International Journal of Research in Advent Technology (IJRAT) (E-ISSN: 2321-9637)
Special Issue
National Conference "CONVERGENCE 2016", 06th-07th April 2016

Survey on Traffic Pattern Discovery System For MANETs

Sanchali S. Pandhare[1]

PG student, Department of Computer Science & engineering ,
Pankaj Laddhad Institute of Technology & Management Studies, Buldana.
sanchalipandhare30@gmail.com

Dr. P. M. Jawandiya

Principal, Pankaj Laddhad Institute of Technology & Management Studies, Buldana.
pmjawandiya@rediffmail.com

Abstract :
Different type of 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, we show that MANETs are still vulnerable under passive statistical traffic analysis attacks. To demonstrate how to discover the communication patterns without decrypting the captured packets, we 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.

Keywords: **Mobile Ad-hoc networks, Encryption, Decryption,Statistical traffic pattern.**

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/ destination anonymity- it is difficult to identify the sources or the destinations of the network flows.

2) End-to-end relationship anonymity- it is difficult to identify the endto- end communication relations.

To achieve anonymous MANET communications, many anonymous routing protocols such as ANODR, MASK , and OLAR have been proposed. Though a variety of anonymity enhancing techniques like onion routing and mix-net are utilized, these protocols mostly rely on packet encryption to hide sensitive information (e.g., nodes, identities and routing information) from the adversaries. However, passive signal detectors can still eavesdrop on the wireless channels, intercept the transmissions, and then perform traffic analysis attacks

The contribution of STARS is two fold: 1) To the best of our knowledge, STARS is the first statistical traffic analysis approach that considers the salient characteristics of MANETs: the broadcasting, ad hoc, and mobile nature; and 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. STARS is a complete attacking system that first identifies all source and destination nodes and then determines their relationship.

2.LITERATURE REVIEW & RELATED WORK

In [3], 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). Moreover, it only uses a naı¨ve accumulative traffic ratio to infer the end-to-end communication relations (e.g., the probability for node j to be the intended destination of node i is computed as the ratio of the traffic from i to j to all traffic coming out from node i), which incurs a lot of inaccuracy in the derived probability distributions

Due to the unique characteristics of MANETs, very limited investigation has been conducted on traffic analysis in the context of MANETs. In [4] He H. Wong proposed a timing-based approach is 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.

In [2], 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 good accuracy in traffic inference, while the mechanism is tightly tied to particular anonymous protocols but not a general approach. Both [4] and [2] are analytical strategies which heavily rely on the deterministic network behaviors.

Traffic analysis attacks against the static wired networks (e.g., Internet) have been well investigated. The brute force attack proposed in [13] tries to track a message by enumerating all possible links a message could traverse. In node flushing 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 (normal) messages. The timing attacks as proposed in [14] focus on the delay on each communication path. 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.

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.

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 indis-tinguishable 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.

The statistical disclosure attacks as mentioned in [10] are similar. A statistical disclosure attack often targets a particular given source node and intends to expose its corresponding destinations. It is assumed that the packets initiated by the source are sent to several destinations with certain prob-ability distribution. The background (covering) traffic also has certain probability distribution (usually assumed to be uniformly distributed).

After a large number of observations, the attackers are able to figure out the possible destinations of the given source.

Nonetheless, the statistical disclosure attacks cannot be applied to MANETs either, because the attackers cannot easily identify the actual source nodes in MANETs. Even if a source node is identified, the attacks can only be performed when the attackers know for sure when the targeted source is originating traffic and can observe the network behaviour in the absence of the source. However, the attackers are prevented from being able to do so by the ad hoc nature of MANETs, i.e., they cannot tell if the source is originating traffic or just forwarding traffic as a relay.

3.EXISTING SYSTEM

Evidence-based statistical traffic analysis model, every captured packet is treated as 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. MANET systems can achieve very restricted communication anonymity under the attack of STARS.

Statistical traffic analysis attacks have attracted broad interests due to their passive nature, i.e., attackers only need toc ollect information and perform analysis quietly without changing the network behavior (such as injecting or modifying packets). The predecessor attacks and disclosure attacks are two representatives.

However, all these previous approaches do not work well to analyze MANET traffic because of the following three natures 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 in curs 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 does not take into consideration the mobility of communication peers, which make the communication relations among mobile nodes more complex.

DISADVANTAGES OF EXISTING SYSTEM:

- Approaches do not work well to analyze MANET traffic.
- The scheme fails to address several important constrains when deriving the end-to-end traffic from the one hop evidences.
- It does not provide a method to identify the actual source and destination nodes (or to calculate the source/destination probability distribution).
- 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.

4.PROPOSED SYSTEM:

We propose a novel STARS(statistical traffic pattern discovery system)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.

