International Journal of Engineering Research and General Science Volume 3, Issue 5, September-October, 2015
ISSN 2091-2730

A Passive Traffic Pattern Discovery Attack to reveal Anonymity of MANET Communication

Gayatri K.A., Aravind S.

Sree Buddha College of Engineering for Women, gaya3.rajasekharan@gmail.com

Abstract— Communication anonymity includes sender anonymity, recipient anonymity and channel anonymity. To break the anonymity in communication traffic analysis is performed. Traffic detection in MANETs is difficult compared to traditional wired infrastructure due to the mobility of the nodes and the lack of a fixed infrastructure. Passive statistical analysis utilize the statistical properties of the captured traffic to reveal the identity of the sender, recipient and the end-to-end link. We analyze Statistical Traffic Pattern Discovery System(STARS) with empirical results.

Keywords— - Anonymity, fpr, fnr, MANET, passive analysis, RREQ, RREP

INTRODUCTION

Secure communication should be built with the pillars of confidentiality, integrity and availability . For applications which require an additional layer of confidentiality (eg; military applications) , the communication should be untraceable or anonymous. Communication anonymity includes sender anonymity, recipient anonymity and link anonymity. The anonymity in communication is defined as an important security property by G.Danezis in [3]. Mobile Adhoc Networks(MANETs) are a set of nodes that form a network dynamically so that any node can join or leave the network any time. The typical features of MANET adhoc nature, ease of deployment, lightweight, high mobility etc. makes it suitable for implementing sensitive communication applications.

Various techniques are used to enhance the anonymity of MANET communication. Anonymous networking techniques like data encryption, encryption of packet headers at different layers using different encryptions function help to protect the traffic content .Using multiplexed traffic and introducing dummy packets or dummy delay provide difficulty for the attacker to analyze the traffic. Anonymous routing protocols hide node identities, relationship between nodes(source/ destination/ neighbouring/ forwarding node) and other routing information using techniques like dynamic pseudonyms, mixing, per-hop encryption,or timing perturberation etc. ANODR(ANonymous On Demand Routing), MASK and SDAR are examples of routing protocols which provide identity-free and on-demand routing, help to protect the anonymity in a mobile environment. Onion-routing, mix-net and DC-net are examples of anonymous communication systems.

Statistical traffic analysis attack discovers the sensitive information by evaluating the statistical characteristics of the captured raw traffic. Predecessor attack and statistical disclosure attacks are examples. But these attacks are suitable for static wired networks. In the case of infrastructure-less MANETS, traffic analysis is difficult because of the three adherent features of MANETS-the broadcasting nature, the adhoc nature and the mobile nature. STARS[1] proposed by Yang Qin et al. is a typical example of such a statistical traffic pattern analysis attack, designed for MANETs. The analysis takes into consideration the broadcasting nature, the adhoc nature and the mobility of the nodes, which are the three special characteristics of MANETs. The communication between the adversary sensors takes place through a separate channel. Thus the signal detection occurs passively, without intervening the actual channel. The adversaries can locate the signal source according to some properties and that they can trace the mobility of the nodes. Traffic matrices are constructed and probability distribution of source, destination and end-to-end links is derived using a heuristic approach. In order to speculate the actual traffic patterns from the probability distributions, the system performance is evaluated in terms of false positive rate (fpr) and false negative rate (fnr).

Extending the work in STARS, the scope of the work includes performance analysis of STARS in terms of average delay, packet drop and packet throughput Simulations were done using NS-2 platform and the result was analyzed. The rest of the paper is organized as follows: Section II describes the related work in the area; section III describes the system architecture; section IV describes the experiments and section V is the conclusion.

RELATED WORKS

The Dining Cyptographer's Network(DC-Net) by Chaum in [1] is one of the early approaches in preserving anonymity of communication. Here one participant among a group of communicating nodes broadcasts a message. The sender encrypts the message and since it is received by all the nodes in the network, recipient anonymity is maintained. Chaum in [5] introduced the concept of Mix-Node, a node is capable of re-arranging the messages that comes in a random order so that it is impossible to correlate between the input and output messages of the node. A network with all the participants are mix-nodes is called a Mix-Network or simply mix-net.

