4-2009

Information Leakage as a Model for Quality of Anonymity Networks

Ye Zhu
Cleveland State University, y.zhu61@csuohio.edu

Riccardo Bettati
Texas A & M University - College Station, bettati@cs.tamu.edu

Publisher's Statement

Original Citation

Ye Zhu; Bettati, R.; , "Information Leakage as a Model for Quality of Anonymity Networks," Parallel and Distributed Systems, IEEE Transactions on , vol.20, no.4, pp.540-552, April 2009 doi: 10.1109/TPDS.2008.100

Repository Citation

Information Leakage as a Model for Quality of Anonymity Networks

Ye Zhu, *Member*, *IEEE Computer Society*, and Riccardo Bettati, *Member*, *IEEE Computer Society*

Abstract—Measures for anonymity in systems must be, on one hand, simple and concise and, on the other hand, reflect the realities of real systems. Such systems are heterogeneous, as are the ways they are used, the deployed anonymity measures, and finally, the possible attack methods. Implementation quality and topologies of the anonymity measures must be considered as well. We therefore propose a new measure for the anonymity degree that takes into account these various aspects of design and operation of anonymity systems. We model the effectiveness of single mixes or of mix networks in terms of information leakage, and we measure it in terms of covert-channel capacity. The relationship between the anonymity degree and information leakage is described, and an example is shown.

Index Terms—Anonymity, covert channel, information leakage.

1 Introduction

Over the recent years, we have experienced an increased perceived need for various forms of privacy-preserving communication settings, be it for pseudonymous publishing, service hiding, or anonymous communication in general. Many high-level privacy-preserving systems such as electronic cash, anonymous publishing, and others require a layer that provides some form of anonymity-preserving communication service in order to protect the identity of the participants. Anonymous communication can be provided in a number of forms: *sender anonymity* protects the identity of the sender, while *receiver anonymity* protects that of the receiver. Another form is sender-receiver anonymity (also called unlinkability), which hides whether a given sender is communicating with a given receiver. Based on these very simple measures, more sophisticated privacy-preserving schemes can be built. A large number of systems to support such privacy-preserving communication have been proposed. Most common such systems rely on so-called mix networks, which perturb the traffic in the network in order to hide the participants in a communication. The traffic is typically perturbed 1) in the payload domain, through encryption, 2) in the routing domain, through rerouting, and 3) in the timing domain, by randomly delaying packets and by generating additional dummy traffic. The objective is to prevent an observer from identifying the participants of a conversation. Examples include Crowds [1] for anonymous Web transactions, Freenet [2] for distributed anonymous information storage and retrieval, Tarzan [3] for P2P networking, and—most prominently—Tor for anonymous communication generally and for service hiding [4].

- *Y. Zhu is with the Department of Electrical and Computer Engineering, Cleveland State University, 2121 Euclid Ave., Cleveland, OH 44115-2214. E-mail: y.zhu61@csuohio.edu.*
- *R. Bettati is with the Department of Computer Science, Texas A&M University, College Station, TX 77843-3112. E-mail: bettati@cs.tamu.edu.*

In order to determine the participants in a communication, the attacker cannot simply inspect the packet or the packet header, since this information is hidden through encryption and rerouting. More sophisticated typically traffic-analysis-based techniques are required. Given the importance of effective privacy-preserving communication systems, the question of how to *quantify* the anonymity provided by an anonymity system is of great relevance. Researchers proposed various definitions to quantify anonymity, such as *anonymity set size* [5], *effective anonymity set size* [6], and entropy-based *anonymity degree* [7], all of which measure the number of possible receivers or senders of a message. While these metrics led to an increasingly better understanding of anonymity, they tend to focus on the anonymity of a *single* message under a *single* anonymity attack. In practice, however, metrics that take into account the realities of today's use of networks are needed: communication settings in real systems range from single messages to message groups, voice-over-IP streams, and FTP transfers. In addition, sophisticated attacks can resort to a variety of techniques to break anonymity: flow correlation attacks [8], intersection attacks [9], trickle attacks [10], and so on. A number of out-of-the-box attacks have been described to attack anonymity systems as well, such as Murdoch's clock-skew attack to locate hidden services [11].

A measure for the anonymity degree should satisfy a number of requirements: First, the anonymity degree should capture the *quality* of an anonymity system. It has been shown, for example, that information-theoretical means such as entropy are more accurate for comparing anonymity systems than, say, anonymity sets. Second, the anonymity degree should take into account the topology of the network or that of any overlay defined by the anonymity system. The topology influences how much information an attacker can gather and thus has an impact on the system anonymity degree. For example, a system of fully connected nodes will provide a different level of anonymity from a chain of nodes. Third, the anonymity degree, as a measure of the effectiveness of the anonymity system, should be independent of the number of users for

two reasons: 1) while a large number of users clearly contributes to anonymity, this does not necessarily reflect on the quality of the anonymity system, and 2) the effectiveness of "hiding in a crowd" is not well understood. For example, we have shown in [12] that crowds can be partitioned through appropriate preconditioning of observation data, e.g., by using Blind Source Separation.

Next, the anonymity measure must be independent of the threat model, as attackers may use a variety of attack techniques or combinations thereof to break the anonymity.

Finally, the anonymity degree should support the *engineering* of anonymity networks. As such, it should allow for *composition* of networks out of single mixes and subnetworks with well understood anonymity degrees into larger networks whose anonymity degree is well understood as well.

Since the goal of anonymity attacks is to infer the communication relations in a system despite countermeasures, it is natural to model such attacks as covert channels, and interest has focused on the interdependence of anonymity and covert channels [13]. The designer of an anonymity system generally faces the question of how much information may leak from the anonymity network given the unavoidable imperfectness of the latter and how this may affect the anonymity degree. This information leakage can be evaluated in form of a covert channel.

The major contributions of our study are summarized as follows: First, we propose an *anonymity degree* to quantify the anonymity provided *by an anonymity network.* This definition generalizes the information-theoretic definitions previously proposed in [6] and [7]. Then, we propose a new class of covert channels, which we call *anonymity-based covert channels.* We formally prove how to establish covert channels of maximum capacity over a single mix based on anonymity attacks on the mix. Finally, we use anonymity-based covert channels to assess the performance of mix networks. We show how the capacity of anonymity-based covert channels can be used to provide simple composable descriptions of nonperfect mix networks and can be used to formulate bounds on the provided anonymity.

The rest of the paper is organized as follows: Section 2 reviews the related work. Section 3 describes the proposed anonymity degree and the relationship with other entropy-based anonymity degree definitions. In Section 4, we define the anonymity-based covert channel. Sections 5, 6, and 7 present the relationship between the covert-channel capacity and anonymity degree for a single-mix case and mix-network case. Section 8 illustrates the results in the previous sections by evaluating and comparing different design decisions for a mix network. We conclude this paper and discuss the future work in Section 9.