In this paper, we propose a novel statistical traffic pattern discovery system (STARS). STARS aims to derive the source/destination probability distribution, i.e., the probability for each node to be a message source/destination, and the end-to-end link probability distribution, i.e., the probability for each pair of nodes to be an end-to-end communication pair.

To achieve its goals, STARS includes two major steps:

1) Construct point-to-point traffic matrices using the time-slicing technique, 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.

ADVANTAGES OF PROPOSED SYSTEM:

The attacker can take advantage of STARS to perform traffic analysis as follows:

- Divide the entire network into multiple regions geographically;
- Deploy sensors along the boundaries of each region
- To monitor the cross-component traffic;
- Treat each region as a super node and use STARS to figure out the sources, destinations, and end-to-end communication relations; and
- Analyze the traffic even when nodes are close to each other by treating the close nodes as a super node.

5.SYSTEM ARCHITECTURE

5.1 COMMUNICATION 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 head-ers 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.

5.2 ATTACK 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 communica-tions. They can monitor every single packet trans-mitted 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 commu-nication between adversaries will not influence the MANET communication.
3. The adversaries can locate the signal source according to certain properties of the detected signal , by using wireless location tracking techniques such as triangulation, nearest sensor. 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) trans-mitted by a node can always be associated with it even when the node moves from one spot to another.

6.STATISTICAL TRAFFIC PATTERN DISCOVERY SYSTEM

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).

To illustrate the basic idea of STARS, 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.

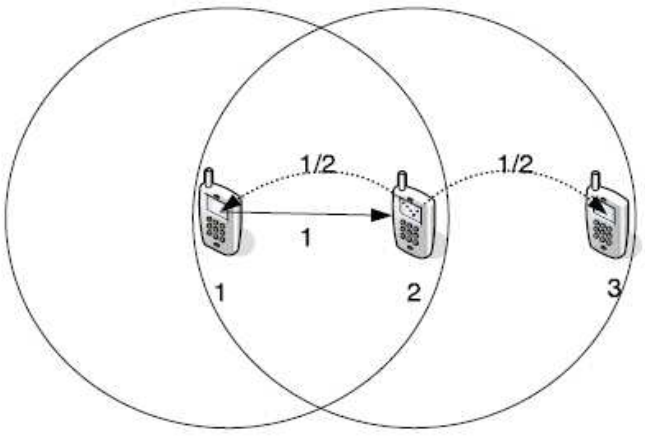

Fig.1. A simple wireless ad hoc network

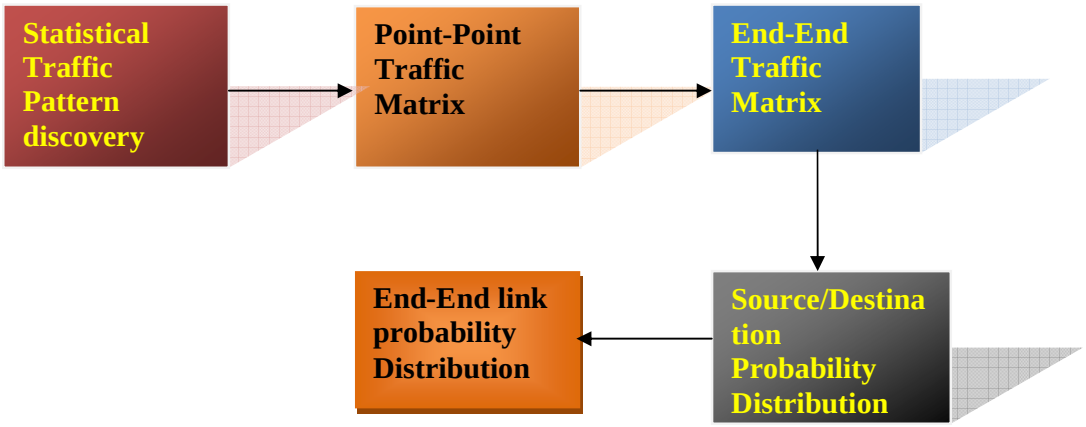

Fig.2. Block diagram of STAR with various modules

7.MODULES DESCRIPTION

7.1 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, 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. 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 Δt_e constructs a slice represented by a traffic matrix W_e, which is an N × N one-hop traffic relation matrix. The length of each time interval Δt_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.

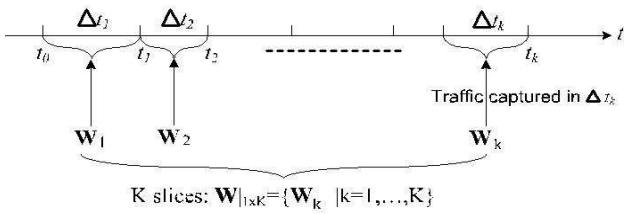

Fig.3 Slicing the domain

7.2 End-to-end traffic matrix

End to end anonymity used to redirect the packets. Given a sequence of point-to-point traffic matrices $W|_{1\times k}$ our goal is to derive the end-to-end traffic matrix R = $(r(i,j))_{N\times N}$, where $(r(i,j))$ is the accumulative traffic volume from node i to

node j, including both the point-to-point traffic captured directly and multihop traffic detect from the point-to-point traffic. In this module, we use the term accumulative traffic matrix and end-to-end traffic matrix interchangeably.