ANODR (ANonymous On Demand Routing) was devised by Kong et al. in [6]. ANODR is a hybrid protocol which uses identity-free routing and on-demand routing as the design principles. The on-demand approach ensures that anonymous routes are set up in real-time as needed, which limits the chance of traffic analyzing to a time-critical control window. Instead of using node identities, ANODR uses one-time cryptographic trap doors to hide node identities, which satisfies the identity-free criterion. MASK[7] is an anonymous on-demand routing protocol, which can accomplish both MAC-layer and network-layer communications without disclosing real IDs of the participating nodes under a rather strong adversary model. MASK offers the anonymity of senders, receivers, and sender-receiver relationships in addition to node unlocatability and untrackability and end-to-end flow untraceability. But MASK is vulnerable to denial-of service attack. It can provide security only against external adversaries. Once becoming internal adversary by compromising certain nodes, it is easy to launch an attack. In [8], Boukerche et al. describes a protocol named SDAR (Secure Distributed Anonymous Routing Protocol) for Wireless and Mobile Ad Hoc Networks. The protocol encrypts routing packet header and abstains from using unreliable intermediate node for preserving anonymity of the established route. The entire process is divided into the path discovery phase, the path reverse phase and the data transfer phase. During the path discovery phase, distributed information gathering about intermediate nodes that can be used along with the anonymous path takes place. Path reverse phase consists of conveying this information to the source node. During the data transfer phase, official data exchange takes place. SDAR provides prevention against active attacks and passive attacks that exploit path discovery path reverse messages. Reed et al. in [9] describes Onion Routing which is an infrastructure to protect anonymity in public networks. An onion is a multi-layered data structure that encapsulates the route of the anonymous connection starting from the onion router for the exit funnel and working backward to the onion router at the entry funnel. The system provides anonymity against eavesdropping and traffic analysis attacks. The authors themselves are stating that 'the implementation of a secure design can be insecure'. Traffic analysis becomes easy if part of the onion network is taken down.

A comprehensive listing of various attacks against mix-nets is provided by Raymond in [10]. In a brute force attack, the attacker follows every possible paths that the message could have taken. The attacker can create a list of possible adversaries and if the network is not well designed, he can easily track the sender and the receiver. The node flushing attack,if the nodes have to wait until they have t messages, before flushing,the attacker can send t-1 messages and easily associate messages leaving the node with those having entered. The route timing information is exploited in timing attacks,i.e., if different routes take different amount of time, the messages in the incoming and outgoing sets of a network can be correlated. Contextual attacks are targeted against real-time interactive communications. In communication pattern attack, the attacker observes the communication pattern over a period of time , making use of the fact that the communicating participants do not talk at the same time. An adversary can count the unusual number of packets sent from a participant and devise a packet counting attack. In intersection attack, attacker having information about what users are active at any given time can, through repeated observations, determine what users communicate with each other. Sender-receiver matching information can be gained by exploiting the fact that user behavior depends on the message received. The nodes not expecting to receive this message will react differently with respect to the nodes expecting the message. In a sting attack, the recipient tries to find the sender's identity and in a "send n' seek" attack, the sender tries to find the recipient's identity.

Distinguished from the above mentioned attacks, the statistical traffic analysis intends to break the anonymity by analyzing the statistical characteristics of the traffic. The predecessor attack and statistical disclosure attack are examples. Reiter and Rubin first described the predecessor attack in [11]. In this attack, the attacker tracks an identifiable stream of communications over a number of rounds. In each round, the attacker simply logs any node that sends a message that is part of the tracked stream. The attack does not always require analysis of the timing or size of packets (although that can speed up the attack), but instead exploits the process of path initialization. The statistical disclosure attack was described in [12] by G.Danezis. The attacker can identify all possible recipients of a message initiated by a particular sender node under this type of attack. This is possible if the attacker has information about the recipient anonymity set, the batch size of the mix and the probability distribution used by all other senders to select their recipients for each round of mixing and the number of observation. An evidence-based statistical traffic analysis was proposed in [2] by D.Huang.

In evidence based statistical traffic analysis each data packets are captured which are considered as evidence that support a point to point transmission between sender and receiver. In this analysis first create a sequence of point-to-point matrices, and then using that matrices derive end-to-end relations between the communication paths. This method fails when deriving the multi-hop traffic from the one hop evidences. This approach does not provide any method to detect the actual source and destination. It utilizes a nave accumulative traffic ratio to detect the multi hop communication which leads a lot of inaccuracies in the derived probability distributions.

