Discovering End-To-End Communication Patterns in Manets Using GSTARS

Ubokobong Esang [1], C Fancy [2]
M.Tech[1], Assistant Professor Ph.D [2]
Information Security and Cyber Forensics
Department of Information Technology
SRM University, Kattankulathur
Tamil Nadu –India

ABSTRACT
There has been proposed, several anonymity enhancing techniques based on packet encryption to protect the communication anonymity of mobile ad hoc networks (MANETs). However, in this paper, we show that MANETs are still very vulnerable under passive statistical traffic analysis attacks. To demonstrate how to securely transmit packets without the attacker being able to decrypt the packets, we present a novel variant of STA RS, generalized statistical traffic pattern analysis discovery system. (GSTARS). GSTARS works in such a way that the adversaries only need to monitor the nodes beside the boundaries of the super nodes. The traffic inside each super node can be cleared as well as ignored, since it will not affect the inter-region traffic patterns. GSTARS is capable of discovering the potential receivers as well as the destinations contained within one or a few super nodes. Empirical studies demonstrate that GSTARS mitigates the inaccuracy of not identifying the actual receiver of a point-to-point transmission within the sender’s transmission range.

Keywords: Anonymity enhancing techniques, mobile ad-hoc networks, point-to-point transmission, generalized statistical traffic pattern analysis (GSTARS)

I. INTRODUCTION

MANET is a system of wireless mobile nodes that can freely and dynamically self-organize in arbitrary and temporary network topologies without the need of a wired backbone or a centralized administration. The mobile nodes can join into the network or can leave from the network only by interaction with other nodes. The mobile nodes communicate over relatively bandwidth constrained wireless links. The routing functionality will be incorporated into mobile nodes; so that all network activity including discovering the topology and delivering messages must be executed by the node itself. Such perceived advantages elicited immediate interest in the field of military disaster and rescue operation.

Compared to wired networks, MANETs are more vulnerable to both active and passive attacks. Wireless transmissions are easy to capture remotely and undetected, while the lack of central management and monitoring make network nodes susceptible to active attacks. Providing security for MANETs is a challenging task, and many researchers have engaged in designing protocols for diverse security related task such as key management, authentication, confidentiality, etc. Recently researchers have also tackled the problem of anonymity in wireless networks. It is clear that providing anonymity in ad hoc networks is important as users may wish to hide the fact that they are accessing some service or communicating with another user.

MANETs introduce two main problems which are not commonly faced by traditional fixed network routing protocols. These are the lack of fixed infrastructure support and the frequent changes in network topology. Such features pose serious privacy issues for user’s and security threats for the information in an adverse environment. Any user wants to communicate with another user, MANET routing protocols should provide a route the users. There are two categories of routing protocols: reactive and proactive. Most of these routing protocols rely on cooperation between nodes due to the lack of a centralized administration and also assumes that all nodes are genuine and trustworthy. These features of MANET provide an opportunity for malicious user to introduce different kinds of attacks [2] at network layer with respect to routing. A malicious user, may falsely advertise good paths to destination node during route discovery process, may drops the packets selectively, may leak confidential or important information to unauthorized nodes in the network.
Related Work

Traffic analysis attacks against the static wired networks (e.g., Internet) have been well investigated. Traffic analysis attacks against MANETs (e.g., maximum hop-count of a packet) when deriving the end-to-end traffic from the one-hop evidences. Second, it does not provide a method to identify the actual source and destination nodes (or to calculate the source/destination probability distribution. In this paper, We call this variant of STA RS as the Generalized STA RS (GSTARS). To perform GSTARS, the adversaries only need to monitor the nodes beside the boundaries of the super nodes. The traffic inside each supernode can be ignored, since it will not affect the inter-region traffic patterns. In addition, GSTARS does not need the signal detectors to be able to precisely locate the signal source. It is only required to determine which supernode (region) the signals are sent from. Unlike in STA RS, the actual receiver of a point-to-point transmission is not identifiable among all the potential receivers within the sender's transmitting range. This inaccuracy can be mitigated in GSTARS because most potential receivers of a packet will be contained within one or a few super nodes. With GSTARS, packets transmitted from different sources to a destination will be securely received inspite of the traffic encountered when the packet passes through a relay node. GSTARS ensures that the traffic is cleared and packets do not have to wait in a queue. In cases of a heavy traffic in the network, the packets are routed through alternative routes. In section two, i have outlined some works related to this area of research. Section three deals with the existing system which I understudied to improve upon. In section four is the proposed system I adopted to enhance the existing system. Section five describes the modules used to implement this paper. Section six is the conclusion and suggested area of improvement to this paper.