7.3 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.

7.4 Source/Destination Probability Distribution

First, derive the original end-to-end traffic matrix R from point –to-point matrix. Then we obtain the original destination probability distribution vector D from the matrix R. Then, the point-to point matrices are modified by eliminating the traffic sent by node i, and the destination probability distribution vector D_ is recomputed. Subtracting D from D results in a vector L0, which indicates the level of each node to be affected by the traffic elimination. Then the normalized vector Ls is a vector of probability for each node to be the intended destination of i. The function Suppress-Sender (i) is used to remove the traffic sent by node i. Accordingly, Suppress-Receiver (j) is used to remove the traffic received by node j

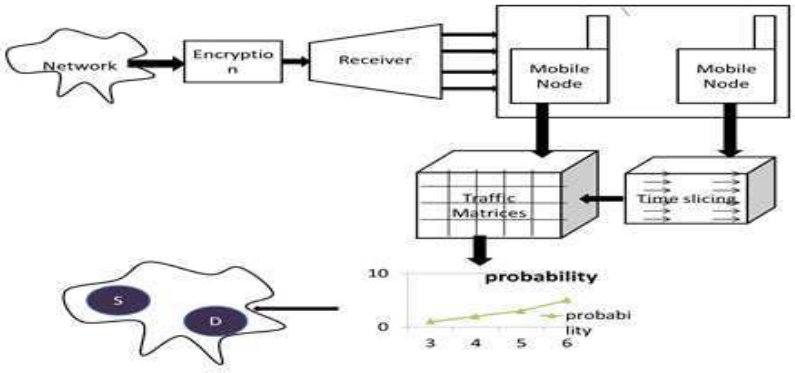

Fig .4. Working model of STAR

8.CONCLUSION

In this survey paper, we discuss the different technics of traffic pattern analysis and propose 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. Our empirical study demonstrates that the existing MANET systems can achieve very restricted communication anonymity under the attack of STARS.

Also STARS method addressed an issue that it is suitable only to small network, but not for large network. So adversaries can also take the advantages of STARS method for analysis of traffic Pattern in mobile ad-hoc networks. It is concluded with the evaluation that the hidden traffic patterns can be discovered by STARS with the good accuracy.

9.REFERENCES

[1] Yang Qin, Dijiang Huang, "STARS: A Statistical Traffic Pattern Discovery System for MANETs", IEEE Transactions On Dependable And Secure Computing, VOL. 11, NO. 2, MARCH/APRIL 2014

[2] 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, 2010.

[3] D. Huang, "Unlinkability Measure for IEEE 802.11 Based MANETs," IEEE Trans. Wireless Comm., vol. 7, no. 3, pp. 1025-1034, Mar. 2008

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

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

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

[7] M. Blaze, J. Ioannidis, A. Keromytis, T. Malkin, and A. Rubin, "WAR: Wireless Anonymous Routing," Proc. Int'l Conf. Security Protocols, pp. 218-232, 2005.

[8] S. Seys and B. Preneel, "ARM: Anonymous Routing Protocol for Mobile Ad Hoc Networks," Proc. IEEE 20th Int'l Conf. Advanced Information Networking and Applications Workshops (AINA Work-shops '06), pp. 133-137, 2006.

[9] D. Figueiredo, P. Nain, and D. Towsley, "On the Analysis of the Predecessor Attack on Anonymity Systems," technical report, Computer Science, pp. 04-65, 2004.

[10] G. Danezis, "Statistical Disclosure Attacks: Traffic Confirmation in Open Environments," Proc. Security and Privacy in the Age of Uncertainty (SEC '03), vol. 122, pp. 421-426, 2003.

[11] G. Danezis, C. Diaz, and C. Troncoso, "Two-Sided Statistical Disclosure Attack," Proc. Seventh Int'l Conf. Privacy Enhancing Technologies, pp. 30-44, 2007.

[12] C. Troncoso, B. Gierlichs, B. Preneel, and I. Verbauwhede, "Perfect Matching Disclosure Attacks," Proc. Eighth Int'l Symp. Privacy Enhancing Technologies, pp. 2-23, 2008.

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

[14] M. Reed, P. Syverson, and D. Goldschlag, "Anonymous Connec-tions and Onion Routing," IEEE J. Selected Areas in Comm., vol. 16, no. 4, pp. 482-494, May 2002.