Efficient and Adjustable Recipient Anonymity in Mobile Ad Hoc Networks

Reza Shokri†, Amir Nayyeri†, Nasser Yazdani†, and Panagiotis Papadimitratos‡
† Router Laboratory, ECE Department, University of Tehran, Tehran, Iran
‡ EPFL, Lausanne, Switzerland
Emails: {r.shokri, a.nayyeri}@ece.ut.ac.ir, yazdani@ut.ac.ir, panos.papadimitratos@epfl.ch

Abstract

The privacy of users of mobile devices has been at stake, with emerging systems based on the mobile ad hoc networking technology raising additional concerns. The establishment of a connection between two nodes could readily reveal information to an eavesdropper. One approach to prevent this is to provide receiver anonymity, i.e., conceal the identity of the receiver, during the establishment of a communication path. In this paper, we introduce such a scheme that improves the efficiency of anonymous discovery, balances its cost among network nodes, and can be adaptive, trading off the degree of anonymity for the receiver.

1 Introduction

The adoption of an increasingly wider range of mobile computing applications raises concerns on the privacy of the users of portable wireless devices. Eavesdropping communications of such devices, either with other devices or with the infrastructure (e.g., access points), can allow an adversary to infer private information. For example, the adversary could track communications and transactions of each node, link those over long periods of time, or associate nodes based on their communication patterns.

In case of multi-hop communication, with the nodes being the mobile ad hoc network (MANET) infrastructure, routing protocols repeatedly discovering paths to specific destinations provide the adversary with easy-to-intercept information. For example, a route request discloses readily the source (querying node) and the destination (sought node or receiver). In addition, depending on the type of the routing protocol, control packets, that is, the route request and route reply for example, may disclose the entire path. Or, the path could be disclosed by the data packet itself, if the communication path is determined by the data sender.

A number of works in the literature have looked at how to make such communication anonymous, that is, conceal the identities of the communicating nodes from an eavesdropping adversary (e.g., [1, 2, 3] and references within). Providing anonymous communication entails a number of mechanisms concealing node identities at different phases of communication. In all cases, querying the network for a specific node is an important part of the system functionality. One approach to provide anonymity is to conceal the identity of the sender. Another approach is to conceal the identity of the receiver; or, in the best case, both identities.

Concealing the identity of the receiver is important, yet it requires a cryptographic mechanism to do: the use of a trapdoor, or, in other words, a function that can be calculated efficiently in one direction only with the knowledge of a specific secret. In our context, such a trapdoor could be set so that it can be opened only by the sought destination node, while none of the other network nodes can identify that node. However, all nodes that receive a message containing such a trapdoor would expend resources attempting to open it. In the case of a reactive MANET routing protocol, including such a trapdoor value to the route request flooded in the network implies that all nodes should undertake the corresponding overhead.

In this paper, we want to address this problem: to reduce the overhead for recipient anonymity in the MANET context. To achieve this, we devise a randomized method that allows nodes to determine with a low cost calculation whether they should attempt to open the trapdoor. In our scheme, the sender selects for each discovery, that is a given destination, a different random subset of the nodes (essentially, node identifiers). It then constructs an identifier, which we denote as the Semi Destination Set (SDS) identifier, to allow nodes receiving the request to determine if they are part of SDS and thus attempt to open the trapdoor. By regulating the size of SDS, the sender can trade off anonymity for efficiency: the smaller the SDS is the lower the protection and the processing cost in the network becomes. At the same time, the random selection a different SDS essentially balances the load among network nodes.

In the rest of this short paper, we introduce our system model, we describe in further detail our scheme, and provide a brief analysis, before we conclude with a brief description of on-going work.

2 System Model

We consider an ad hoc network composed of nodes which have unique identifiers (e.g., IP addresses). Moreover, every node may have multiple pseudonyms (random identifiers) corresponding to each node communicat-

1-4244-1455-5/07/$25.00

ing with. We assume that each node V has a public and private key pair, PuK_V, PrK_V, issued by a Certification Authority (CA) and also the public key of the CA. Nodes are equipped with symmetric key hash functions and encryption/decryption algorithms. We assume that adversaries are passive: they eavesdrop and can intercept all communications, and decrypt those for which they have the necessary cryptographic keys.

