ISSN:2348-2079

Volume-5 Issue-2

International Journal of Intellectual Advancements and Research in Engineering Computations

Finding end to end communication for discovery system in manet using anonymous supernode

1. Ms. K.E.Eswari M.C.A.,M.Phil.,M.E., Associate Professor
2. Ms. G.Kiruthika, Final M.C.A.,
Department of MCA, Nandha Engineering College (Autonomous), Erode-52.
kiruthikabsc2263@gmail.com

Abstract--Many anonymity enhancing techniques have been proposed based on packet encryption to protect the communication anonymity of Mobile Ad Hoc Networks(MANETs). However, in this paper, MANETs are still vulnerable under passive statistical traffic analysis attacks. To demonstrate how to discover the communication patterns without decrypting the captured packets and present a novel Statistical Traffic Pattern Discovery System (STARS). STARS works passively to perform traffic analysis based on statistical characteristics of captured raw traffic. STARS is capable of discovering the sources, the destinations, and the end-to-end communication relations. Empirical studies demonstrate that STARS achieves good accuracy in disclosing the hidden traffic patterns.

1. INTRODUCTION

Mobile ad hoc networks (MANETs) are originally designed for military tactic environments. Communication anonymity is a critical issue in MANETs, which generally consists of the following aspects: 1)Source or destination anonymity—it is difficult to identify the sources or the destinations of the network flows. 2) The End-to-end relationship anonymity—it is difficult to identify the end- to-end communication relations

In this paper presents a novel statistical traffic pattern discovery system (STPDS). STPDS works passively to perform traffic analysis is based on the statistical characteristics of captured raw traffic. STPDS is capable of discovering the sources, and destinations [1], and the end-to-end communication relations. Empirical studies demonstrate that STPDS achieves good accuracy in disclosing the hidden traffic patterns system. The contribution of STPDS is twofold: 1) STPDS is the first statistical traffic analysis approach that considers the salient characteristics of MANETs: the broadcasting, ad hoc, and mobile nature; and 2) The 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 or source nodes for given Particular source or destination nodes. STPDS is a complete attacking system that first identifies all source and destination nodes and then determines their relationship [1][2].

In addition, the STPDS is extended as GSTPDS which 1) The entire networks are divides into multiple regions geographically 2) The deploys sensors along the boundaries of each region to monitor the cross-component traffic 3)It treats each region as a super node and use STPDS to figure out the sources, destinations, and end-to-end communication relations and 4) analyzes the traffic even when nodes are close to each other by treating the close nodes as a super node.

The contribution of STARS is twofold: 1) To the best of our knowledge, this STARS is the first statistical traffic analysis approach that considers the salient characteristics of MANETs: the broadcasting, ad hoc, and mobile nature.

2) 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[3]. STARS is a complete attacking system that first identifies all source and destination nodes and then determines their relationship.

2. SYSTEM MODELS

In this section, present the fundamental system models adopted assumed by STARS.

2.1 Communication Model

The anonymity enhancing techniques are used to protect the MANETs. However, these techniques are designed to different levels of anonymity[6].

To focus on the statistical traffic analysis, based on that a combination of these techniques is applied and the targeted MANET communication system is subject to the following model:

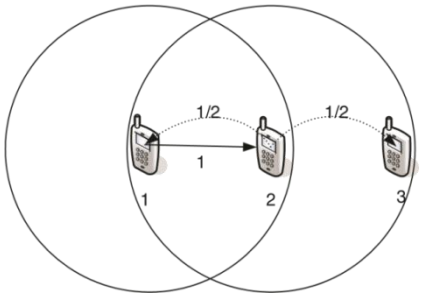
1. The PHY/MAC layer is controlled by the commonly used 802.11(a/b/g) protocol. But all MAC frames packets are encrypted so that the adversaries cannot decrypt them to look into the contents.
2. Padding is applied so that all MAC frames packets have the same size. Nobody can trace a packet according 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[8].
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.

2.2Attack 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.
3. Therefore, the communication between adversaries will not influence the MANET communication.
4. The adversaries can locate the signal source according to certain properties e.g., transmission power and direction of the detected signal.
5. By using wireless location tracking techniques such as triangulation, nearest sensor, or RF fingerprinting. Note that none of these techniques can identify the source of a signal from several nodes very close to each other.
6. Hence, this assumption actually indicates that the targeted networks are sparse in terms of the node density.

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.



A simple wireless ad hoc network

3. STATISTICAL TRAFFIC PATTERN DISCOVERYSYSTEM

To disclose the hidden traffic patterns in a MANET communication system, STARS includes two major steps. 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[7].