II. RELATED WORK

Traffic analysis attacks against the static wired networks (e.g., Internet) have been well investigated. The brute force attack proposed in [11] tries to track a message by enumerating all possible links a message could traverse, namely, the brute force approach. Recently, statistical traffic analysis attacks have attracted broad interests due to their passive nature, i.e., attackers only need to collect information and perform analysis quietly without changing the network behavior (such as injecting or modifying packets). However, all these previous approaches do not work well to analyze MANET traffic because of the following three nature of MANETs:

1. The broadcasting nature: In wired networks, a point-to-point message transmission usually has only one possible receiver. While in wireless networks, a message is broadcasted which can have multiple possible receivers and so incurs additional uncertainty.

2. The Ad hoc nature: MANETs lack network infrastructure, and each mobile node can serve as both a host and a router. Thus, it is difficult to determine the role of a mobile node to be a source, a destination, or just a relay.

3. The Mobile nature: Most of existing traffic analysis models do not take into consideration the mobility of communication peers, which make the communication relations among mobile nodes more complex.

In [21], Huang devised an evidence-based statistical traffic analysis model specially for MANETs. In this model, every captured packet is treated as an evidence supporting a point-to-point (one-hop) transmission between the sender and the receiver. A sequence of point-to-point traffic matrices is created, and then they are used to derive end-to-end (multihop) relations. This approach provides a practical attacking framework against MANETs but still leaves substantial information about the communication patterns undiscovered. First, the scheme fails to address several important constrains (e.g., maximum hop-count of a packet) when deriving the end-to-end traffic from the one-hop evidences. Second, it does not provide a method to identify the actual source and destination nodes (or to calculate the source/destination probability distribution. In this paper, We call this variant of STA RS as the Generalized STA RS (GSTARS). To perform GSTARS, the adversaries only need to monitor the nodes beside the boundaries of the super nodes. The traffic inside each supernode can be ignored, since it will not affect the inter-region traffic patterns. In addition, GSTARS does not need the signal detectors to be able to precisely locate the signal source. It is only required to determine which supernode (region) the signals are sent from. Unlike in STA RS, the actual receiver of a point-to-point transmission is not identifiable among all the potential receivers within the sender's transmitting range. This inaccuracy can be mitigated in GSTARS because most potential receivers of a packet will be contained within one or a few super nodes. With GSTARS, packets transmitted from different sources to a destination will be securely received inspite of the traffic encountered when the packets pass through a relay node. GSTARS ensures that the traffic is cleared and packets do not have to wait in a queue. In cases of a heavy traffic in the network, the packets are routed through alternative routes. In section two, i have outlined some works related to this area of research. Section three deals with the existing system which I understudied to improve upon. In section four is the proposed system I adopted to enhance the existing system. Section five describes the modules used to implement this paper. Section six is the conclusion and suggested area of improvement to this paper.
latencies. The message tagging attacks (e.g., [12]) require attackers to occupy at least one node that works as a router in the communication path so that they can tag some of the forwarded messages for traffic analysis. By recognizing the tags in latter transmission hops, attackers can track the traffic flow.