2 Related Work

A large number of systems for anonymous communication have been proposed and developed over the last few decades, both for latency-tolerant and low-latency communication. Chaum [14] pioneered the idea of anonymity in 1981. Since then, researchers have applied the idea to different applications such as message-based e-mail and flow-based low-latency communications, and they have invented new defense techniques as more attacks have been proposed. For anonymous e-mail applications, Chaum proposed to use relay servers, called mixes, that reroute messages. Messages are encrypted to prevent their tracking by simple payload inspection.

Helsingius [15] implemented the first Internet anonymous remailer, which is a single application proxy and replaces the original e-mail's source address with the remailer's address. Gülcü and Tsudik [16] developed a relatively complete anonymous e-mail system, called Babel. Cottrell [17] developed Mixmaster, which counters a global passive attack by using message padding. It counters trickle and flood attacks [16], [10] by using a pool batching strategy. Danezis et al. [18] developed Mixminion. Mixminion's design considers a large set of attacks that researchers have found [19], [10]. The authors suggest a list of research topics for future study. Tor [4], the second-generation onion router, has been developed for circuit-based low-latency anonymous communication recently. It can provide perfect forward secrecy.

To evaluate the effectiveness of such anonymity systems under anonymity attacks, a number of different anonymity degree definitions have been proposed. The anonymity degree proposed in [1] is defined as the probability of not being identified by the attacker. It focuses on each user separately and does not capture the anonymity of the whole system. Berthold et al. [19] propose an anonymity degree based on the number of the users of an anonymity system. There is an ongoing debate about what is the role of the number of users in providing anonymity. Intuitively, the larger the crowd, the easier it is for an individual to hide in it. In practice, however, attacks proceed by isolating users or groups of users that are more likely to be participants in a communication. This was first considered in the *anonymity set*, introduced in [20]. The anonymity set describes the set of *suspected* senders or receivers of a message. The size of the anonymity set is used in [5] as the anonymity degree.

A big step forward was done by Serjantov and Danezis [6] and by Diaz et al. [7] by proposing anonymity measures that consider probability distributions in the anonymity set. Both measures are based on entropy and can differentiate two anonymity sets that have identical sizes but different distributions. The measure in [7] normalizes the anonymity degree to discount for the anonymity set size.

A number of efforts have studied the relation between covert channels and anonymity systems. Moskowitz et al. [21] focus on the covert channel over a mix-firewall between two enclaves. The covert channel in this case is established by the channel receiver determining whether an anonymized sender is transmitting packets. Newman et al. [22] focus on the covert channel over a timed mix. The authors in [13] make a series of excellent observations about the relation between covert channels and anonymity systems. They illustrate this relation by describing the linkage between the lack of complete anonymity (quasianonymity) and the covert communication over different types of mixes. Finally, they propose to use this covert-channel capacity as a metric for anonymity.

The work presented in this paper takes a system-level view of covert channels and anonymity and differs from previous work such as [13], [21], and [22] in two ways. First, we assume

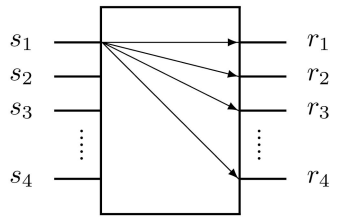

Fig. 1. Model of a mix.

that the existence of various sources of information leakage in the elements (mixes, batchers, padders, ...) of an anonymity system are a reality that system designers and operators have to deal with. Some of the resulting covert channels can be identified and either measured or analyzed using the techniques described in [21] and [22]. Statistical techniques can be used as well, as we describe in Section 3. In addition, any cautious anonymity system designer or operator must assume that mixes presumed to be perfect are not so, even if the particular weakness is not known a priori. In this paper, we use *covert-channel capacity* as a generic measure to model weaknesses (known or unknown) in the anonymity system infrastructure. This gives a tool for designers or operators to uniformly describe both known weaknesses (i.e., results of attacks) or merely suspected ones and to analyze their effect on the anonymity provided by the system. Second, the anonymity degree of the mix network is a result of system-level effects: changes in the user population or application mix affect the anonymity provided and so do topology of the anonymity system and routing preferences within the system. As a result, there is no one-to-one mapping from the anonymity degree to covert-channel capacities of elements in a mix network and vice versa. In this paper, we investigate the relationship between anonymity degree and covert-channel capacity in terms of what effect one has on the other.

3 Anonymity Degree

A number of attacks have been described recently that give rise to moderately high-capacity channels on mixes. Several attacks to simple mixes lend themselves to an accurate analysis of the exploited covert channels, such as in [13], [21], and [22]. For other attacks, the covert-channel capacity can be merely estimated, using statistical means. Examples are intersection attacks [9], timing attacks [23], Danezis's attack on the continuous mixes [24], and the flow correlation attack [8]. The timing attack [23] uses cross-correlation to match flows given the packet time stamps of the flow. Danezis's attack on the continuous mix [24] uses likelihood ratios to detect a flow in aggregate traffic. The flow correlation attack [8] employs statistical methods to detect TCP flows in aggregate traffic. The flow correlation attack can achieve high detection rates for all the mixes described in [10] and for continuous mixes.

3.1 Attack Model

We model a single mix (Fig. 1) as a communication node that connects m senders $S = (s_1, s_2, s_3, \ldots, s_m)$ to n receivers $R = (r_1, r_2, r_3, \ldots, r_n)$. Every sender s_i may communicate to every receiver r_j. We say that a *communication* exists between s_i and r_j whenever s_i communicates to r_j. A communication between s_i and r_j is denoted by the term $[s_i, r_j]$. It can consist of either a single packet being sent or of an established flow.

We model an *attack* to such a node in terms of its effectiveness in determining who is talking to whom: the set of probabilities $p([s_u, r_v]_s | [s_i, r_j]_a)$ denotes the probability that communication $[s_u, r_v]_s$ is suspected, given that communication $[s_i, r_j]_a$ is actually taking place. In other words, a probability $p([s_u, r_v]_s | [s_i, r_j]_a)$ denotes the probability of erroneously suspecting s_u sending to r_v when in actuality s_i is sending to r_j. This model lends itself to accurate descriptions of many different attacks, as the probability $p([\cdot, \cdot]_s | [\cdot, \cdot]_a)$ can be defined based on the observation of single packets, a number of packets, a flow, or a session, depending on the particular attack method used.

For example, the passive attack described in [25] successfully determines a flow when the flow is alone on a link. Therefore, the probability $p([s_i, r_j]_s | [s_i, r_j]_a)$ of correctly identifying communication $[s_i, r_j]$ is equal to the chance that the flow is alone on the output link from the mix to receiver r_j. Alternatively, for Danezis's attack on the continuous mix [24], the probability $p([s_i, r_j]_s | [s_i, r_j]_a)$ is the probability that the likelihood of the hypothesis assuming that the flow of interest is going through the link between the mix and receiver r_j is greater than any other hypothesis, assuming that the flow of interest is going to any other receiver. Finally, for the flow correlation attack [8], the probability of $p([s_i, r_j]_s | [s_i, r_j]_a)$ is equal to the probability that the mutual information between the flow of interest and the aggregate traffic on the link between the mix and receiver r_j is larger than the mutual information between the flow of interest and the aggregate traffic on any other outgoing link.