To illustrate the basic idea of STARS, we use a simple scenario. 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.

3.1 Traffic Matrices Construction

3.1.1 Point-to-Point Traffic Matrix

With the captured point-to-point one-hop traffic in a certain period T, we first need to build point-to-point traffic matrices such that each traffic matrix only contains "independent" one-hop packets.

Note that two packets captured at different time could be the same packet appearing at different locations, such as the two packets sent by node 1 and node 2 consecutively in Fig. 1, so they are "dependent" on each other.

To avoid a single point-to-point traffic matrix from containing two dependent packets, we apply a "time slicing" technique as shown in Fig. 2. That is, we take snapshots of the network, and each snapshot is triggered by a captured packet. A sequence of snapshots during a time interval t_e constructs a slice represented by

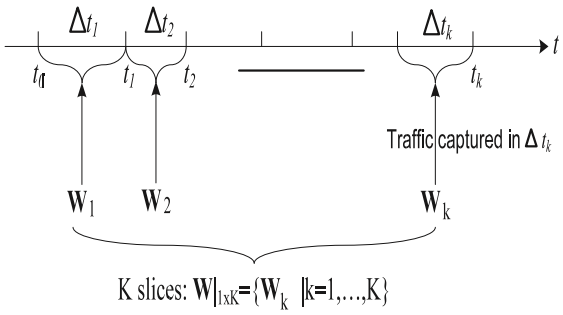


Fig. 2. Slicing the time domain

a traffic matrix W_e, which is an N one-hop traffic relation matrix. The length of each time interval t_e is determined by two criteria:

1) A node can be either a sender or a receiver within this time interval. But it cannot be both.

2) Each traffic matrix must correctly represent the one-hop transmissions during the corresponding time interval. In this way, the construction of matrices Wj_{1e} ¼ ðW1; W2;...; WKÞwill automatically involve mobility in the traffic matrices constructions. For example, traffic matrix W_e ¼ðweði; jÞÞNNis created for direct transmissions between nodes during time interval t_e. Since each snapshot of the network is triggered by capturing a packet, as long as potential receiver j is located within sender i's communication range i.e., $d_{i;j}$ r, a small change of distance $d_{i;j}$ due to mobility will not alter the value assigned to $w_e$ði;jÞ

In addition to the "time slicing," we need to follow the three rules listed below: 1) The number of captured packets rather than the actual size of payloads is considered as the "traffic volume," since the size of payloads does not affect the traffic pattern and we assumed all MAC frames are of the same length due to the application of padding.

2) All nodes within the transmitting range of a packet have the same probability to be the actual receiver. For example, if a node i broadcasts a packet in the time interval t_e, and nodes $j_1;j_2;...j_n$ are all within i's transmitting range, then the entries $w_e$ði;j_1Þ, $w_e$ði;j_2Þ...$w_e$ði;j_nÞ should be all equally increased by 1=n. This is equivalent to dividing a packet into n subpackets and each sent to one neighboring node. For simplicity, in the remainder of the paper, we denote the original packet as "virtual size" 1 and each of the subpackets as "virtual size" 1=n. 3

Each packet p in $w_e$ði;jÞ:pkt, 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.

3.1.2 End-to-End Traffic Matrix

Given a sequence of point-to-point traffic matrices Wj_{1K}, our goal is to derive the end-to-end traffic matrix R ¼ ðrði;jÞÞNN, where rði;jÞ is the accumulative traffic volume from node i to node j, including both the point-to-point traffic captured directly and multihop 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 Wj_{1K} as the inputs to derive the accumulative traffic matrix R.

Algorithm 1. —fðWj_{1K}Þ.

1: R ¼ W_1

2: for e ¼ 1 to K 1 do

3: R ¼ gðR;$W_{eþ1}$ÞÞ$W_{eþ1}$

4: end for

5: return R

In this algorithm, each update to R line includes the multihop 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 singlehop 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) $W_{eþ1}$ is the next point-to-point traffic matrix. The output is the end-to-end traffic matrix derived from W_1 to $W_{eþ1}$.

For each packet x recorded in $W_{eþ1}$, 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

multihop 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 ¼ W_1. Then gðR;W_2Þ 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 multihop flow from node 1 to node 3. Since the volume of a multihop flow consisting of a sequence of one-hop transmissions cannot exceed the volume of any of the transmissions, we have $p_{1;3}$:vsize ¼ min$f p_{1;2}$:vsize;$p_{2;3}$:vsizeg¼ The timing threshold T must be at least the value of the maximum retransmission time.

