Quantification of Anonymity for Mobile Ad Hoc Networks

Marie Elisabeth Gaup Moe [1,2]

Centre for Quantifiable Quality of Service in Communication Systems
Norwegian University of Science and Technology
Trondheim, Norway

Abstract

We propose a probabilistic system model for anonymous ad hoc routing protocols that takes into account the a priori knowledge of the adversary, and illustrate how the information theoretical entropy can be used for quantification of the anonymity offered by a routing protocol as the adversary captures an increasing number of nodes in the network. The proposed measurement schema is applied to ANODR and ARM routing protocols.

Keywords: Ad-hoc networks, Anonymity, Routing, Security Model

1 Introduction

There is a need to provide secure cryptographic services in dynamic network environments with untrusted parties and a changing net topology. Cryptographic services are security services implemented by cryptographic mechanisms, examples of such services are confidentiality, authenticity, integrity, privacy, accountability, accessibility and nonrepudiation. Privacy is a service that is often difficult to realize at the same time as other cryptographic services, like authenticity, accountability and nonrepudiation. Parties in dynamic networking environments like mobile ad hoc networks, where each node is acting as a combined terminal and router, would be particularly exposed to threats against their privacy since they have no control over the trustworthiness of network nodes that handle the messages sent. Appropriate privacy enhancing cryptographic mechanisms, that can be trusted to work as intended, are required to handle this problem. Privacy has become an increasing concern for users of communication services. As communication networks are becoming more complex and diverse, the trustworthiness of network nodes, like

[1] "Centre for Quantifiable Quality of Service in Communication Systems, Centre of Excellence" appointed by The Research Council of Norway, funded by the Research Council, NTNU and UNINETT. http://www.q2s.ntnu.no

[2] Email: marieeli@q2s.ntnu.no

This paper is electronically published in
Electronic Notes in Theoretical Computer Science
URL: www.elsevier.nl/locate/entcs

routers, cannot always be guaranteed. In order to properly define secure anonymous routing it is useful to have a security model that represents the system, and to have some sort of measure that can quantify the amount of anonymity offered by the protocol.

In this paper we propose a probabilistic system model for anonymous ad hoc routing protocols that takes into account the a priori knowledge of the adversary and illustrate how the information theoretical entropy measure can be used for quantification of the anonymity of the system as the adversary captures an increasing number of nodes in the network.

2 Background: Anonymity Metrics

In this section we will give a short survey of the state of art on quantification of anonymity.

2.1 Defining Anonymity

Before we can start with measuring anonymity we need to have a clear understanding of what anonymity means. We adopt the definitions by Pfitzmann and Hansen [10]:

- *Anonymity* is the state of being not identifiable within a set of subjects, the *anonymity set*.
- *Unlinkability* of two or more items within a defined system means that these items are no more and no less related than they are related concerning the a-priori knowledge.

For the ad hoc routing setting an anonymous routing protocol should ideally offer sender and recipient anonymity, meaning that the sender of a message or recipient of a message remains unidentifiable under the assumed adversary model. We also want to achieve *realationship anonymity* between the sender and the recipient of a message, so that an observer cannot determine which nodes are taking part in a specific communication flow. In other words, sender and recipient are unlinkable. Note that the unlinkability property is weaker than the anonymity property, as anonymity of both sender and recipient implies the unlinkabilty between them. For the remaining of this paper we will focus on anonymity in the context of ad hoc routing, and in particular look at sender and recipient anonymity in the route discovery part of ad hoc routing protocols.

An *identity* is defined in [10] as any subset of attributes of an individual which identifies this individual within any set of individuals. So usually there is no such thing as the identity, but several of them. The nodes in an ad hoc network could be identified in terms of a node *identifier*. A node identifier could for instance be the node's mac address or ip address, or the identity of the user controlling the node at the time. We will assume that every node in a network of N nodes has a unique authenticated node identifier N_i, $1 \leq i \leq N$. A node could also be identified by its *location*, signal positioning could easily be used to determine an approximate location of a transmitting node. To obtain *location privacy* of a sender node a packet should not reveal the number of hops it has travelled. The packet should also not

reveal how many hops it has left to traverse before arriving at the destination in order to obtain recipient location privacy.