Mobile ad-hoc networks (MANETs) have great potential in hostile battlefield-like environments without a wired communication infrastructure. The shared wireless medium in a MANET unfortunately enables passive, adversarial eavesdropping on arbitrary radio transmissions. The adversary can then run traffic analysis on overheard transmissions to infer the network traffic pattern, which consists of a set of end-to-end flows with each described by a 6-tuple, (source, destination, start-time, end-time, rate, path). The disclosure of the traffic pattern and its changes is often devastating for a mission-critical MANET. For example, a node as the source or destination of many end-to-end flows may be a VIP node which often issues tactical commands or collects tactical information for making critical decisions. In addition, high-rate flows may imply the relationships of the two end nodes in terms of rank (a node may be allowed to communicate with others with rank just above or below itself). Also, an unexpected change of the traffic pattern in a tactical MANET may indicate a forthcoming action, a chain of commands, or a state change of network alertness [1]. The adversary can then exploit the obtained information to launch various targeted attacks such as compromising the VIP nodes. Anonymous routing protocols were proposed as a countermeasure against malicious traffic analysis in MANETs. They were aimed at preventing inferring the traffic pattern by hiding the real sources, real destinations, and source destination pairs of overheard packets. These schemes can withstand a local adversary who is incapable of overhearing every radio transmission to various degrees. It remains unclear whether they can defeat a global adversary who is able to eavesdrop on every radio transmission. An improvement over the previously known disclosure attack is presented that allows, using statistical methods, to effectively de-anonymize users of a mix system. Furthermore the statistical disclosure attack is computationally efficient, and the conditions for it to be possible and accurate are much better understood. The new attack can be generalized easily to a variety of anonymity systems beyond mix networks. The statistical disclosure attack only relies on collecting observations and performing trivial operations on vectors, and therefore is computationally cheap and scales very well. Therefore the collection of observations, and the calculation of anonymity sets corresponding to messages to be the main computational bottleneck of an attacker. Algorithms used here are the probability distribution algorithm, uniform distribution, and recipient anonymity. The statistical disclosure attack does not simply provide a computational improvement over the disclosure attack, but also presents important new features. The conditions for it to be possible can be expressed in closed algebraic form, as presented above, and therefore no simulations are required to decide when it is applicable and effective. An important improvement over the previous work is also the fact that the statistical disclosure attack can be applied when the probability distributions described by vectors \( v, u \) and \( o_i \) are not uniform, but are skewed.

III. EXISTING SYSTEM

In existing system, the brute force attack tries to track a message by enumerating all possible links a message could traverse. In blending attacks, attacker easily modifies messages and reordered by the system. If the attacker can monitor the latency of each path, he can correlate the messages coming in and out of the system by analyzing their transmission latencies. Moreover, in a MANET protected by anonymity enhancing techniques, it is a difficult task itself to identify an actual destination node as the target due to the ad hoc nature. The adversaries are not able to determine whether a particular node is a destination depending on whether the node sends out traffic. Nonetheless, the statistical disclosure attacks cannot be applied to MANETs either, because the attackers cannot easily identify the actual source nodes in MANETs.

Limitations of the existing system

1. It is fails to address several important constrains (e.g., maximum hop-count of a packet) when deriving the end-to-end traffic from the one hop evidences.
2. It is does not provide a method to identify the actual source and destination nodes (or to calculate the source/destination probability distribution).
3. The source node handles path resolution and packet routing.

IV. PROPOSED WORK

This paper aims to derive the source/destination probability distribution, each node to be a message source/destination, and the end-to-end link probability distribution, i.e., the probability for each pair of To achieve its goals, GSTARS makes uses an improvement of STARS which includes two major steps: 1) The time slicing technique which is used to construct point-to-point traffic matrices, and then derive the
end-to-end traffic matrix with a set of traffic filtering rules; and 2) Apply a heuristic approach to identify the actual source and destination nodes, and then correlate the source nodes with their corresponding destinations. The contribution of STARS is most of the previous approaches are partial attacks in the sense that they either only try to identify the source (or destination) nodes or to find out the corresponding destination (source) nodes for given particular source (destination) nodes. STARS are a complete attacking system that first identifies all source and destination nodes and then determines their relationship.