We note that the attacker may use different attack methods to estimate the probability $p([s_u, r_v]_s | [s_i, r_j]_a)$ for different communications on different mixes or even on the same mix. In addition, a priori information that the attacker may gather about potential participants can be included as well. For example, the presence of dummy traffic may indicate the existence of some anonymized traffic. Similarly, some active participants are known to have communicated in the past, which may reflect on their a priori probability that they are communicating in the present.

The model above describes attacks on sender-receiver anonymity, where both the sender and the receiver are anonymous. It can be easily extended to sender anonymity or receiver anonymity, that is, cases where the sender only or the receiver only is anonymous, respectively. For example, we can describe the results of a sender-anonymity attack in terms of $p([s_u, *]_s | [s_i, *]_a)$ or just $p([s_u]_s | [s_i]_a)$. To keep the following discussion simple and general, we will focus on sender-receiver anonymity, with the understanding that sender anonymity or receiver anonymity can be modeled just as well.

3.2 Proposed Anonymity Degree

We define a new measure D for the anonymity degree based on the following rationale: Let the random variable $[S, R]_a$ indicate the *actual* sender and receiver pair and the random variable $[S, R]_s$ in turn indicate the *suspected* sender and receiver pair. If the attack identifies the communicating pairs with high accuracy, then the dependence between the two random variables $[S, R]_a$ and $[S, R]_s$ will be high.

In general, the dependence of two random variables can be measured using the *mutual information* of the two random variables. The mutual information $I(X;Y)$ of two random variables X and Y is a function of the entropies of X and Y as follows:

$$I(X;Y) = H(X) - H(X|Y). \quad (1)$$

Therefore, the effectiveness of an attack can be described in terms of the mutual information $I([S,R]_a;[S,R]_s)$.

To give a more figurative interpretation of mutual information as a measure of the attack effectiveness, we use an analogy to communication channels: mutual information is typically used to describe the amount of information sent across a channel from a sender X to a receiver Y, where $H(X)$ is the information at the input of the channel, and $H(X|Y)$ describes the information attenuation caused by the noise on the channel. (See [26] for an easy-to-read introduction to the information theory used in this context.) This gives an intuition of why mutual information describes the effectiveness of an anonymity attack. Let $[S,R]_a$ be the random variable that describes the actual sender and receiver pair. Let an attacker's estimate of $[S,R]_a$ through observation of the system be $[S,R]_s$. The information carried through the observation channel provided by an attack is therefore $I([S,R]_a;(S,R]_s)$. The higher this carried information, the more accurate the anonymity attack.[1] Using the textbook definition for entropy, the effectiveness of an anonymity attack can be described as follows:

$$\begin{aligned} I([S,R]_a;[S,R]_s) &= H([S,R]_a) - H([S,R]_a|[S,R]_s) \\ &= \sum_{[s,r]_a,[s,r]_s} p([s,r]_a,[s,r]_s)\log\frac{p([s,r]_s|[s,r]_a)}{p([s,r]_s)}. \end{aligned} \quad (2)$$

In (2), we let $p([s,r]_a,[s,r]_s) = p([s,r]_a)p([s,r]_s|[s,r]_a)$ and $p([s,r]_s) = \sum_{[s,r]_a} p([s,r]_a,[s,r]_s)$. We let $p([s,r]_a)$ denote the a priori probability of s communicating to r, typically derived from the expected traffic from s to r.

We can now formulate the *anonymity degree* D as a function of the attack effectiveness as follows:

$$D = 1 - \frac{I([S,R]_a;[S,R]_s)}{\log(m \cdot n)}. \quad (3)$$

Since $I([S,R]_a;[S,R]_s) \le H([S,R]_a) \le \log(m \cdot n)$, we use $\log(m \cdot n)$ to normalize the anonymity degree into the range of [0, 1] in (3). Alternatively, one could choose $H([S,R]_a)$ as the normalization factor. However, the latter depends on the a priori probability of communication between each pair of sender and receiver. The impact of this a priori probability has been taken into account by the term $p([s,r]_a)$ in (2).

The equality $I([S,R]_a;[S,R]_s) = H([S,R]_a)$ holds when perfect identification of the sender-receiver pair is achieved, that is, $p([s_i,r_j]_s|[s_i,r_j]_a) = 1$ for each pair of sender and receiver. This corresponds to the situation where anonymity is totally broken, in which case the anonymity degree measure D is zero.

1. Strictly speaking, we measure the upper bound over all anonymity attacks. If an attack is biased and consistently makes wrong decisions, some other attack may make consistently better decisions.

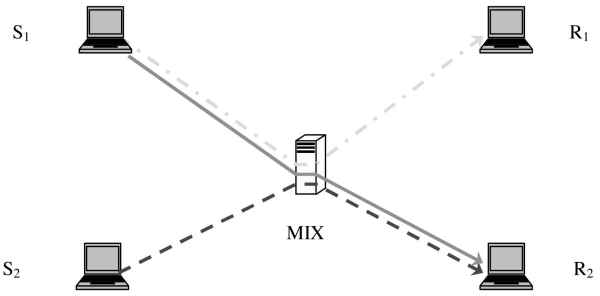

Fig. 2. Single-mix scenario.

On the other hand, the anonymity degree D is one whenever no information about sender-receiver relationships can be gathered in addition to a priori information: in this case, $I([S,R]_a;[S,R]_s) = 0$.

3.3 Relationship to Previous Anonymity Degree Definitions

The anonymity degree definition D is a generalization of the entropy-based definitions proposed in [6] and [7]. In fact, we can rewrite the attack effectiveness $I([S,R]_a;[S,R]_s)$ as

$$\begin{aligned} I([S,R]_a;[S,R]_s) &= H([S,R]_s) - H([S,R]_s|[S,R]_a) \\ &= H([S,R]_s) \\ &\quad - \sum_{[s,r]_a} p([s,r]_a) H([S,R]_s|[S,R]_a = [s,r]_a). \end{aligned} \quad (4)$$

In (4), the term $H([S,R]_s|[S,R]_a = [s,r]_a)$ represents the conditional entropy of the suspected sender-receiver pair distribution given the communication $[s,r]$. This corresponds to the anonymity degree definition described in [6] and also to the core of the anonymity degree defined in [7].

In our mutual-information-based anonymity degree, the entropy-based degree is included by averaging according to $p([s,r]_a)$, the a priori probability of traffic between each pair. In comparison to the entropy-based definitions above, our proposed definition describes the anonymity provided by a network of mixes.

4 Anonymity-Based Covert Channels

Less-than-perfect anonymity systems give rise to a form of covert channel that is exploited by anonymity attacks. We call this form of covert channel an *anonymity-based* covert channel. The input symbols of this type of covert channel are the *actual* sender-receiver pairs $[s,r]_a$, and the channel output symbols are the *suspected* sender-receiver pairs $[s,r]_s$. The channel transition probability $p([s,r]_s|[s,r]_a)$ (i.e., the probability that $[s,r]_s$ is suspected as communication given that $[s,r]_a$ is the actual communication) describes the result of the anonymity attack.