2.2 Measuring Anonymity

The classic way of quantifying the degree of anonymity is done by simply measuring the size of the anonymity set [3]. The size of the anonymity set is intuitively an indication of the degree of anonymity, as the more members of the set of potential senders/receivers, the less is the probability that a randomly chosen member of the set was the actual sender/receiver. But we should take into account that anonymity is stronger the more evenly distributed the sending and receiving of messages by the subjects within that set is.

Reiter and Rubin [11] give a qualitative scale for degrees of anonymity ranging from *absolute privacy* to *provable exposed*:

- *absolute privacy* means that sending a message is unobservable for the attacker
- *beyond suspicion* means that even though the attacker can see evidence of a sent message, the sender appears no more likely to be the originator than any other potential sender in the system
- *probable innocence* means that to the attacker, the sender appears no more likely to be the originator than to not be the originator
- *exposed* means that the attacker can identify the sender of a message
- *provable exposed* means that the attacker can also prove the identity of the sender to others

The degree of anonymity could also be quantified in terms of the information theoretical *entropy* of the probability distribution that the attacker assigns to each possible sender as being the originator of a message, after observing the system. In a system with N users, let p_i be the probability assigned by the attacker for user i to be the sender/recipient of a message, and let X be the discrete random variable taking the possible values $x_1, x_2, \ldots, x_N$ with probabilities $p_1, p_2, \ldots, p_N$ respectively, the entropy $H(X)$ of the probability distribution can be calculated by

$$H(X) = -\sum_{i=1}^{N} p_i \log_2(p_i). \tag{1}$$

The entropy can be interpreted as the number of bits of additional information that the attacker needs in order to definitely identify a user, or as the effective decrease in uncertainty. This information-theoretic measure of anonymity was proposed independently by Diaz et al [5] and Serjantov and Danezis [12]. For quantification of the degree of anonymity Diaz et al compared the information obtained by the attacker after observing the system against the optimal situation where all users are equally likely to have sent/received the message. The degree of anonymity is denoted d and defined as

$$d = 1 - \frac{H_{\max} - H(X)}{H_{\max}} = \frac{H(X)}{H_{\max}},$$

where $H_{\max} = \log_2(N)$ is the maximum entropy for the system. This measure tells

us how evenly distributed the probabilities within the anonymity set are.

Entropy may be used as a measure of how evenly the probabilities are distributed within each distribution, but two distributions with the same entropy could still have very different qualitative anonymity. In particular the *beyond suspicion* property could be broken even with high entropy, since a distribution of high entropy does not necessarily guarantee that a particular sender or recipient does not have a much higher probability to have sent or received a message than the rest of the potential senders/receivers. Some examples of such probability distributions are given by Tóth et al [16]. To capture this they suggest to use the worst case metric *minimum entropy* $H_{\min}$, which denotes the probability of the most likely sender/receiver within the anonymity set

$$H_{\min} = -\log_2(\max_{1 \leq i \leq N}(p_i)).$$

This measure was also used by Shmatikov and Wang [15] to calculate the relationship anonymity between sender and recipient in several simulations of mix networks, where they take into account the route selection mechanisms and the distribution of message destinations.

Another problem of the entropy measure is that it does not take into account the a priori knowledge of the adversary. In an ad hoc routing setting, if we consider an adversary that has both a global and local perspective on a network, we could imagine that the adversary has some a priori knowledge of the communication patterns of network nodes, derived from traffic analysis or from application-layer contexts. The global adversary could for instance know about the frequency of route request transmissions from all nodes, which give rise to a probability distribution over the potential senders of a particular message. This a priori knowledge could then be combined with the information the adversary obtains by local observations, as suggested by Clauß and Schiffner [4]. As noted by Diaz et al [6], the problem of how to combine the entropy measures from two different sources has not yet been fully addressed. It is not necessarily true that the entropy decreases when an adversary gets access to more information in a given attack scenario. However, if we take the weighted average of all possible entropies that the adversary can obtain after observing the system, given the a priori knowledge, this entropy, defined by Shannon as the *conditional entropy*, will always be equal or less to the entropy of the a priori probability distribution.