Advantages:

- It is easy to identify the destination.
- Traffic between the mobile nodes can be detected.
- Point-to-point traffic matrices are constructed using the time-slicing technique to forward the packets to its destination.
- We can derive the end-to-end traffic matrix with a set of traffic filtering rules.
- Packets are encrypted and transmitted to the destination.

1. Point-to-point algorithm which is used to detect the traffic.
2. End-to-end anonymity which is used to redirect the packet.

THE TRAFFIC PATTERN DISCOVERY MODEL

To discover the hidden traffic patterns in a MANET system, this model includes two steps. Initially, it captures the raw traffics to construct point-to-point traffic matrices and then derives the end-to-end traffic matrices. Second it calculates the probability for each node to be a source/destination, for further analyzing the end-to-end traffic matrices. From that each pair of source and destination nodes of end-to-end probability distribution can be obtained.

In this paper, we present a fundamental system adopted (assumed) by GSTARS such as:

Attackers Model

The attackers’ goal is to discover the traffic patterns among mobile nodes. Particularly, we have the following four assumptions for attackers:

1. The adversaries are passive signal detectors, i.e., they are not actively involved in the communications. They can monitor every single packet transmitted through the network.

2. The adversary nodes are connected through an additional channel which is different from the one used by the target MANET. Therefore, the communication between adversaries will not influence the MANET communication.

3. The adversaries can locate the signal source according to certain properties (e.g., transmission power and direction) of the detected signal, by using wireless location tracking techniques [25] such as triangulation, nearest sensor, or RF fingerprinting. Take note that none of these techniques can identify the source of a signal from several nodes very close to each other. Hence, this assumption actually indicates that the targeted networks are sparse in terms of the node density. In other words, any two nodes in such a network are distant from each other so that the location tracking techniques in use are able to uniquely identify the source of a wireless signal. In the following of this paper, unless specifically denoted as "signal source" or "source of signal," the word “source” indicates the source of a network flow.

4. The adversaries can trace the movement of each mobile node, by using cameras or other types of sensors. In this case, the signals (packets) transmitted by a node can always be associated with it even when the node moves from one spot to another.

1. Point-to-point algorithm which is used to detect the traffic.
2. End-to-end anonymity which is used to redirect the packet.

Fig 1: Work flow label

Communication Model

We assume the anonymity enhancing techniques (such as [1],[2],[3]) are used to protect the MANETs. However, these techniques are designed to different levels of anonymity. To focus on the statistical traffic analysis, we assume, based on [21], that a combination of these techniques is applied and the targeted MANET communication system is subject to the following model:

1. The PHY/MAC layer is controlled by the commonly used 802.11(a/b/g) protocol. But all MAC frames (packets) are encrypted so that the adversaries cannot decrypt them to look into the contents.
2. Padding is applied so that all MAC frames (packets) have the same size. Nobody can trace a packet according to its unique size.

3. The “virtual carrier sensing” option is disabled. The source/destination addresses in MAC and IP headers are set to a broadcasting address (i.e., all “1”) or to use identifier changing techniques. In this case, adversaries are prevented from identifying point-to-point communication relations.

4. No information about the traffic patterns is disclosed from the routing layer and above.

5. Dummy traffic and dummy delay are not used due to the highly restricted resources in MANETs.

In order to clear the traffic patterns in a MANET communication system, GSTARS includes two major steps adopted from STARS. First, it uses the captured traffic to construct a sequence of point-to-point traffic matrices and then derives the end-to-end traffic matrix. Second, further analyzing the end-to-end traffic matrix, it calculates the probability for each node to be a source/destination (the source/destination probability distribution) and that for each pair of node to be an end-to-end communication link (the end-to-end link probability distribution).

