Hoc Networks. *Proc. of the 1st Annual International Conference on Mobile and Ubiquitous Systems: Networking and Services* (MobiQuitous'04), pp. 52-61, 2004.

[10] M. Zapata, Secure Ad Hoc On-Demand Distance Vector (SAODV). Internet draft, draft-guerrero-manet-saodv-01.txt, 2002.

[11] Y. Hu, D. Johnson, and A. Perrig, SEAD: Secure Efficient Distance Vector Routing in Mobile Wireless Ad-Hoc Networks. *Proc. of the 4th IEEE Workshop on Mobile Computing Systems and Applications* (WMCSA'02), pp. 3-13, 2002.

[12] A. Perrig, R. Canetti, J. Tygar, and D. Song, The TESLA Broadcast Authentication Protocol. Internet Draft, 2000.

[13] P. Papadimitratos and Z. Haas, Secure Routing for Mobile Ad Hoc Networks. *Proc. of the SCS Communication Networks and Distributed Systems Modeling and Simulation Conference* (CNDS 2002), 2002.

[14] W. Mehuron, Digital Signature Standard (DSS). U.S. Department of Commerce, National Institute of Standards and Technology (NIST), Information Technology Laboratory (ITL). FIPS PEB 186, 1994.

[15] Y. Hu, A Perrig, and D. Johnson, Packet Leashes: A Defense Against Wormhole Attacks in Wireless Ad Hoc Networks. *Proc. of IEEE INFORCOM*, 2002.

[16] H. Deng, W. Li, and D. Agrawal, Routing Security in Wireless Ad Hoc Networks. *IEEE Communications Magazine*, vol. 40, no. 10, 2002.

[17] B. Awerbuch, D. Holmer, C. Nita-Rotaru, and H. Rubens, An On-demand Secure Routing Protocol Resilient to Byzantine Failures. *Proceedings of the ACM Workshop on Wireless Security*, pp. 21-30, 2002.

[18] P. Papadimitratos and Z. Haas, Secure Data Transmission in Mobile Ad Hoc Networks. *Proc. of the 2003 ACM Workshop on Wireless Security*, pp. 41-50, 2003.

[19] Y. Hu, A. Perrig, and D. Johnson, Rushing Attacks and Defense in Wireless Ad Hoc Network Routing Protocols. *Proc. of the ACM Workshop on Wireless Security* (WiSe), pp. 30-40, 2003.

[20] Y. Hu, A. Perrig, and D. Johnson, Ariadne: A Secure On-Demand Routing for Ad Hoc Networks. *Proc. of MobiCom 2002*, Atlanta, 2002.

[21] H. Yang, H. Luo, F. Ye, S. Lu, and L. Zhang, Security in Mobile Ad Hoc Networks: Challenges and Solutions. *IEEE Wireless Communications*, pp. 38-47, 2004.

[22] C. Perkins, *Ad Hoc Networks*, Addison-Wesley, 2001.

[23] R. Oppliger, *Internet and Intranet Security*, Artech House, 1998.

[24] B. Wu, J. Wu, E. Fernandez, S. Magliveras, and M. Ilyas, Secure and Efficient Key Management in Mobile Ad Hoc Networks. *Proc. of 19th IEEE International Parallel & Distributed Processing Symposium*, Denver, 2005.

[25] L. Buttyan and J. Hubaux, Report on Working Session on Security in Wireless Ad Hoc Networks. *Mobile Computing and Communications Review*, vol. 6, 2002.

[26] S. Ravi, A. Raghunathan, and N. Potlapally, Secure Wireless Data: System Architecture Challenges. *Proc. of International Conference on System Synthesis*, 2002.

[27] W. Stallings, *Wireless Communication and Networks*, Pearson Education, 2002.

[28] N. Borisov, I. Goldberg and D.Wagner, Interception Mobile Communications: The Insecurity of 802.11. *Conference of Mobile Computing and Networking*, 2001.

[29] P. Kyasanur and N. Vaidya, Detection and Handling of MAC Layer Misbehavior in Wireless Networks. *Proc. of the International Conference on Dependable Systems and Networks*, pp. 173-182, 2003.

[30] A. Crdenas, S. Radosavac, and J. Baras, Detection and Prevention of MAC layer Misbehavior in Ad Hoc Networks. *Proc. of the 2nd ACM Workshop on Security of Ad Hoc and Sensor Networks*, pp. 17-22, 2004.

[31] C. Murthy and B. Manoj, *Ad Hoc Wireless Networks: Architectures and Protocols*, Prentice Hall PTR, 2005.

[32] K. Sanzgiri, B. Dahill, B. Levine, C. Shields, and E. Belding-Royer, A Secure Routing Protocol for Ad Hoc Networks. *Proc. of IEEE International Conference on Network Protocols* (ICNP), pp. 78-87, 2002

[33] K. Ng and W. Seah, Routing Security and Data Confidentiality for Mobile Ad Hoc Networks. *Proc. of Vehicular Technology Conference*(VTC), Jeju, Korea, 2003.

[34] M. Jakobsson, S. Wetzel, and B. Yener, Stealth Attacks on Ad Hoc Wireless Networks. *Proc. of IEEE Vehicular Technology Conference* (VTC), 2003.

[35] Y. Hu and A. Perrig, A Survey of Secure Wireless Ad Hoc Routing. *IEEE Security & Privacy*, pp. 28-39, 2004.

[36] S. Marti, T. Giuli, K. Lai, and M. Baker, Mitigating Routing Misbehavior in Mobile Ad Hoc Networks, *Proc. of the Sixth Annual International Conference on Mobile Computing and Networking* (MOBICOM), Boston, 2000.