3 Model Description

The security model for anonymous ad hoc routing introduced in this paper is a probabilistic information theoretical model based on the models for anonymity in mix-networks proposed by Diaz et al [5] and Serjantov and Danezis [12]. The novelty of our approach is that we apply the conditional entropy measure of anonymity to ad hoc networks, that we take into account the a priori knowledge of the adversary and that we quantify the amount of additional information the adversary will gain by taking over more nodes in the network.

3.1 Adversary Model

The ad hoc network consists of a collection of nodes that can come and go into the network, the nodes simultaneously act as senders, recipients and routers. An adversary model usually distinguishes between external/internal, passive/active and global/local adversaries. An external adversary can only capture the communication between nodes while the internal adversary has access to all internal information of compromised nodes. A passive adversary can only eavesdrop on the communication or read the internal information of nodes, while an active adversary may insert, delete or modify messages or alter internal information in nodes. A global adversary has full information of the network while a local adversary only controls part of the network. The most common adversary model used when analyzing the anonymity offered by ad hoc routing protocols is an external passive global adversary (an eavesdropper on the wireless communication of all nodes in the network), that possibly cooperates with one or more internal passive or active local adversaries (malicious nodes inside the network). The proposals for anonymous routing protocols by Zhang et al [17], Boukerche et al [1], Kong and Hong [9] and Seys and Preneel [14] all use variants of this adversary model.

Hu and Perrig [7] propose to characterize an adversary based on the number of nodes it owns in the network and the number of nodes it has compromised, they suggest to use the notation Active-n-m for an active adversary that has compromised n nodes and owns m nodes. We do not wish to separate between owned and compromised nodes, as we assume that a compromised node is fully controlled by the adversary. We are interested in knowing how many nodes in the network that can be overtaken by an adversary before the anonymity offered by the routing protocol gets unacceptably low. In order to achieve this we need to have a quantification of the anonymity offered by the protocol in relation to the number of compromised nodes as well as the total number of nodes in the network. We propose to use the term Passive-c/n for an adversary that is an external passive local or global adversary for the whole network, which is consisting of N nodes, of which this adversary can eavesdrop on the communication of a subset of n nodes, and that has compromised or owns c nodes inside the network, in other words the local or global external adversary cooperates with a local internal passive adversary that controls c nodes. As we are focusing on the anonymity aspects of the routing protocol we do not in our model take into account an *active* adversary that could inject, drop or modify packets in order to disturb the routing mechanisms or to launch a denial of service attack.

We assume that the adversary carries out a probabilistic attack, this means that the adversary obtains a probability distribution over the potential sender or recipient nodes in the network that could have sent or is the recipient of a particular message. Depending on the number of nodes controlled by the adversary, this probability distribution will vary. The worst case scenario is a Passive-N/N adversary, which is a rather uninteresting case since the adversary controls all nodes in the network. For the case study used in this paper we assume that the external adversary has a global view of the network. This means that the weakest adversary in our model will be a Passive-$0/N$ adversary, which is a global external adversary without any knowledge of any internal node's information.

3.2 Network Topology Model

We choose to use an analytical probabilistic model of the ad hoc network topology, because we want the measurement model to be as general as possible to be able to compare different protocols not only for specific network topologies and specific attack scenarios. An alternative to our analytical approach could be to use simulations, where the anonymity measure is calculated over many different simulated network topologies and routes. When concerning the mobility of nodes this would indeed be a better solution and will be investigated in our further work. It should be noted that the proposed measurement model is resistant on the net topology, so our approach could still be applied to other network topology models.