We use the simple scenario shown in Fig. 2 as an example. We assume that the attacker can collect data at the output ports of the mix as well as some additional information about incoming traffic from the senders. The details on how this information is collected and evaluated depend on the particular attack. (We described a few examples in Section 3.1.) Given sufficient collected data, the attacker can detect individual communications such as

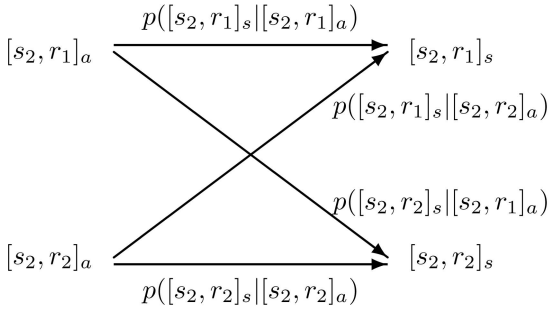

Fig. 3. Anonymity-based covert-channel model.

$[s_2,r_2]$ with some non-negligible probability, despite the anonymity-preserving countermeasures in the mix.

The fact that the attacker is able to gain information about communications indicates that a covert channel of the following form exists: A covert-channel sender can send a symbol by establishing a communication from some sender s_2 to receiver r_1 and send another symbol by establishing a communication from sender s_2 to another receiver r_2. The covert-channel receiver can use the anonymity attack to detect the flow's direction and then make the decision. The channel model is as shown in Fig. 3. For the sake of simplicity, in this example, we limit the covert-channel sender to establishing communications from sender s_2. Allowing communications from sender s_1 increases the set of input symbols accordingly.

We compute the capacity of the (anonymity-based) covert channel in a textbook fashion by maximizing the mutual information over all input symbol distributions:

$$\begin{aligned} C &= \max_{p([s_2,r]_a)} I([s_2,R]_a;[s_2,R]_s) \\ &= \max_{p([s_2,r]_a)} \sum_{i=1}^{2}\sum_{j=1}^{2} p([s_2,r_i]_a,[s_2,r_j]_s) \\ &\qquad \cdot \log \frac{p([s_2,r_j]_a,[s_2,r_i]_s)}{p([s_2,r_i]_a)p([s_2,r_j]_s)}. \end{aligned} \tag{5}$$

While the previously defined anonymity degree D only incidentally has an information-theoretic definition, the similarity between the anonymity degree D and the capacity of the anonymity-based covert channel is not accidental: The anonymity degree D describes the "expected" information leakage through an anonymity network, whereas the anonymity-based covert channel describes the leakage that is maximally achievable by carefully controlling the communication patterns through the network. The channel capacity C therefore describes a low bound on the "quality" of the anonymity network. We will describe in the following how to derive the provided anonymity degree D from a bound on C and what "quality" levels are needed to reach a predefined anonymity level D.

The covert channels previously proposed in the context of mix networks [13], [21], [22] are not anonymity based in the sense described above, as the signal is not received across the channel as a result of an anonymity attack. Rather, they describe information leakage in low-level mechanisms that are used to realize mixes, such as the batching mechanisms in [13] and [21]. These covert channels are then exploited by the anonymity attacks, which in turn can be used to establish the type of anonymity-based covert channels described in this paper.

5 Single-Mix Case

In a mix with a single sender s_1, a covert-channel sender can establish a covert channel by having s_1 communicate with any combination of j among the n receivers. For this covert channel, the set of input symbols is $\{[s_1,r_k]_a : 1 \quad k \quad n\}$, and the set of output symbols is $\{[s_u,r_v]_s : 1 \quad u \quad m, 1 \quad v \quad n\}$. We can include all communications into the set of output symbols because the *improbability* of any particular communication being declared as suspected by a particular attack can be appropriately reflected by a zero transition probability.

Therefore, $\sum_{j=1}^{n}\binom{n}{j}$ different covert channels can be established. Similarly, if the covert-channel sender has control over multiple senders, there are at least $\sum_{i=1}^{m}\binom{m}{i}\sum_{j=1}^{n}\binom{n}{j}$ different covert channels that can be established. Which of these $\sum_{i=1}^{m}\binom{m}{i}\sum_{j=1}^{n}\binom{n}{j}$ covert channels has the maximum capacity?

Lemma 1. *For a single sender s_i on a single mix, the maximum covert-channel capacity is achieved when s_i can communicate to all receivers.*

Proof. By having s_i communicate to all receivers, the covert-channel sender can send all the possible symbols $[s_i,r_j]_a$, $1 \quad j \quad n$. We call this covert channel x. By definition, the capacity of channel x is the maximal mutual information over the distributions $p([s_i,r_1]_a), p([s_i,r_2]_a), \cdots, p([s_i,r_n]_a)$, where $\sum_{j=1}^{n} p([s_i,r_j]_a) = 1$, that is

$$C_x = \max_{\substack{p([s_i,r_1]_a),p([s_i,r_2]_a),\\ \cdots,p([s_i,r_n]_a)}} I([S,R]_a;[S,R]_s). \tag{6}$$

The range of the max operator in (6) (and the value for C_x) is clearly maximized when s_i can communicate to all receivers, and therefore, none of the $p([s_i,r_j]_a)$ is zero.

Hence, the capacity of channel x communicating to all receivers is larger or equal to the capacity of all other covert channels that communicate to only a subset of receivers. □

Theorem 1. *For a single mix, the maximum covert-channel capacity is achieved when the covert-channel sender controls all the senders $s_1, s_2, \ldots, s_m$ and the input symbols of the corresponding channel include all the possible pairs $[s_i,r_j]_a$.*

The proof of Theorem 1 follows the same approach as the proof of Lemma 1.

From Theorem 1, we can derive the following corollary:

Corollary 1. *For the single-mix model shown in Fig. 1, the maximum covert-channel capacity is*

$$C = \max_{p([s,r]_a)} I([S,R]_a;[S,R]_s).$$

From Corollary 1 and (3), we derive the relationship between the *quality* of a single mix (i.e., the capacity of any covert channel that allows information to leak from the mix) and its *anonymity degree.* (Note that this relationship is

trivial for the single-mix case. However, we make use of this result in the analysis of networks of mixes.)

Lemma 2. *Given a single mix with a potential maximum information leakage that is upper bounded by C_{upper}, the anonymity degree of the single mix is lower bounded by $1 - \frac{C_{upper}}{\log(m \cdot n)}$. Similarly, given that the anonymity degree provided by a single mix is upper bounded by D_{upper}, the maximum information leakage of the mix is lower bounded by $(1 - D_{upper}) \log(m \cdot n)$.*

Proof. If the covert-channel capacity is upper bounded by C_{upper}

$$\begin{aligned} D &= 1 - \frac{I([S,R]_a;[S,R]_s)}{\log(m \cdot n)} \\ &\geq 1 - \frac{C}{\log(m \cdot n)} \\ &\geq 1 - \frac{C_{upper}}{\log(m \cdot n)}. \end{aligned}$$

If the anonymity degree is upper-bounded by D_{upper},

$$\begin{aligned} C &= max\big(I([S,R]_a;[S,R]_s)\big) \\ &\geq I([S,R]_a;[S,R]_s) \\ &= (1 - D)\log(m \cdot n) \\ &\geq (1 - D_{upper})\log(m \cdot n). \end{aligned}$$

□

Lemma 2 describes how the design and implementation quality of a mix affects effectiveness. In the following sections, we will describe this relation for the case of mix networks.