To illustrate the basic idea of GSTARS, we use a simple scenario shown in Fig. 1 as an example. In this network, there are three wireless nodes (1, 2, and 3). Node 2 is located in the transmission range of node 1, and node 3 is located in the transmission range of node 2 (but not the transmission range of node 1). Two consecutive packets are detected: node 1 broadcasts a packet and then node 2 broadcasts a packet.

V. MODULES

1. MANET communication system
2. Point-to-point traffic matrices
3. End-to-end traffic matrix
4. Advanced encryption Standard
2. All nodes within the transmitting range of a packet have the same probability to be the actual receiver. This is achieved by dividing a packet into n sub-packets and each sent to one neighboring node. For simplicity, we denote the original packet as “virtual size” 1 and each of the subpackets as “virtual size” 1/n.

3. Each packet p has three associated features: p:vsize, p:time, and p:hop, denoting the “virtual size,” transmitting time, and hop count of this packet, respectively. A packet’s hop count is set to 1 when added to the point-to-point traffic matrix.

3. End-to-end traffic matrices
Given a sequence of point-to-point traffic matrices Wj1:K, our goal is to derive the end-to-end traffic matrix R ¼ is the accumulative traffic volume from node i to node j, including both the point-to-point traffic captured directly and multi-hop traffic deduced from the point-to-point traffic. In this paper, we use the term accumulative traffic matrix and end-to-end traffic matrix interchangeably. The following Algorithm 1(function f) takes \[ W_{1:K} \] as the input to derive the accumulative traffic matrix R.

Algorithm 1. \( f(W_{1:K}) \).
1. \( R = W_1 \)
2. for \( e = 1 \) to \( K - 1 \)
3. \( R = g(R, W_{e+1}) + W_{e+1} \)
4. end for
5. return \( R \)

In algorithm 2 as described below, each update to \( R \) (line 3) includes the multi-hop traffic derivation function \( g \) shown as in Algorithm 2, and the addition of the point-to-point traffic matrix which is the evidence of possible direct (single-hop) communication.

Function \( g \) takes two inputs: 1) \( R \) is an end-to-end traffic matrix derived from point-to-point matrices \( W_1 \) to \( W_e \), and 2) \( W_{e+1} \) is the next point-to-point traffic matrix. The output is the end-to-end traffic matrix derived from \( W_1 \) to \( W_{e+1} \). For each packet x recorded in \( W_{e+1} \), the function tries to find a packet y in R that is potentially the same packet transmitted at x’s previous hop. If such a packet y exists, then a multi-hop flow (packet) from the source of y to the destination of x should be derived. For instance, in our example scenario, we first let \( R \) \( \frac{1}{4}W_1 \). Then \( g(R, W_2) \) should derive all possible end-to-end flows. \( W_2 \) contains two packets, sent from node 2 to nodes 1 and 3, respectively. Let \( p(2,1) \) and \( p(2,3) \) denote these two packets. The current R contains only one packet \( p(1;2) \) sent from node 1 to node 2. Thus, it is possible that \( p(1;2) \) and \( p(2;3) \) are the same packet appearing at different hops. In this case, a new packet \( p(1;3) \) is derived to represent a multi-hop flow from node 1 to node 3. Since the volume of a multi-hop flow consists of a sequence of one-hop transmissions cannot exceed the volume of any of the transmissions, we have \( p1,3.vsize = \min\{p1;2:vsize; p2;3:vsize\} = 0.5 \). Two constraints are considered for reasonable traffic inference: The difference between the transmitting times of a packet at two consecutive hops cannot be too large and the hop-count of a packet cannot exceed a maximum value.