The analytical network topology model requires some simplifying assumptions. Inspired by the topological model used in [13] we assume that at any given time the network nodes are evenly distributed on a two-dimensional plane and that all nodes have an equal transmission range and communicate through a wireless symmetric channel. We also assume that routes follow shortest distance paths, so that a message transmitted from node N_1 to node N_2 could not have originated from a node closer to N_2 than to N_1. We refer to the *node density* ρ as the number of nodes that lie within each node's transmission range. Let c_1 be the number of nodes that are one hop away from any particular network node, c_2 denotes the number of nodes two hops away and so on. We define $c_0 = 1$, as the only node zero hops away from any node is the node itself. As the hop-count increases from $k-1$ hops to k hops, the number of nodes grows proportionally according to the number of nodes contained within the area difference of two concentric circles with radii k and $k-1$. The number of nodes k hops away from a particular node will be $c_k = (1/2)\rho(2k-1)$.

3.3 Measurement Model

When evaluating the anonymity offered by a routing protocol, we are interested in knowing how resistant the protocol is against possibly colluding malicious nodes. To achieve this we measure the anonymity in terms of entropy based on the external global view of the adversary before any nodes have been compromised, and then quantify the average gain in information of the adversary as it controls an increasing number of nodes in the network, using the conditional entropy measure and following some of the discussion about this measure by Diaz et al [6]. In the following we will only discuss sender anonymity, with minor adjustments the same reasoning can also be applied to recipient anonymity.

Let X be a discrete random variable with probability mass function $p_i = P(X = x_i)$, x_i corresponds to a node N_i in the network and p_i is the probability that N_i will be sending a message m, as viewed by the adversary before any internal nodes have been captured. Let $\mathcal{P}_0$ be the discrete a priori probability distribution with values p_i for $1 \leq i \leq N$ in a network with N nodes. This a priori probability distribution could for instance be based on traffic analysis performed by the global external adversary. The anonymity of the nodes with respect to this external adversary could be measured in terms of the entropy of $\mathcal{P}_0$, as given by Equation 1. In the case where an ad hoc routing protocol is resistant to this kind of analysis by means of extensive use of dummy traffic we could imagine that this a priori distribution is

a uniform distribution with entropy $H_{\max} = \log_2(N)$.

Assume that a node N_j is taken over by the adversary, and that this node receives the message m. With the internal information of this node the adversary could then possibly gain some new information about which node that originated m, so that the probability distribution $\mathcal{P}_0$ can be updated to $\mathcal{P}_1$. As will be illustrated by the examples later this new information can for instance be about how many hops away the message was originated. If the node internal processing of the message m reveals the number of hops it has travelled or how many hops away to the destination it has left to travel, the adversary can in the worst case locate the position of the sending or receiving node of this message, e.g. the message was originated one hop away. If the message reveals that it was generated k hops away, the size of the anonymity set for the sending node will be

$$c_k = (1/2)\rho(2k-1).$$

In our measurement model we want to combine the probabilities assigned to each node in this anonymity set with the a priori knowledge of the adversary, to form the new probability distribution $\mathcal{P}_1$.

Let Y denote the discrete random variable with probability mass function $q_k = P(Y = y_k)$, where q_k is the probability that a message m, received by the adversary node N_j, was originated at node k hops away, according to the knowledge the adversary can derive from the internal information of node N_j. Assume that there is a maximum path length λ in the ad hoc network, measured in number of hops. If we assume the local node adversary to have no a priori knowledge of the probability of other network nodes as being the originator of the received message m, the probability that m was originated at a node k hops away is given by:

$$q_k = \frac{c_k}{\sum_{i=1}^{\lambda} c_i} \tag{2}$$

The entropy $H(Y) = -\sum_{k=1}^{\lambda} q_k \log_2(q_k)$ will express the adversary's uncertainty on which node that originated the message m, viewed *locally* from node N_j, we will combine this entropy measure with the measure of the a priori *global* view of the adversary using the Shannon *conditional entropy* $H(X|Y)$. The conditional entropy is not a measure of the uncertainty of the adversary in a specific attack scenario, but rather a measure of the adversary's average uncertainty given all possible local observations:

$$\begin{aligned} H(X|Y) &= -\sum_{i,k} P(x_i, y_k) \log_2 P(x_i|y_k) \\ &= -\sum_{k} q_k \sum_{i} P(x_i|y_k) \log_2 P(x_i|y_k). \end{aligned}$$