6 Mix-Network Case

6.1 Anonymity Degree of a Mix Network

We generalize the anonymity degree for a single mix defined in (3) to the network case by observing that the effectiveness of a mix network can be represented similarly to that of a "supermix." Let R_M and S_M represent the set of senders and receivers of the supermix, respectively. The anonymity degree of the supermix (and of the mix network) is

$$D = 1 - \frac{I([S_M,R_M]_a;[S_M,R_M]_s)}{\log(m \cdot n)}, \tag{7}$$

where similar to the single-mix case

$$I([S_M,R_M]_a;[S_M,R_M]_s) = \sum_{[s_i,r_j]_a,[s_u,r_v]_s} p([s_i,r_j]_a,[s_u,r_v]_s)\left(\log\frac{p([s_u,r_v]_s|[s_i,r_j]_a)}{p([s_u,r_v]_s)}\right). \tag{8}$$

$I([S_M,R_M]_a;[S_M,R_M]_s)$ is determined by $p([s_i,r_j]_a)$ and $p([s_u,r_v]_s|[s_i,r_j]_a)$, where probability $p([s_i,r_j]_a)$ is the proportion of traffic between s_i and r_j, and the probability $p([s_u,r_v]_s|[s_i,r_j]_a)$ is determined by the results of the anonymity attack at one or more mixes in the mix network. In the following sections, we describe how to make use of the single-mix attack result to describe the effectiveness of a mix network.

6.2 Effectiveness of Single Mix versus Supermix

In the following, we use the term $p_h([s_u,r_v]_s|[s_i,r_j]_a)$ to represent the transition probabilities that are the result of some anonymity attack on mix M_h and $p([s_u,r_v]_s|[s_i,r_j]_a)$ to represent the end-to-end transition probability for the supermix. Without loss of generality, we assume in the following that the supermix transition probability we are interested in is $p([s_u,r_v]_s|[s_i,r_j]_a)$. The process to determine the relationship between $p_h([s_u,r_v]_s|[s_i,r_j]_a)$ and $p([s_u,r_v]_s|[s_i,r_j]_a)$ can be divided into three steps.

Step 1. Find the set P_{uv} of all the possible paths between s_u and r_v. Clearly,

$$p\big([s_u,r_v]_s|[s_i,r_j]_a\big) = \sum_{P_a \in P_{uv}} p\Big([s_u,r_v]_{s,P_a}|[s_i,r_j]_a\Big), \tag{9}$$

where $p([s_u,r_v]_{s,P_a}|[s_i,r_j]_a)$ denotes the probability of suspecting communication $[s_i,r_j]_a$ to be communication $[s_u,r_v]_s$ *over path* P_a. Note that the actual communication between s_i and r_j takes only one path, which we call path P_0.

Step 2. Determine the probability of suspecting an actual communication over path P_0 to be the communication over another path P_a. Depending on how path P_a and path P_0 overlap, we distinguish three situations: 1) there is only one segment where the two paths overlap, 2) the two paths share multiple segments, and 3) there is no overlap between the two paths. Since there is no overlap in situation 3, the probability of suspecting a communication over path P_0 to be the communication over path P_a is zero. Hence, we only need to further pursue situations 1 and 2.

Situation 1 can be divided into four subcases.

Case 1. P_0 and P_a are identical. This implies that $s_u = s_i$ and $r_v = r_j$. In this case, the probability of suspecting correctly is the product of the probabilities of locally suspecting correctly at all mixes along path P_0. If we denote the mixes on path P_0 to be $M_1, M_2, \ldots, M_l$, then

$$\begin{aligned} p\Big([s_i,r_j]_{s,P_0}|[s_i,r_j]\Big) = {} & p_1\big([s_i,M_2]_s|[s_i,M_2]_a\big) \\ & \cdot \prod_{d=2}^{l-1} p_d\big([M_{d-1},M_{d+1}]_s|[M_{d-1},M_{d+1}]_a\big) \\ & \cdot p_l\big([M_{l-1},r_j]_s|[M_{l-1},r_j]_a\big). \end{aligned} \tag{10}$$

This follows directly from the fact that correct guesses at each mix on the path cause the attacker to correctly suspect the actual path.

Case 2. P_0 and P_a share the same path from s_i through the first mix M_1 to some mix M_l and then diverge due to an error at mix M_l. This is illustrated in Fig. 4, where in order to emphasize the path P_0 and P_a, other possible connections among the mixes and other possible mixes are ignored. The fact that P_0 and P_a share the same path from s_i means that s_i is correctly suspected, i.e., $s_u = s_i$.

In this subcase, the probability of erroneously suspecting some receiver r_v other than r_j is the result of correctly identifying the path up to some mix M_{l-1} and then making a mistake at mix M_l. Once an error has been made, the remaining mixes on the path to any erroneously suspected

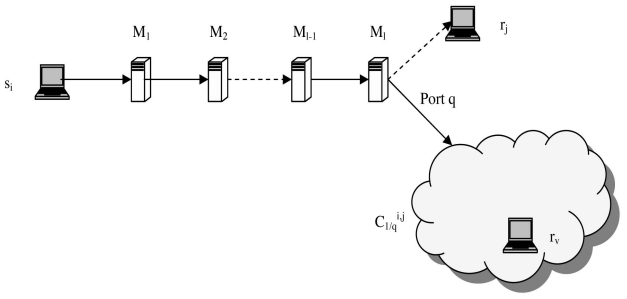

Fig. 4. Case 2.