[37] Y. Zhang and W. Lee, Intrusion Detection in Wireless Ad-hoc Networks, *Proc. of the Sixth Annual International Conference on Mobile Computing and Networking* (MOBICOM), Boston, 2000.

[38] P. Kyasanur and N. Vaidya, Detection and Handling of MAC Layer Misbehavior in Wireless Networks, *Proc. of Dependable Computing and Communications Symposium (DCC) at the International Conference on Dependable Systems and Networks (DSN)*, 2003.

[39] A. Cardenas, N. Benammar, G. Papageorgiou, and J. Baras, Cross-Layered Security Analysis of Wireless Ad Hoc Networks, *Proc. of 24th Army Science Conference*, 2004.

[40] H. Yang, X. Meng, and S. Lu, Self-Organized Network-Layer Security in Mobile Ad Hoc Networks. *Proc. of ACM MOBICOM Wireless Security Workshop* (WiSe'02), Atlanta, 2002.

[41] S. Buchegger and J. Boudec, Nodes Bearing Grudges: Towards Routing Security, Fairness, and Robustness in Mobile Ad Hoc Networks, *Proc. of the 10th*

Euromicro Workshop on Parallel, Distributed and Network-based Processing, Canary Islands, Spain, 2002.

[42] L. Hu and D. Evans, Using Directional Antennas to Prevent Wormhole Attacks. *Proc. of Networks and Distributed System Security Symposium* (NDSS), 2004.

[43] P. Ning and K. Sun, How to Misuse AODV: A Case Study of Inside Attacks against Mobile Ad-Hoc Routing Protocols, *Proceedings of the 2003 IEEE Workshop on Information Assurance, United States Military Academy*, West Point, NY, 2003.

[44] V. Park and S. Corson, Temporally-Ordered Routing Algorithm (TORA) Ver. 1 Functional Specification, IETF draft, 2001.

[45] T. Clausen and P. Jacquet, Optimized Link State Routing Protocol (OLSR) Project, Hipercom, INRIA, www.ietf.org/rfc/rfc3626.txt, RFC-3626, 2003.

[46] X. Wang, D. Feng, X. Lai, and H. Yu, Collisions for Hash Functions MD4, MD5, HAVAL-128 and RIPEMD, *Cryptology ePrint Archive*, Report 2004/199, http://eprint.iacr.org/, 2004.

[47] T. Karygiannis and L. Owens, Wireless Network Security-802.11, Bluetooth and Handheld Devices. National Institute of Standards and Technology. Technology Administration, U.S Department of Commerce, *Special Publication* 800-848, 2002.

[48] R. Nichols and P. Lekkas, *Wireless Security-Models, Threats, and Solutions*, McGraw-Hill, Chapter 7, 2002.

[49] H. Hsieh and R. Sivakumar, Transport Over Wireless Networks. *Handbook of Wireless Networks and Mobile Computing*, Edited by Ivan Stojmenovic. John Wiley and Sons, Inc., 2002.

[50] N. Weaver, V. Paxson, S. Staniford, and R. Cunningham, "A Taxonomy of Computer Worms", *First Workshop on Rapid Malcode* (WORM), 2003.

[51] C. Kaufman, R. Perlman, and M. Speciner, *Network Security Private Communication in a Public World*, Prentice Hall PTR, A division of Pearson Education, Inc., 2002

[52] S. Capkun, L. Buttyan, and J. Hubaux, Sector: Secure Tracking of Node Encounters in Multi-hop Wireless Networks. *Proc. of the ACM Workshop on Security of Ad Hoc and Sensor Networks*, 2003.

[53] W. Wang, B. Bhargava, Y. Lu, and X. Wu, Defending Against Wormhole Attacks in Mobile Ad Hoc Networks, under review at Wiley Journal Wireless Communication and Mobile Computing (WCMC).

[54] S. Yi, P. Naldurg, and R. Kravets, Security-Aware Ad-hoc Routing for Wireless Networks. Report No.UIUCDCS-R-2002-2290, UIUC, 2002.

[55] V. Gupta, S. V. Krishnamurthy, and M. Faloutsos, Denial of Service Attacks at the MAC Layer in Wireless Ad Hoc Networks. *In Proc. of MILCOM*, 2002.

[56] I. Aad, J. Hubaux, and E. W. Knightly, Denial of Service Resilience in Ad Hoc Networks, *In Proc. of 10th Ann. Int'l Conf. Mobile Computing and Networking* (MobiCom 2004), pp. 202 - 215, ACM Press, 2004.

[57] J. Hubaux, L. Buttyan, and S. Capkun, The Quest for Security in Mobile Ad Hoc Networks, *In Proc. of the ACM Symposium on Mobile Ad Hoc Networking & Computing* (MobiHoc 2001), Long Beach, CA, Oct. 2001.

[58] H. Yang, X. Meng, and S. Lu, Self-organized Network Layer Security in Mobile Ad Hoc Networks, *ACM MOBICOM Wireless Security Workshop* (WiSe'02).

[59] L. Buttyan and J. Hubaux, Nuglets: A Virtual Currency to Simulate Cooperation in Self-organized Ad Hoc Networks. Technial Report DSC/2001/001, Swiss Federal Institute of Technology - Lausanne, 2001.

[60] P. Michiardi and R. Molva, Core: A Collaborative Reputation Mechanism to Enforce Node Cooperation in Mobile Ad Hoc Networks, *IFIP-Communication and Multimedia Security Conference 2002*.

[61] S. Bansal and M. Baker, Observation-based Cooperation Enforcement in Ad Hoc Networks, http://arxiv.rog/pdf/cs.NI/0307012, July 2003.