The conditional entropy measure is the average entropy of X, given Y, weighted according to the probability of getting a particular observation y_k. Let Z denote the discrete random variable describing the conditional probability that node N_i originated a message, given the observation y_k. Thus we have that $P_k(z_i) = P(x_i|y_k)$

and

$$H(X|Y) = \sum_k q_k H_k(Z), \tag{3}$$

where $H_k(Z)$ denotes the entropy of Z, given the observation y_k. In a specific attack scenario $P_k(z_i)$ would be the probability that N_i was the sending node, derived by an adversary that has an a priori knowledge of $\mathcal{P}_0$, and that by the capturing of the message m can see that it was originated k hops away.

In the case where $\mathcal{P}_0$ is uniformly distributed, the adversary only controls one network node and the adversary can derive that the message m, received by the network node controlled by the adversary, was generated k hops away, this observation will limit the set of potential sending nodes to only the nodes that are located k hops away. In this case the entropy measure will be reduced from $H(X) = H_{\max} = \log_2(N)$ without any observations, to $H_k(Z) = \log_2(c_k)$ given this particular observation.

If an adversary controls more than one node in the network, the anonymity set of senders, given an observed message arriving at one of the adversarial nodes, could be further reduced. If the adversary controls half of the nodes in the network, that is we have a Passive-$\frac{N}{2}/N$ adversary, we could assume that on average half of the nodes in the anonymity set would be adversarial. In that case we can derive $H_k(Z) = \log_2(\frac{c_k}{2})$, and insert this into the conditional entropy measure given in Equation 3. More generally, if the adversary controls c out of N network nodes we get the measure $H_k(Z) = \log_2((1 - \frac{c}{N})c_k)$.

If the adversary has some a priori knowledge of the node's communication patterns, $\mathcal{P}_0$ will not be uniformly distributed. In this case we will have to find the value of $P_k(z_i) = P(x_i|y_k)$, which can be rewritten using Bayes' rule as

$$P(x_i|y_k) = \frac{P(y_k|x_i)P(x_i)}{\sum_{i=1}^{N} P(y_k|x_i)P(x_i)}, \tag{4}$$

where $P(y_k|x_i)$ is the probability that a node observes that a message was originated k hops away, given that node N_i generated this message. As we will see in the examples in the following section, this probability can be derived from properties of the specific routing protocol being used.

4 Examples of Measuring Anonymity

In this section we will illustrate by two examples of anonymous ad hoc routing protocols how the entropy measure can be used for quantification of the anonymity of an ad hoc routing protocol with respect to the previously described adversary model. We will first introduce the concept of *onion routing*, which is a technique used in different variations in many proposed anonymous routing protocols.

4.1 Onion routing

Onion routing is a variant of Chaum's mix-networks [3], where messages are wrapped in layers of encryption with the keys of all intermediate nodes on the route to the destination. At each node a layer of encryption is peeled off before the node forwards

the messages in random order. If for example a message m is to be sent from the node N_1 to N_4 via the intermediate nodes N_2 and N_3, the message sent to N_2 from N_1 would be

$$\{N_3, \{N_4, \{m\}_{k_4}\}_{k_3}\}_{k_2},$$

where the k_i are secret keys shared between N_1 and all the other nodes on the route. This message is called an onion, Some padding also has to be added to the onion, so that it has a constant size, otherwise the size of the onion would reveal the distance in number of hops from the sender to the recipient. The privacy of the sender and the receiver of a message relies on the fact that there should be no correspondence between incoming and outgoing messages from a node. In practice an external passive global adversary could just track the flow of messages through the network. To prevent this, an addition of dummy traffic and different mixing strategies are applied as extra measures beside the routing protocol.