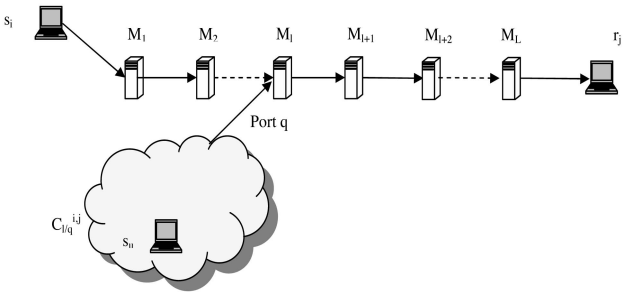

Fig. 5. Case 3.

receiver r_v are not on path P_0. According to the attack model described in Section 3, no differentiation can be made between r_v and any other receiver that can be reached after making an error at mix M_l. We therefore aggregate all receivers that can be reached after an error at mix M_l into what we call a *cloud* of receivers. We denote by $C^{ij}_{l/q}$ the cloud that is a result of an error at mix M_l, where communication $[s_i, r_j]_a$ is incorrectly identified because port q was erroneously selected instead of the port taken by $[s_i, r_j]_a$. For the example in Fig. 4, the probability of suspecting a receiver to be inside cloud $C^{ij}_{l/q}$ is

$$p\Big(\Big[s_i, C^{ij}_{l/q}\Big]_s \Big| [s_i, r_j]\Big) = p_1\big([s_i, M_2]_s | [s_i, M_2]_a\big) \cdot \prod_{d=2}^{l-1} p_d\big([M_{d-1}, M_{d+1}]_s | [M_{d-1}, M_{d+1}]_a\big) \cdot p_l\Big(\Big[M_{l-1}, C^{ij}_{l/q}\Big]_s \Big| [M_{l-1}, r_j]_a\Big). \quad (11)$$

Since we are only interested in *receivers* in the cloud, we call $C^{ij}_{l/q}$ a *receiver cloud* in this case. Whenever the context requires, we distinguish between sender clouds and receiver clouds, denoted SC and RC, respectively. We aggregate receivers into clouds because without additional evidence about the actual flow, it is impossible to differentiate suspects in a cloud by assigning different probabilities. More sophisticated anonymity attacks may make it possible to better differentiate receivers and senders in local attacks on mixes. In such a case, we would modify our detector model and extend (11) accordingly. In some cases, a cloud can consist of a single receiver or sender.

The dashed line between mix M_l and receiver r_j in Fig. 4 is to emphasize that the existence of intermediate mixes after M_l will not further contribute to suspecting communication $[s_i, r_j]_a$ as communication $[s_i, C^{ij}_{l/q}]_s$.

Case 3. P_0 and P_a share the same path from some mix M_l to the receiver. Similar to Case 2, we introduce a *sender cloud* $C^{ij}_{l/q}$, which is connected to the (input) port q of mix M_l. Since the anonymity attacks from mix M_1 to mix M_{l-1} may make a wrong decision to suspect communication $[s_i, r_j]_a$ as communications from senders attached to mixes M_1 to M_{l-1}, the probability of suspecting communication $[s_i, r_j]_a$ as communications from senders attached to the mixes after M_{l-1} will be $p_1([s_i, M_2]_s | [s_i, M_2]_a) \cdot (\prod_{d=2}^{l-1} p_d([M_{d-1},\ M_{d+1}]_s | [M_{d-1}, M_{d+1}]_a))$. Then, a wrong guess at mix M_l and correct guesses till the end of path will result in the suspected communication $[SC^{ij}_{l/q}, r_j]_s$. For the situation in Fig. 5, the probability of suspecting communication $[C^{ij}_{l/q}, r_j]_s$ is

$$p\Big(\Big[C^{ij}_{l/q}, r_j\Big]_s \Big| [s_i, r_j]\Big) = p_1\big([s_i, M_2]_s | [s_i, M_2]_a\big) \cdot \prod_{d=2}^{l-1} p_d\big([M_{d-1}, M_{d+1}]_s | [M_{d-1}, M_{d+1}]_a\big) \cdot p_l\Big(\Big[C^{ij}_{l/q}, M_{l+1}\Big]_s \Big| [M_{l-1}, M_{l+1}]_a\Big) \cdot \prod_{d=l+1}^{L-1} p_d\big([M_{d-1}, M_{d+1}]_s | [M_{d-1}, M_{d+1}]_a\big) \cdot p_L\big([M_{L-1}, r_j]_s | [M_{L-1}, r_j]_a\big). \quad (12)$$

Case 4. P_0 and P_a only share their path in middle of each path, as shown in Fig. 6.

In this case, we combine Cases 2 and 3 as follows:

$$p\Big(\Big[SC^{ij}_{l/p}, RC^{ij}_{L/q}\Big]_s \Big| [s_i, r_j]\Big) = p_1\big([s_i, M_2]_s | [s_i, M_2]_a\big) \cdot \prod_{d=2}^{l-1} p_d\big([M_{d-1}, M_{d+1}]_s | [M_{d-1}, M_{d+1}]_a\big) \cdot p_l\Big(\Big[SC^{ij}_{l/p}, M_{l+1}\Big]_s \Big| [M_{l-1}, M_{l+1}]_a\Big) \cdot \prod_{d=l+1}^{L-1} p_d\big([M_{d-1}, M_{d+1}]_s | [M_{d-1}, M_{d+1}]_a\big) \cdot p_L\Big(\Big[M_{L-1}, RC^{ij}_{L/q}\Big]_s \Big| [M_{L-1}, r_j]_a\Big). \quad (13)$$

We point out that Cases 1, 2, and 3 can all be regarded as special cases of Case 4. In Case 1, both sender cloud and receiver clouds have only one sender and one receiver, respectively. In Case 2, the sender cloud has only one sender, while in Case 3, the receiver cloud has only one receiver.

Situation 2 can have two or more overlaps between path P_0 and P_a. However, the attacker loses the ability to

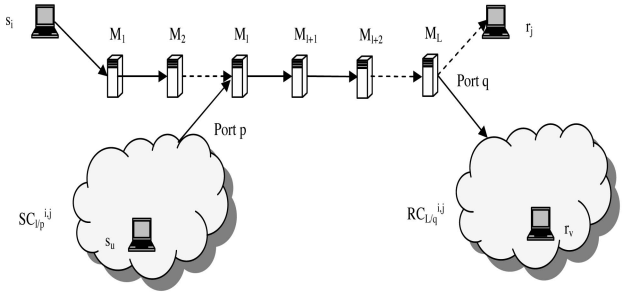

Fig. 6. Case 4.

infer anything about communication $[s_i, r_j]_a$ after the first mistake, where the two paths split. All the nodes reachable after the first mistake have to be aggregated in a receiver cloud. This situation is therefore no different than the single-overlap situation described above.

The result of Step 2 is the probability $p([SC^{ij}_{l/p}, RC^{ij}_{L/q}]_s | [s_i, r_j])$ of suspecting communication $[s_i, r_j]_a$ as communication $[SC^{ij}_{l/p}, RC^{ij}_{L/q}]_s$.

Step 3. In Steps 1 and 2, we determined path-dependent end-to-end transition probabilities of the form $p([SC^{ij}_{l/p}, RC^{ij}_{L/q}]_s | [s_i, r_j]_a)$ from the local transition probabilities at the mixes. This allows us to determine the end-to-end transition probabilities of the supermix (and —as a side result—the anonymity degree of the mix network) by solving the following optimization problem:

Given

- local transition probabilities $p_h([\cdot]_s | [\cdot]_a)$ at each mix M_h in the network,
- path-dependent transition probabilities $p([SC^{ij}_{l/p}, RC^{ij}_{L/q}]_s | [s_i, r_j]_a)$, and
- traffic volume in form of a priori probability $p([s_i, r_j]_a)$.

Objective function. Minimize the anonymity degree D in (3). This is equivalent to maximizing the mutual information $I([S, R]_a; [S, R]_s)$ in (2).