STARS(Statistical Traffic Pattern Discovery System) was proposed by Yang Qin et al. in [4]. The analysis takes into consideration the broadcasting nature, the adhoc nature and the mobility of the nodes, which are the three special characteristics of MANETs. The attack model assumes the adversaries as passive signal detectors,who are connected through an additional channel which is different from the one used by the target MANET,the adversaries can locate the signal source according to some properties and that they can trace the mobility of the nodes. The source/destination probability distribution and the end-to-end link probability distribution are derived. A sequence of point-to-point traffic matrices are constructed from which end-to-end traffic matrices are derived in the first step. During the second step, a heuristic approach is used to identify the actual source/destination and then correlate the source node with the corresponding destination.

THE PASSIVE TRAFFIC DISCOVERY SYSTEM

Assumptions about the Network

The attacker nodes in a passive traffic analysis system do not directly involve in the communication that is flowing through the network. Their goal is to detect the traffic and to figure out the source, destination and link. These nodes make use of wireless location tracking techniques to find out the source of the detected signal. This demands that the targeted network should have limited node density, otherwise the source could not be correctly located from the set of close nodes. A separate channel is used by the adversary nodes for their communication. Encrypted packets having unique size are send by all the noble nodes through the channel so that the attacker cannot decrypt the content nor determine the source with the size of the packet. The mobility of the nodes is traced by the attacker using sensors. The system uses AODV routing protocol with random way-point mobility model. The physical/ MAC layer is controlled by IEEE 802.11(a/b/g) protocol. Every mobile node in the adhoc network maintains a routing table which has information about the next hop router to a destination node. A particular source node, in the absence of a valid, next hop path to the destination, initiates a Route Request Procedure. Since the message is broadcast over the network, the nodes having valid route replies with a Route Reply (RREP) message.

System Model

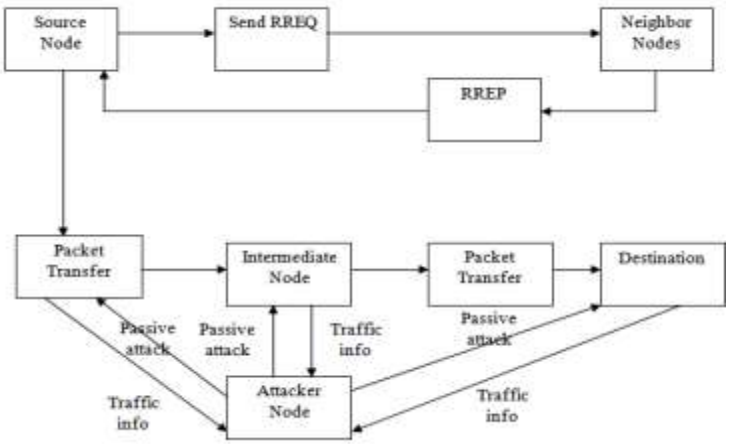

Fig.1 System Model

When the source node wishes to communicate with another, it broadcast an RREQ message in the network. When the RREP message is received from the neighboring nodes, the source node selects the path with minimum number of hops to the destination. The source node begins to send packets to the destination through intermediate nodes. The attacker nodes are deployed in a distributed manner in

the network. If an attacker node is present around an intermediate node through which the packet passes, then the attacker can detect the signal and use STARS to find out the source, destination and routing path. The probabilistic approach used in STARS, after constructing a sequence of traffic matrices yields more or less a complete attacking system. Now we analyze the system in terms of empirical parameters.

EXPERIMENTS

Experiments were done using NS-2 simulation tool with tcl coding in the front-end and C++ coding in the back end. The scope of the work is limited to finding the source probability distribution based on STARS and then evaluating the system performance in terms of average delay, packet drop and packet throughput .

Demonstration

We create a network consisting of a set of mobile nodes, deployed in 800 x 800 m^2 area. The number of nodes and the number of sources among them can be fixed by the user. One of the nodes is kept as the sink node and we consider that there are multiple source nodes. Fig.2 demonstrate the source probability distribution. The nodes having maximum probability are considered to be the source of the traffic. From the probability distribution, the nodes 5,18,24 and 26 have the highest probability to be the source node.