Most proposed anonymous ad hoc routing protocols, e.g. ANODR [9] and ARM [14], are on demand routing protocols that use onions in some way or another. The main idea of these protocols is that the source node N_s that is to send a message to the recipient node N_r, broadcasts a route request message that contains some information that only the recipient node can recognize (typically some information encrypted with a shared key between N_s and N_r). When nodes that are on the route, but not the recipient receives this route request they either keep some state information of this route request, or they add some encrypted information to the route request, so that later when the recipient node broadcasts the route reply message they know how to process and forward this message. When the source node N_s receives the route reply from N_r it can start to send data messages along the established route.

4.2 The ANODR Protocol

The route discovery part of ANODR uses a variant of onion routing where the source node broadcasts a route request message containing the inner core of an onion, as the route reply is forwarded throughout the network each node on the route adds a layer to this onion so that when the request reaches the recipient node the onion is wrapped with layers of encryption of all the intermediate nodes on the route. When the route reply is sent from the recipient node it contains this onion, and as the route reply traverses the route back to the source every node on the route peels off one layer of encryption from the onion. The onion is padded with random bits so that its size does not reveal the number of hops from the source or recipient node, but as noted by the authors of [18], this padding only protects against external adversaries. An internal adversary controlling one of the nodes on the route will see the size of the onion and can from this knowledge deduce the number of hops away the message was originated.

When measuring the anonymity offered by the ANODR protocol in terms of the conditional entropy, assuming one compromised network node, and given the adversary's a priori knowledge $\mathcal{P}_0$, the term $P(y_k|x_i)$ in Equation 4 is equal to 1 if the node N_i is k hops away from the adversarial node receiving the message m, and equal to 0 otherwise. This means that we are simply reducing the anonymity

set to the nodes k hops away and scaling the probabilities according to the a priori probability distribution.

If more than one node is compromised we need to exclude a number of nodes from the anonymity set according to the network proportion of adversarial nodes. One way of doing this when the a priori distribution is not uniform is to weight the sending probability of each node in the anonymity set according to the proportion of adversarial nodes as well as the a priori sending probability. If the adversary controls c out of N network nodes we would then get the conditional entropy measure:

$$H(X|Y) = -\sum_k q_k \sum_i (1 - \frac{c}{N}) P(x_i|y_k) \log_2((1 - \frac{c}{N}) P(x_i|y_k)). \quad (5)$$

4.3 The ARM Protocol

The ARM protocol uses a probabilistic padding of onions and a probabilistic time-to-live scheme in the route discovery part of the protocol. The length of route request messages grows as they traverse the network, so in order to prevent the disclosure of the distance the message has travelled, the source node N_s randomly selects a padding length of the route request message according to a specific probability distribution. This means that a neighbor node of N_s can calculate the probability that N_s was the originator of this route request message. For the route reply and data messages every node on the route chooses a time-to-live value according to a specific probability distribution, similarly in this case a neighbor node can calculate the probability that this message originated from the broadcasting node. Corresponding route request and route reply messages carry the same pseudonym identifier, this allows an adversary to correlate the internal information about these particular messages in a probabilistic attack with an increasing number of malicious nodes as described in our adversary model.

If we only look at the route request messages, assume the padding length is drawn from the discrete probability distribution $\mathcal{R}$, where r_l is the probability that the padding length $l_{min} \leq l \leq l_{max}$ is chosen. A padding length of l means that the route request appears to a neighboring node of N_s to have been originated l hops from the real source node.