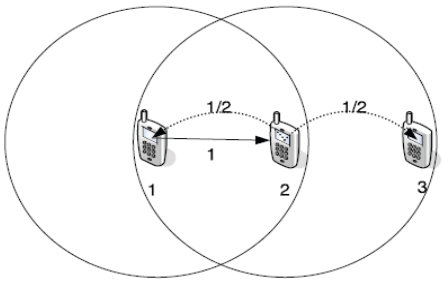
3.2 Traffic Pattern Discovery

The traffic matrix R tells us the deduced end-to-end traffic volume between each pair of nodes. However, we still need to perform further investigation to discover the actual source/destination probability distribution and end-to-end link probability distribution, that is, to figure out who are the actual sources and destinations and who are communicating with whom.

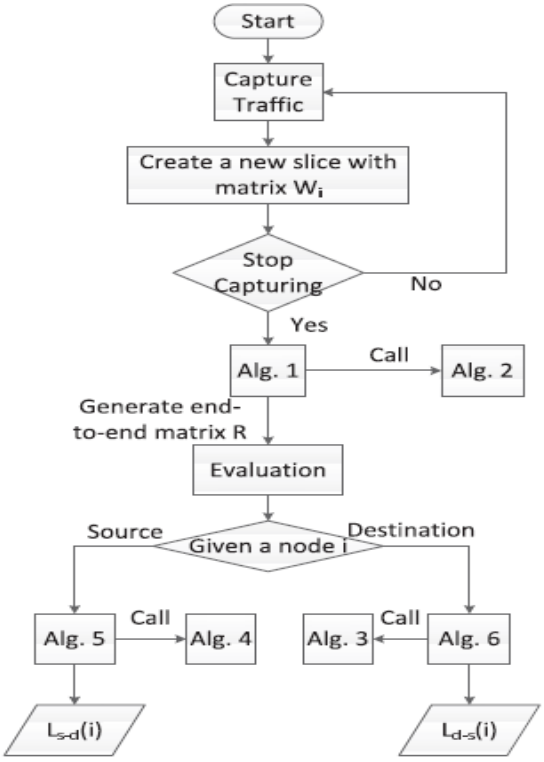
3.2.1 Source /Destination Probability Distribution

We denote the actual source and destination probability distribution, respectively, as two vectors S ¼ðsð1Þ;sð2Þ;...; sðNÞÞ and D ¼ðdð1Þ;dð2Þ;...;dðNÞÞ, where sðiÞ and dðiÞ i ¼ 1 to N represent the probability for node i to be an actual source and destination, respectively. Note that if the total number of source nodes is m, then we should have P ^{N}i¼1 sðiÞ¼ m for S. However, since we only care about the relative order among all possibilities to know which nodes are more possible to be the actual sources but not the total number m, we can always assume m ¼ 1. It is the same case for D and all the probability vectors we will calculate later in this paper. That is, all probability distribution vectors in this paper are normalized[1] and only the relative orders among the elements of each vector actually make sense.

SYSTEM ARCHITECTURE



WORK FLOW



4. DISCUSSION AND FUTURE WORK

The adversarial model presented in Section 3.2 assumes that the adversaries can globally monitor the traffic across the entire network region. This assumption is conservative from the network users' point of view. Usually, it is difficult for the attackers to perform such a global traffic detection. However, even though the adversaries are not able to monitor the entire network, they can monitor several parts of the network simultaneously

This paper proposes a novel STARS for MANETs. STARS is basically an attacking system, which only needs to capture the raw traffic from the PHY/MAC layer without looking into the contents of the intercepted packets. From the captured packets, STARS constructs a sequence of point-to-point traffic matrices to derive the end-to-end traffic matrix, and then uses a heuristic data

processing model to reveal the hidden traffic patterns from the end-to-end matrix.

The empirical study demonstrates that the existing MANET systems can achieve very restricted communication anonymity under the attack of STARS. In addition, the STPDS is extended as GSTPDS which divides the entire network into multiple regions geographically and deploys sensors along the boundaries of each region to monitor the cross-component traffic.

Also it treats each region as a super node and use STPDS to figure out the sources, destinations, and end-to-end communication relations.

5. RELATED WORK

Traffic analysis attacks against the static wired networks e.g., Internet have been well investigated. The brute force attack proposed in tries to track a message by enumerating all possible links a message could traverse. In node flushing attacks blending attacks, n1 attacks, the attacker sends a large quantity of messages to the targeted anonymous system which is called a mix-net. Since most of the messages modified and reordered by the system are generated by the attacker, the attacker can track the rest a few messages. The timing attacks as proposed in focus on the delay on each communication path.