2. Advanced Encryption Standard
The Advanced Encryption Standard (AES) is a symmetric-key block cipher algorithm and U.S. government standard for secure and classified data encryption and decryption. AES is based on a design principle known as a substitution-permutation network, combination of both substitution and permutation, and is fast in both software and hardware. In this paper the information will be sent in an encrypted form. For that, we use AES Algorithm to encrypt the data based on key size. So whatever the data sent to the input to the AES Algorithm will be converted to cipher text with 128 bit cycle.

VI. EXPERIMENTS
In this section, we carry out an empirical study of MANETs under the concept of GSTARS in relation to STARS, consisting of two components: demonstration and experimental results (evaluation). The network environment is simulated using Fedora with a tool called Network simulator which are all mounted on a VMware LINUX platform.

Demonstration
The MANET for demonstration is comprised of 40 mobile nodes randomly deployed in an 800 x 800m² area. In this experiment we demonstrate three source nodes and three
destinations in order to demonstrate the point-to-point and end-to-end communication and efficient packet routing in this MANET using the adhoc on-demand distance vector routing algorithm. AODV is a routing protocol designed for ad hoc mobile networks. AODV is capable of both unicast and multicast routing. It is an on demand algorithm, meaning that it builds routes between nodes only as desired by source nodes. It maintains these routes as long as they are needed by the sources. Additionally, AODV forms trees which connect multicast group members. The trees are composed of the group members and the nodes needed to connect the members. AODV uses sequence numbers to ensure the freshness of routes. It is loop-free, self-starting, and scales to large numbers of mobile nodes.

AODV builds routes using a route request / route reply query cycle. When a source node desires a route to a destination for which it does not already have a route, it broadcasts a route request (RREQ) packet across the network. Nodes receiving this packet update their information for the source node and set up backwards pointers to the source node in the route tables. In addition to the source node's IP address, current sequence number, and broadcast ID, the RREQ also contains the most recent sequence number for the destination of which the source node is aware. A node receiving the RREQ may send a route reply (RREP) if it is either the destination or if it has a route to the destination with corresponding sequence number greater than or equal to that contained in the RREQ. If this is the case, it unicasts a RREP back to the source. Otherwise, it rebroadcasts the RREQ. Nodes keep track of the RREQ's source IP address and broadcast ID. If they receive a RREQ which they have already processed, they discard the RREQ and do not forward it. As the RREP propagates back to the source, nodes set up forward pointers to the destination. Once the source node receives the RREP, it may begin to forward data packets to the destination. If the source later receives a RREP containing a greater sequence number or contains the same sequence number with a smaller hop-count, it may update its routing information for that destination and begin using the better route. As long as the route remains active, it will continue to be maintained. A route is considered active as long as there are data packets periodically travelling from the source to the destination along that path. Once the source stops sending data packets, the links will time out and eventually be deleted from the intermediate node routing tables. If a link break occurs while the route is active, the node upstream of the break propagates a route error (RERR) message to the source node to inform it of the now unreachable destination(s). After receiving the RERR, if the source node still desires the route, it can reinitiate route discovery.

Experimental Results

After a successful execution of this experiment, the following results and evaluations were made with a thorough comparison of GSTARS and STARS. The results derived are based on the following matrix: Bandwidth, Average delay, energy consumed, overhead, packet delivery ration, message drop.
VII. CONCLUSION AND FUTURE WORK

In this paper, there is some tendency to propose a completely unique GSTARS model for MANETs. GSTARS makes use of the supernode (neighbouring node) concept to perform efficient routing in an end-to-end system. From the captured packets, GSTARS constructs a sequence of point-to-point traffic matrices to derive the end-to-end traffic matrix. To conclude the evaluation, the hidden traffic patterns can be discovered in good accuracy using GSTARS, even without the number of actual sources, destinations, and end-to-end communication relations known to the traffic analyzers and also secure the data packets by the way of applying AES.

Our study demonstrates an existing system that is able to do ensure an efficient and effective communication even during the attack of a typical MANET.

Before sending the file or packets, we should be able to check the traffic status and if the network is suitable for transmission.

REFERENCES