Constraints. The optimization problem is subject to the following three sets of constraints:

[**Constraint set 1**]. The sum of all path-independent transition probabilities to all the end nodes in a group of clouds is identical to the sum of path-dependent end-to-end transition probabilities to the clouds in the group. For simplicity of notation, we formulate this for the special case of a correctly suspected sender s_i. The extension to the general case is cumbersome, but straightforward. Let $GR^{i,j}_v$ be the smallest set of receiver clouds that contain r_v and all receivers in $GR^{i,j}_v$:

$$\forall r_v : \sum_{r_w \in GR^{i,j}_v} p\big([s_i, r_w]_s | [s_i, r_j]_a\big) = \sum_{RC^{i,j}_{l/q} \in GR^{i,j}_v} p\Big[s_i, RC^{i,j}_{l/q}\Big]_{s,P}\Big([s_i, r_j]_a \quad . \tag{14}$$

[**Constraint set 2**]. The sum of all path-independent transition probabilities to a subgroup of receivers is larger than the sum of the path-dependent end-to-end transition probabilities to clouds that only contain the receivers in the subgroup. This is true because one receiver in the subgroup may be contained in another cloud that contains the receivers not in the subgroup. Let R_{sub} be a subset of the set R of all receivers. Define $H^{i,j}_{R_{sub}}$ to be the set of all clouds that contain *only* receivers in R_{sub}. For the simple case of a correctly suspected sender s_i

$$\forall R_{sub} : \sum_{r_v \in R_{sub}} p\big([s_i, r_v]_s | [s_i, r_j]_a\big) \geq \sum_{RC^{i,j}_{l/q} \in H^{i,j}_{R_{sub}}} p\Big([s_i, RC^{i,j}_{l/q}]_{s,P_b} | [s_i, r_j]_a\Big). \tag{15}$$

[**Constraint set 3**]. The sum of all path-independent transition probabilities to a subgroup of receivers is less than the sum of the path-dependent end-to-end transition probabilities to those clouds that have at least one receiver in the subgroup. This holds because these clouds may have other receivers that are not in the subgroup. Let R_{sub} be a subset of the set R of all receivers. Define $I^{i,j}_{R_{sub}}$ to be the set of all clouds that contains *at least one* of the receivers in R_{sub}. We can conclude that

$$\forall R_{sub} : \sum_{r_v \in R_{sub}} p\big([s_i, r_v]_s | [s_i, r_j]_a\big) \sum_{RC^{i,j}_{l/q} \in I^{i,j}_{R_{sub}}} p\Big([s_i, RC^{i,j}_{l/q}]_{s,P_b} | [s_i, r_j]_a\Big). \tag{16}$$

[**Constraint set 4**]. The end-to-end transition probabilities for all suspects for all actual communications sum up to 1:

$$\forall i, j \sum_{s_u, r_v} p\big([s_u, r_v]_s | [s_i, r_j]_a\big) = 1. \tag{17}$$

The solution of this optimization problem is the set of the end-to-end transition probabilities of the supermix that minimize the anonymity degree of the mix network.

6.3 A Small Example

We use the example mix network displayed in Fig. 7 to illustrate how to compute end-to-end transition probabilities as described in Step 2 of Section 6.2.

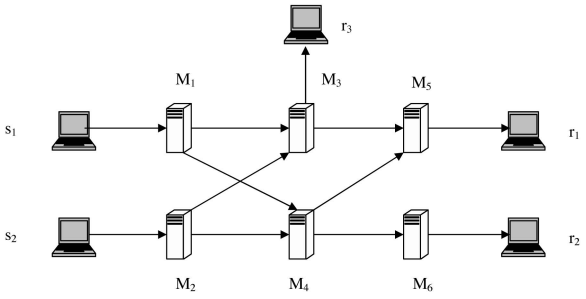

Fig. 7. A small example.

We focus on communication $[s_1, r_1]$. Suppose the actual communication takes the route $P_0 : s_1 \to M_1 \to M_3 \to M_5 \to r_1$. In this case, the probability of (erroneously) suspecting communications $[s_1, r_3]$ is computed as follows:

$$p\big([s_1, r_3]_s | [s_1, r_1]_a\big) = p_1\big([s_1, M_3]_s | [s_1, M_3]_a\big) \cdot p_3\big([M_1, r_3]_s | [M_1, M_3]_a\big). \tag{18}$$

This computation is simple, since there is only one path from s_1 to r_3.

The situation of (correctly) suspecting communication $[s_1, r_1]_a$ is more complicated, because two paths can be taken. One is $P_0 : s_1 \to M_1 \to M_3 \to M_5 \to r_1$, and the other is $P_1 : s_1 \to M_1 \to M_4 \to M_5 \to r_1$. Clearly, we have

$$p\Big([s_1, r_1]_{s,P_0} | [s_1, r_1]\Big) = p_1\big([s_1, M_3]_s | [s_1, M_3]_a\big) \cdot p_3\big([M_1, M_5]_s | [M_1, M_5]_a\big) \cdot p_5\big([M_3, r_1]_s | [M_3, M_1]_a\big) \tag{19}$$

of suspecting $[s_1, r_1]$ over path P_0.

For path P_1, we cannot get expression $p([s_1, r_1]_{s,P_1} | [s_1, r_1]_a)$ directly in terms of anonymity attack result at mixes, because

the wrong guess at mix M_1 will possibly lead to two receivers r_1 and r_2. Therefore, we have to aggregate receivers r_1 and r_2 in receiver cloud $C_{1/q}^{1,1}$, where q denotes the wrongly selected output port at mix M_1. Therefore, what we can get is

$$p\left(\left[s_1, C_{1/q}^{1,1}\right]_s | [s_1, r_1]_a\right) = p_1\left([s_1, M_4]_s | [s_1, M_1]_a\right), \quad (20)$$

where the erroneous selection of port q on mix M_1 leads to the suspected path $s_1 \rightarrow M_1 \rightarrow M_4$. Clearly, both receiver r_1 and receiver r_2 can be reached after selecting port q on mix M_1.

In turn, by following (14), we can get

$$\begin{aligned} & p\left([s_1, r_1]_s | [s_1, r_1]_a\right) + p\left([s_1, r_2]_s | [s_1, r_1]_a\right) \\ & = p_1\left([s_1, M_4]_s | [s_1, M_1]_a\right) + p_1\left([s_1, M_3]_s | [s_1, M_3]_a\right) \\ & \cdot p_3\left([M_1, M_5]_s | [M_1, M_5]_a\right) \\ & \cdot p_5\left([M_3, r_1]_s | [M_3, M_1]_a\right). \end{aligned} \quad (21)$$

After repeating this for all possible sender-receiver pairs, expressions for the end-to-end transition can be formulated, and the optimization described in Step 3 of Section 6.2 can be used to determine the anonymity degree of the network.