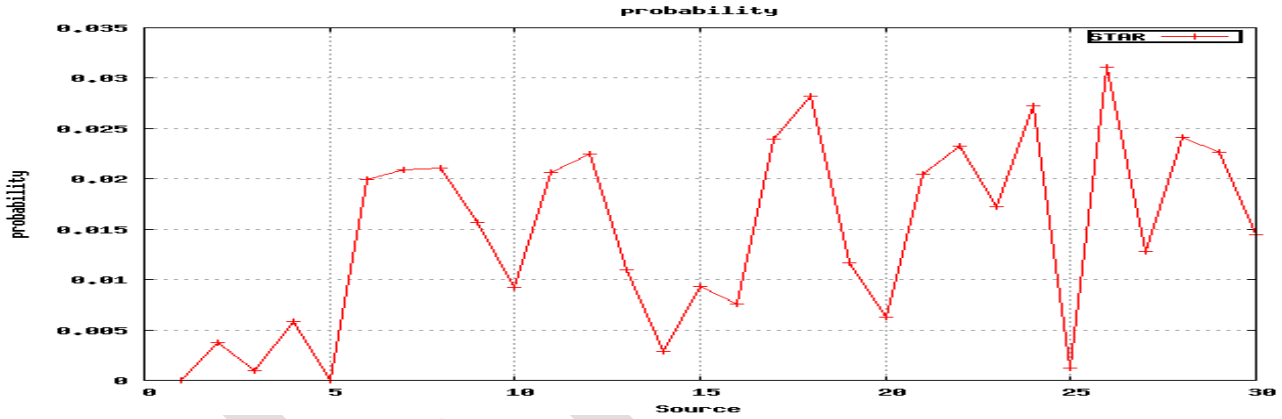

Fig.2 Source Probability Distribution

Performance evaluation

Once we execute the script, the output values are stored on to trace files. The values read from the trace file are used to plot the graph. We first analyze the system in terms of average delay in the network. Average delay is calculated as the ratio of total delay in the network to the number of packets. The graph is plotted against the number of nodes. This is depicted in Fig. 3. From the figure, it is clear that the delay increases with the node density. Fig.4 demonstrates the packet drop in the network. The packet drop retains a small value until there are about 35 nodes in the network. Thereafter the drop increases abruptly. The packet throughput of the network is depicted in Fig.5 . The throughput increases linearly until the number of nodes reaches 45, thereafter it decreases abruptly. From the performance evaluation, it is clear that when the STARS attack pattern achieves good performance when the number of nodes is limited to about 45. The system shows good performance in terms of average delay, drop and throughput until the number of nodes reaches 45.