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 [7]. The message tagging attacks 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. The watermarking attacks are actually variants of the message tagging attacks. They reveal the end-to-end communication relations by purposely introducing latency to selected packets.

Different from the attacks mentioned above, statistical traffic analysis intends to discover sensitive information from the statistical characteristics of the network traffic, for example, the traffic volume. The adversaries usually do not change the network behavior such as injecting or modifying packets. The only thing they do is to quietly collect traffic information and perform statistical calculations[7].

In a typical predecessor attack, the attackers act exactly as legitimate nodes in the network communications. They collectively maintain a single predecessor counter for each legitimate node in the system. When an attacker finds himself to be on an anonymous path to the targeted destination, he increments the shared counter for its predecessor node in this path. The counters are then used for the attackers to infer the possible source nodes of the given destination. Obviously, to launch such an attack, a large number of legitimate nodes must first be compromised and controlled by the attackers. This is usually not achievable in MANETs. 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. That is, destinations are indistinguishable from other nodes e.g., relays in a MANET. In fact, they usually act as relay nodes as well, forwarding traffic for others. The adversaries are not able to determine whether a particular node is a destination depending on whether the node sends out traffic. This is totally different from the situation in traditional infrastructural networks where the role of every node is determined.

Due to the unique characteristics of MANETs, very limited investigation has been conducted on traffic analysis in the context of MANETs. A proposed timing-based approach in to trace down the potential destinations given a known source. In this approach, assuming the transmission delays are bounded at each relay node, they estimate the flow rates of communication paths using packet matching. Then based on the estimated flow rates, a set of nodes that partition the network into two parts, one part to which the source can communicate in sufficient rate and the other to which it cannot, are identified to estimate the potential destinations. Liu et al designed a traffic inference algorithm (TIA) for MANETs based on the assumption that the difference between data frames, routing frames, and MAC control frames is visible to the passive adversaries, so that they can recognize the point-to-point traffic using the MAC control frames, identify the end-to-end flows by tracing the routing frames, and then infer the actual traffic pattern using the data frames. The TIA achieves good accuracy in traffic

inference, while the mechanism is tightly tied to particular anonymous routing protocols but not a general approach. Both analytical strategies which heavily rely on the deterministic network behaviors.

6. CONCLUSION

This paper surveys and compares various Existing Methods available to analyze statistical traffic from MANETs. Our empirical study demonstrates that the existing MANET systems can achieve very restricted communication anonymity under the attack of STARS. In STARS, the actual receiver of a point-to-point transmission is not identifiable among all the potential receivers within the sender's transmitting range.

REFERENCES

[1] J. Kong, X. Hong, and M. Gerla, "An Identity-Free and OnDemand Routing Scheme against Anonymity Threats in Mobile Ad Hoc Networks," IEEE Trans. Mobile Computing, vol. 6, no. 8 , pp. 888-902, Aug. 2007.

[2] Y. Zhang, W. Liu, W. Lou, and Y. Fang, "MASK: Anonymous OnDemand Routing in Mobile Ad Hoc Networks," IEEE Trans.
Wireless Comm., vol. 5, no. 9, pp. 2376-2385, Sept. 2006.

[3] Y. Qin and D. Huang, "OLAR: On-Demand Lightweight Anonymous Routing in MANETs," Proc. Fourth Int'l Conf. Mobile Computing and Ubiquitous Networking (ICMU '08), pp. 72-79, 2008.

[4] Arunkumar R, Bharateshhegde, Ganeshprasad, Manoj C Jagatap, Vishwas S, "MTPD: MANET Traffic Pattern Discovery –A HeuristicAproach" Volume No.02, Issue No. 06, June 2014

[5] T. He, H. Wong, and K. Lee, "Traffic Analysis in Anonymous MANETs," Proc. Military Comm. Conf. (MILCOM '08), pp. 1-7, 2008.

[6] Y. Liu, R. Zhang, J. Shi, and Y. Zhang, "Traffic Inference in Anonymous MANETs," Proc. IEEE Seventh Ann. Comm. Soc. Conf. Sensor Mesh and Ad Hoc Comm. and Networks (SECON '10), pp. 1-9.

[7] D.Chaum, "Untraceable Electronic Mail, Return Addresses, and Digital Pseudonyms," Comm. ACM, vol. 24, no. 2, pp. 84-88, 1981.

[8] J. Raymond, "Traffic Analysis: Protocols, Attacks, Design Issues, and Open Problems," Proc. Int'l Workshop Designing Privacy Enhancing Technologies: Design Issues in Anonymity and Unobservability, pp. 10-29, 2001.