In our measurement model this would mean that if a node observes that according to the onion length the message was generated k hops away, the message could have been originated at a node as far as $k + l_{max}$ hops away. The probability $P(y_k|x_i)$ in Equation 4 would then be equal to r_l if the node N_i is $k + l$ hops away. The anonymity set for the possible sender nodes would also increase in size giving:

$$c_k = (1/2)\rho \sum_{i=k}^{k+l_{max}} (2i - 1)$$

So for the ARM protocol we are reducing the anonymity set to all nodes between k and $k + l_{max}$ hops away, and scaling the probabilities according to the a priori probability distribution as well as the probability distribution for the padding scheme. We can now analyze the anonymity in terms of the conditional entropy for different numbers of adversarial nodes by using Equation 5 as explained above. As an

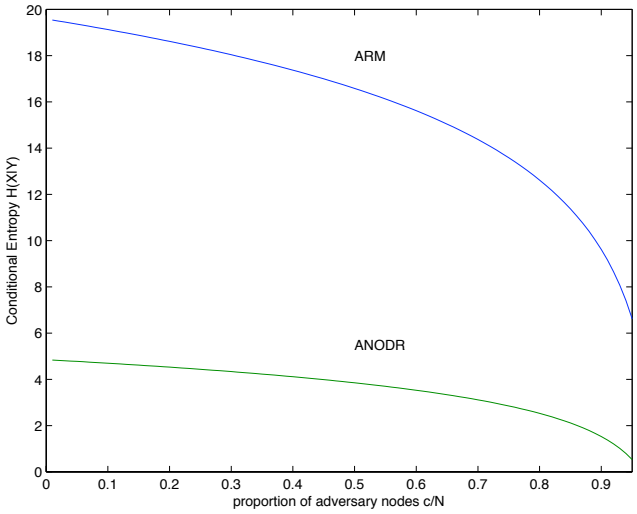

Fig. 1. Illustration of the conditional entropy measure applied to the ANODR and ARM protocols

illustration of the anonymity measure we have in Figure 1 plotted the conditional entropy measure as a function of the proportion of adversary nodes for the ANODR and ARM protocols. In our calculations we used the parameters $\rho = 8$, $\lambda = 6$ and $l_{max} = 3$, for simplicity we assumed that the distributions $\mathcal{P}_0$ and $\mathcal{R}$ were uniform.

5 Discussion and Conclusions

We have proposed a probabilistic system model for anonymous ad hoc routing protocols and showed how the information theoretical measure conditional entropy could be used for quantification of the average anonymity of the system as the adversary captures an increasing number of nodes in the network. We illustrated our approach by the examples of the ANODR and the ARM protocol, but the approach could be generally applied to ad hoc routing protocols that are using probabilistic mechanisms to achieve anonymity.

It should be noted that the weakness of the padding of onions in the ANODR protocol, allowing for an internal node to deduce the number of hops from source node, was fixed in an updated version of the protocol. To achieve secure onion routing the padding of onions should be done in such a way that a node receiving a padded onion is unable to tell if it was padded or not, the cryptographic issues involved in such a padding scheme were treated formally by Camenisch and Lysyanskaya [2]. However, in an ad hoc routing setting we need to be concerned about the efficiency of computations, so there is always a trade-off between the security and usability of a protocol, which sometimes rules out the use of provable secure but computationally heavy solutions.

There are many possible directions for further research based on this approach. When designers of a protocol want to achieve a statistical notion of anonymity, meaning that the probability of determining the sender or recipient of a message should not exceed a certain threshold, as described by [8] and [16], our approach

could possibly be used in an analysis for maximising anonymity while minimising the computational cost. We proposed an analytical model for calculating anonymity in terms of entropy, giving a weighted average entropy measure. We used a simple network topology model for our calculations, to further improve the measurement model we could in our future work replace the topology model with simulations of many different network topologies and routes, with a varying number of adversarial nodes.

References

[1] Azzedine Boukerche, Khalil El-Khatib, Li Xu, and Larry Korba. SDAR: a secure distributed anonymous routing protocol for wireless and mobile ad hoc networks. In *29th Annual IEEE International Conference on Local Computer Network, LCN'04*, pages 618–624. IEEE, 2004.