3 The Proposed Scheme

In a route discovery, the request sender, S provides the identity of the sought destination (receiver) D. To conceal this, for example, IP address, we use the public key of the destination or a symmetric key shared by S and D to encrypt this information (D address). This is essentially a trapdoor that only D can open, as it possesses the corresponding private key or the symmetric key itself. As a result, among all nodes that receive the route request, only the designated destination can reply to the sender.

To reduce the overall overhead of opening the trapdoor (decrypting the corresponding request field), we devise a method that essentially induces this operation only to a subset of the network nodes. We denote this set and its identifier as the Semi Destination Set (SDS) and SDS_ID respectively. S does not actually select among the nodes but only constructs an SDS_ID which implies the identities of nodes that potentially belong to the SDS.

The sought number of nodes in SDS, as a fraction of the (potential) total of network nodes, is a protocol-selectable parameter, φ, between 0 and 1. This essentially trades off anonymity for efficiency, and can be updated and agreed upon between S and D. For example, for nodes the sender has to provide a high degree of anonymity, e.g., based on a security policy, higher values for φ are used.

To further reduce overhead, in subsequent route discoveries, the SDS_ID can be replaced with the pseudonym identifier of destination. The pseudonym is a large random number that S and D generate as its identifiers in their connections with each other and exchanges securely, as an encrypted field of D response, thus being unknown to an observer and in fact any other node. Then, only D bearing a specific pseudonym is the only one that attempts to open the trapdoor. Next, we describe in further details the construction of the SDS_ID, how its membership function works, as well as the request and reply phases in a route discovery.

SDS. To uniformly select between the nodes in the network and place them in a same group with the destination, we manipulate the IP address of the destination to form the SDS_ID. The SDS_ID is composed of two parts: the manipulated IP of the destination, which we term as the *Masked IP* (MIP), along with a threshold, termed as the *Distance Threshold* (DTH). All the IP addresses whose distance to the MIP is less than the DTH are members of SDS. The distance is calculated simply by finding the different bits between two addresses. The MIP is produced by flipping a number of bits in the the destination's IP address: S selects a number $1 \leq \alpha \leq DTH$ and flips randomly α bits in the IP address of D. Actually, it has bto be done on the host part of the IP address (bits which are 0 in the subnet mask of the network).

To calculate the DTH, we consider the factor $\varphi \leq 1$, and the random variable X uniformly distributed over the set of network's IP addresses. Based on the definition of SDS, the minimum z such that for any address a we have $Pr(DIFF(a, X) < z) \geq \varphi$ will be assigned to the DTH, where $DIFF(IP_a, IP_b)$ implies the different bits between two addresses IP_a and IP_b.

For every two S, D nodes that already performed an SDS-based discovery, a destination pseudonym will be assigned to the MIP and DTH will be set to 0.

Request. The route request message contains a $Trapdoor$ field, in addition to the SDS_ID. The $Trapdoor$ contains IP addresses of S and D to identify the IP address of source node to D and to convince D that it is the destination of the message. Moreover, SqN (sequence number) and $Nonce$ fields protect the protocol in front of replay attacks. K_{DS} is the shared key suggested by the node S for symmetric cryptographic operations for future communications. The pseudonym identifier of sender, PID_S, also is added. Finally, to authenticate the sender, a digital signature of node S (based on its private key, over the $Trapdoor$'s fields), DS_S should be put into the trapdoor. After encrypting those fields using the public key of D, the sender will broadcast the message into the network. It is worth noting that, after the first route discovery, the $Trapdoor$ will be encrypted using the shared key between S and D. The structure of $Trapdoor$ is as follows:

$$E_{PuK_D}(SqN \| Nonce \| IP_D \| IP_S \| K_{DS} \| PID_S \| DS_S).$$

Request processing and Reply. Every node that receives the request message, checks the SDS_ID; if it is in SDS, it opens the $Trapdoor$ using its private key. In the other cases in which the DTH equals to 0, the receiver node checks if the MIP matches to one of its pseudonyms. If so, the $Trapdoor$ will be opened using the shared key corresponding to that pseudonym, and thus the corresponding source or querying node S. If the node handling the request is not in SDS or cannot open the $Trapdoor$, it relays the request.