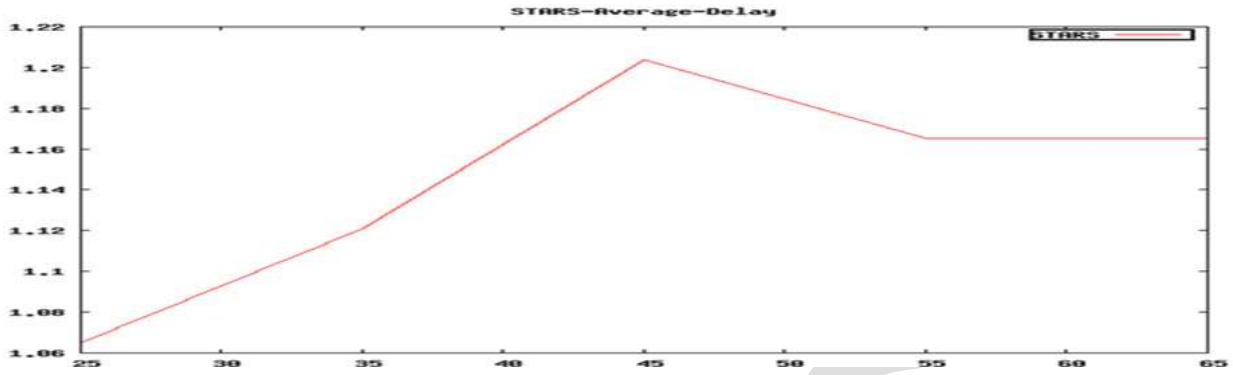

Fig.3 Average delay versus number of nodes

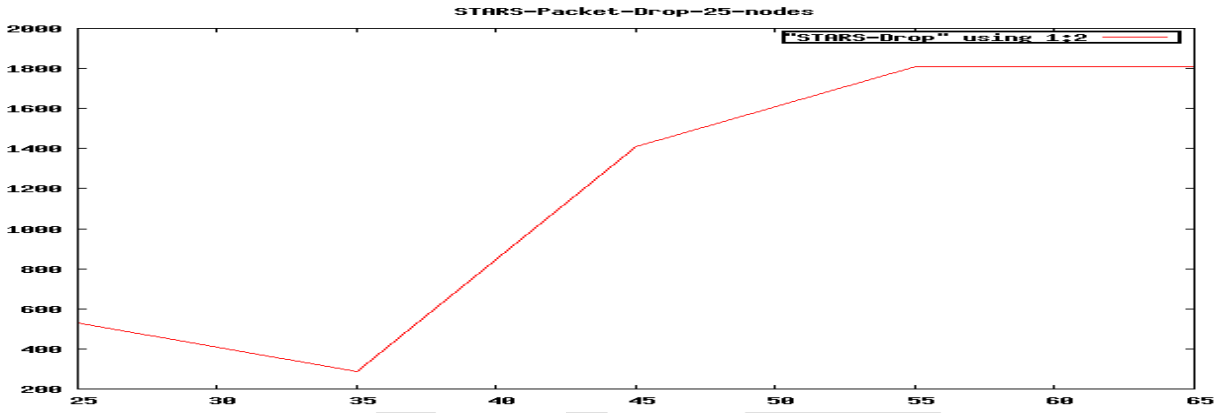

Fig.4 Packet drop versus number of nodes

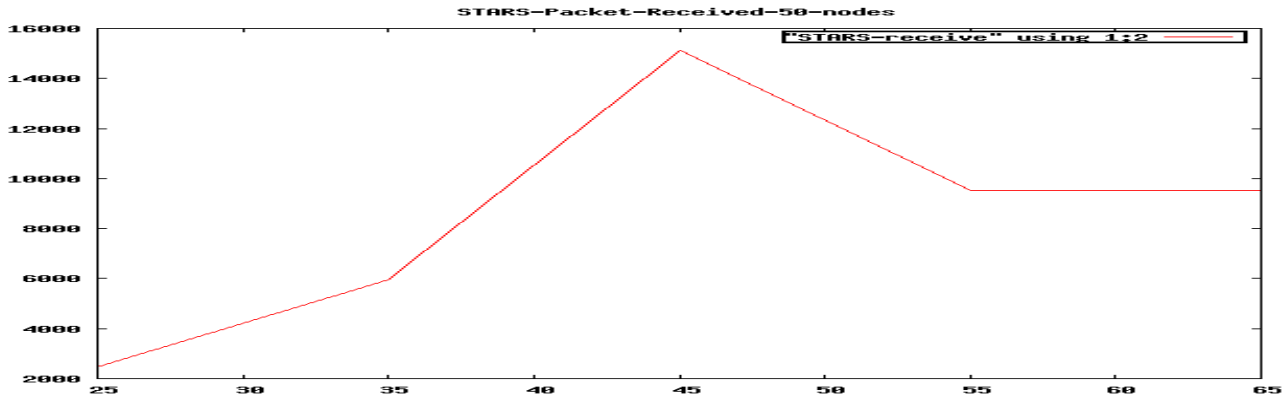

Fig.5 Packet throughput versus number of nodes

CONCLUSION

STARS attack model was analyzed in terms empirical parameters against the number of nodes in the network. From the study, it is revealed that STARS is able to find out the source of a traffic flow accurately. The system shows good performance when the number of nodes in the network is limited. But there are certain limitations of the system. Some of the routing nodes can be incorrectly determined to be the source node. Also the assumption about the adversary nodes having a global view of the network is difficult to implement as it requires deployment of large number of sensor nodes . The future work includes designing a complete attacking system addressing the above requirements.

REFERENCES:

[1] D. Chaum,The Dining Cryptographers Problem: Unconditional Sender and Recipient Untraceability, J. Cryptology, vol. 1, no. 1,pp. 65-75, 1988.

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

[3] G. Danezis, "Better Anonymous Communications," PhD thesis, University of Cambridge, January 2004.

[4] Yang Qin, Dijiang Huang and Bing Li, STARS: A Statistical Traffic Pattern Discovery System for MANETs, IEEE Trans. on Dependable and Secure Computing , vol. 11,no. 2,March/April 2014.

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

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

[7] 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.

[8] A. Boukerche, K. El-Khatib, L. Xu, and L. Korba,SDAR: A Secure Distributed Anonymous Routing Protocol for Wireless and Mobile Ad Hoc Networks,Proc. IEEE 29th Ann. Intl Conf. Local Computer Networks (LCN 04), pp. 618-624, 2004.

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

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

[11] M. Reiter and A. Rubin,Crowds: Anonymity for Web Transactions, ACM Trans.Information and System Security, vol. 1, no. 1,pp. 66-92, 1998

[12] 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