[2] Jan Camenisch and Anna Lysyanskaya. A Formal Treatment of Onion Routing. In *Advances in Cryptology, Crypto 2005*, LNCS 3621, pages 169–187. Springer, 2005.

[3] David L. Chaum. Untraceable electronic mail, return addresses, and digital pseudonyms. *Communications of the ACM*, 24(2):84–88, February 1981.

[4] Sebastian Clauß and Stefan Schiffner. Structuring anonymity metrics. In *DIM '06: Proceedings of the second ACM workshop on Digital identity management*, pages 55–62, New York, NY, USA, 2006. ACM.

[5] Claudia Diaz, Stefaan Seys, Joris Claessens, and Bart Preneel. Towards measuring anonymity. In Roger Dingledine and Paul Syverson, editors, *Proceedings of Privacy Enhancing Technologies Workshop (PET 2002)*. Springer-Verlag, LNCS 2482, April 2002.

[6] Claudia Diaz, Carmela Troncoso, and George Danezis. Does additional information always reduce anonymity? In *WPES '07: Proceedings of the 2007 ACM workshop on Privacy in electronic society*, pages 72–75, New York, NY, USA, 2007. ACM.

[7] Yih-Chun Hu and Adrian Perrig. A survey of secure wireless ad hoc routing. *IEEE Security & Privacy*, 2(3):28–39, 2004.

[8] Dogan Kesdogan, Jan Egner, and Roland Büschkes. Stop-and-go MIXes: Providing probabilistic anonymity in an open system. In *Proceedings of Information Hiding Workshop (IH 1998)*. Springer-Verlag, LNCS 1525, 1998.

[9] Jiejun Kong and Xiaoyan Hong. ANODR: ANonymous On Demand Routing with Untraceable Routes for Mobile Ad-hoc Networks. In *ACM MobiHoc'03*. ACM, 2003.

[10] Andreas Pfitzmann and Marit Hansen. Anonymity, unobservability, and pseudonymity: A consolidated proposal for terminology. Draft, December 2005.

[11] Michael Reiter and Aviel Rubin. Crowds: Anonymity for web transactions. *ACM Transactions on Information and System Security*, 1(1), June 1998.

[12] Andrei Serjantov and George Danezis. Towards an information theoretic metric for anonymity. In Roger Dingledine and Paul Syverson, editors, *Proceedings of Privacy Enhancing Technologies Workshop (PET 2002)*. Springer-Verlag, LNCS 2482, April 2002.

[13] Stefaan Seys. *Cryptographic Algorithms and Protocols for Security and Privacy in Ad Hoc Networks*. PhD thesis, Katholieke Universiteit Leuven, 2006.

[14] Stefaan Seys and Bart Preneel. ARM: Anonymous routing protocol for mobile ad hoc networks. In *Proceedings of the 20th IEEE International Conference on Advanced Information Networking and Applications - Workshops (AINA 2006)*. IEEE, 2006.

[15] Vitaly Shmatikov and Ming-Hsiu Wang. Measuring relationship anonymity in mix networks. In *WPES '06: Proceedings of the 5th ACM workshop on Privacy in electronic society*, pages 59–62, New York, NY, USA, 2006. ACM.

[16] Gergely Tóth, Zoltán Hornák, and Ferenc Vajda. Measuring anonymity revisited. In Sanna Liimatainen and Teemupekka Virtanen, editors, *Proceedings of the Ninth Nordic Workshop on Secure IT Systems*, pages 85–90, Espoo, Finland, November 2004.

[17] Yanchao Zhang, Wei Liu, and Wenjing Lou. Anonymous Communications in Mobile Ad Hoc Networks. In *IEEE INFOCOM 2005*. IEEE, 2005.

[18] Bo Zhu, Zhiguo Wan, Mohan S. Kankanhalli, Feng Bao, and Robert H. Deng. Anonymous secure routing in mobile ad-hoc networks. In *29th Annual IEEE International Conference on Local Computer Network, LCN'04*, pages 102–108. IEEE, 2004.