D will update its connection table using provided fields by the sender after opening the $Trapdoor$. D can communicate with S using their last pseudonyms and the symmetric key provided by S. The destination node puts the last pseudonym of the sender on the reply message accompanied with the $Response$ field that is constructed as following:

$$E_{K_{DS}}(SqN \| IP_D \| IP_S \| PID_D) \| Nonce + 1).$$

4 Analysis

In this section, we analyze the proposed method in terms of the degree of recipient anonymity the system provides and the imposed overhead on the network. N stands for the number of nodes in the network and k is the number of connections (route discoveries) that a given pair of nodes, S and D, perform over their presence in the network.

4.1 Anonymity

Information Theoretic anonymity metrics to measure the degree of anonymity were proposed in [4]. In this paper, we use Entropy [5] to measure the degree of provided anonymity in a communication system. To calculate the entropy after the attacker observes the communication, the probability to determine whether a node is the destination of a given message or not, must be calculated. Considering a message contains MIP, the percentage of nodes those are within the SDS (equals to the φ) will open the message. Clearly, for the other nodes the probability of being the destination is 0. The degree of destination anonymity in the network for only the first connection, that is, one that uses the SDS concept, denoted as d_1, is calculated as following:

$$d_1 = \frac{-\sum_{i=1}^{\varphi.N} \varphi \log_2(\varphi)}{\log_2(N)}. \quad (1)$$

For the messages related to subsequent connections, the attacker can not obtain any information about the destination IP address because the end nodes use private random one-time use pseudonym identifiers. Because all the nodes are equally probable to be the destination the degree of anonymity in the next connections, d_n, equals to 1.

As no connection between the SDS-based message and the subsequent ones can be made, the observer can not link between them. Therefore, to obtain the overall *degree of anonymity*, d, based on the average of d_1 and d_n, we calculate it as following:

$$d = \frac{d_1 + (k-1) \times d_n}{k}. \quad (2)$$

It is obvious that as the number of messages in the network increases, the proportion of first contact becomes negligible and the degree of anonymity approaches to 1 (fully anonymous).

Even if multiple SDS-based discoveries are necessary, e.g., because of long source-destination disconnections, those will infrequent. Moreover, the randomized SDS construction makes the linking of two SDS-based discoveries for the same pair of nodes hard. Due to space limitations, we will present this analysis in the full version of the paper.

4.2 Performance

In this section, we analyze approximately the overhead our method imposes to the network, taking into consideration the overhead of symmetric (O_{Sym}) and asymmetric (O_{Asym}) cryptographic operations. Since the reply message is unicasted, the overhead of our method in a route discovery (initiated by S searching for D) can be calculated as following:

$$O_{Total} = (\varphi.N + 1)O_{Asym} + [4k - 2]O_{Sym}.$$

Note that the average overhead for each connection can become lower over time, as the use of SDS-based discoveries decreases proportion to the total number of route discoveries in the network.

Considering I as the average number of nodes in the path between S and D, the overhead of using traditional broadcasting of the request with all nodes opening the trapdoor, e.g., as in [2, 3], can be calculated as:

$$O'_{Total} = k[(2N + I)O_{Asym} + 2O_{Sym}].$$

To compare the results, with k growing, we have:

$$\frac{O_{Total}}{O'_{Total}} \simeq \frac{O_{Sym}}{N \times O_{Asym}} \quad (3)$$

The main advantage of the proposed method is the weaker connection of its overhead to the size of the network.

5 On-going Work

Discovering the destination node is a significant part in an anonymous routing protocol. In this paper, we propose an efficient solution and provide a brief analysis. We are currently developing a a broader analysis, including simulation evaluations, also considering the method integrated in a system that provides anonymity protection for other parts of communications, as well as stronger adversary models.

References

[1] Y. Zhang, W. Liu, and W. Lou, *Anonymous Communications in Mobile Ad Hoc Networks*, In Proceedings of IEEE Infocom, 2005.

[2] J Kong and X Hong, *ANODR: ANonymous On Demand Routing with Untraceable Routes for Mobile Ad-hoc Networks*, In Proceedings of MOBIHOC 2003.

[3] R. Shokri, N. Yazdani, and A. Khonsari,*Chain-based Anonymous Routing for Wireless Ad Hoc Networks*, in Proceedings of IEEE CCNC 2007.

[4] Claudia Diaz, Stefaan Seys, Joris Claessens, and Bart Preneel, *Towards measuring anonymity*, In Proceedings of PET 2002.

[5] Claude Shannon, *A mathematical theory of communication*, The Bell System Technical Journal, 27:379423:623656, 1948.