7 Covert-Channel Capacity versus Anonymity Degree in Mix Networks

The analysis of the effectiveness of anonymity networks is rendered difficult for two reasons primarily: First, attacks on such networks are typically out-of-the-box attacks (for example, none of the intersection attacks, trickle attacks, or others target measures taken by the mix network). Second, it is unknown where and how traffic information is collected. Is the attack targeting individual mixes or clusters of mixes? Is the information collected on a per-mix or a per-link basis?

In this section, we describe how the anonymity in mix networks can be systematically analyzed and bounded based on estimates of either per-mix weakness (using local covert channels) or the entire mix network (using network-wide covert channels). For this purpose, we investigate the relation between the covert-channel capacity of a mix network and the anonymity provided by the network.

7.1 Upper Bound on the Covert-Channel Capacity in Mix Networks

Let the mix network have K mixes. For Mix M_h, we use S_h and R_h to represent the set of senders and receivers of mix M_h, respectively. Any anonymity attack on mix M_h will lead to a set of probabilities of the form $p_h([s_u, r_v]_s | [s_i, r_j]_a)$ with s_u and s_i in S_h and r_v and r_j in R_h.

In a mix network, there are various ways to establish covert channels. For example, in the mix network shown in Fig. 8, there are at least two ways to establish the covert channels using the two mixes M_A and M_B. One way is to establish one covert channel on M_A and M_B separately. Alternatively, one can establish a covert channel on the supermix containing both M_A and M_B. We assume each mix can only be contained in one covert channel as before. In the following, we use the notation $cc(\mathcal{M})$ to denote the covert channel that can be established over the set of the mixes $\mathcal{M}$. If we denote the capacities of $cc(\{M_A\})$ and $cc(\{M_B\})$ to be C_A and C_B, respectively, then the sum of the covert-channel capacity clearly is $C_A + C_B$. We have the following lemma.

Lemma 3. *The capacity of $cc(\{M_A, M_B\})$ will be no greater than $C_A + C_B$.*

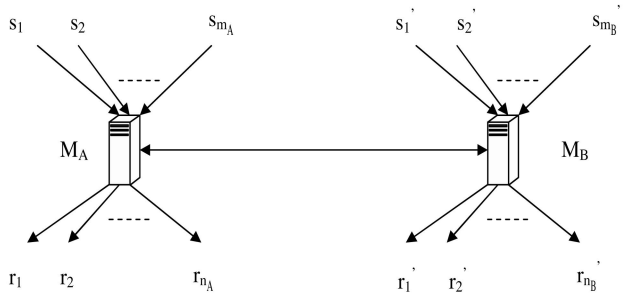

Fig. 8. Mix network of two mixes.

Proof. The input and output alphabet of $cc(\{M_A\})$ are $\{[s,r]_a : s \in S_A, r \in R_A\}$, where $S_A = \{s_1, s_2, \cdots, s_{m_A}, M_B\}$, and $R_A = \{r_1, r_2, \cdots, r_{n_A}, M_B\}$. Please note that mix M_B can be both a sender and a receiver for mix M_A and vice versa. We can construct a new channel v_1 from $cc(\{M_A\})$ with reduced set of input symbols. The input symbols of channel v_1 are $\{[s,r]_a : s \in S_A - \{M_B\}, r \in R_A\} \bigcup \{[M_B, M_B]_a\}$. According to Theorem 1, the capacity of $cc(\{M_A\})$ will be no less than the capacity of channel v_1.

Now, we consider the following covert channel v_2. The covert-channel sender of v_2 controls all the senders $s \in S_A - \{M_B\}$ attached to mix M_A to communicate with any receiver r attached to both mixes, $r \in R_A \bigcup R_B - \{M_A, M_B\}$, where $R_B = \{r'_1, r'_2, \ldots, r'_{n_B}, M_A\}$. Let I_2 denote the set $\{[s,r]_a : s \in S_A - \{M_B\}, r \in R_A \bigcup R_B - \{M_A, M_B\}\}$. Assuming the covert-channel sender can also send the symbol $[M_B, M_B]_a$, the input symbols of v_2 are $I_2 \bigcup \{[M_B, M_B]_a\}$. The receiver of the covert channel v_2 can only observe all the links connected to mix M_A. Therefore, the channel output symbols are $\{[s,r]_s : s \in S_A, r \in R_A\}$. The transition probability for channel v_2 is fully determined by the anonymity attack on mix M_A. For example, for input symbol $[s_1, r_1]_a$ and output symbol $[s_1, r_1]_s$, the transition probability is $p_{M_A}([s_1, r_1]_s | [s_1, r_1]_a)$. Please note that

$$\begin{aligned} p\left([s_1, r_x]_s | \left[s_1, r'^{\leftarrow}_i\right]_a\right) &= p\left([s_1, r_x]_s | \left[s_1, r'^{\leftarrow}_j\right]_a\right) \\ &= p_{M_A}([s_1, r_x]_s | [s_1, M_B]_a), \end{aligned} \quad (22)$$

where $r_x \in R_A$, $r'^{\leftarrow}_i \in R_B$, and $r'^{\leftarrow}_j \in R_B$.

We can observe that because of (22), we can get channel v_1 by aggregating channel v_2's input symbols $[s_x, r'^{\leftarrow}_1]_a$, $[s_x, r'^{\leftarrow}_2]_a, \ldots, [s_x, r'^{\leftarrow}_{n_B}]_a$ ($s_x \in S_A - \{M_B\}$) into $[s_x, M_B]_a$. It is obvious that

$$\sum_{i=1}^{n_B} p\left([s_x, r'^{\leftarrow}_i]_a\right) = p([s_x, M_B]_a). \quad (23)$$

The mutual information $I(X;Y)$ is a concave function of $p(x)$ for fixed $p(y|x)$. From *Jensen's inequality* [27], we can infer that the mutual information between channel v_1's input and output will be no less than the mutual information between channel v_2's input and output. Therefore, the capacity of channel v_1, denoted as C_{v_1}, is no less than the capacity of channel v_2, denoted as C_{v_2}.

Furthermore, we can extend the output symbols of channel v_2. The extension is described as follows: 1) extend

$[s_x, M_B]_s$ to $[s_x, r'_1]_a, [s_x, r'_2]_a, \ldots, [s_x, r'_{n_B}]_a$, 2) extend $[M_B, r_y]_s$ to $[s'_1, r_y]_s, [s'_2, r_y]_s, \ldots, [s'_{m_B}, r_y]_s$, and 3) extend $[M_B, M_B]_s$ to $\{[s, r] : s \in S_B - \{M_A\}, r \in R_B - \{M_A\}\}$.

We can construct channel v_3 as follows: Its input symbols are the output symbols of channel v_2, and its output symbols are the extended output symbols of channel v_2. Clearly, the transition probabilities of channel v_3 are determined by the anonymity attack on mix M_B. Thus, channel v_3's output is determined by channel v_3's input, and it is independent of channel v_2's input given channel v_3's input. Therefore, we have the *Markov Chain*: channel v_2's input $\rightarrow$ channel v_2's output, i.e., channel v_3's input $\rightarrow$ channel v_3's output.

According to the *data processing inequality* [27], the mutual information between channel v_2's input and channel v_2's output will be no less than the mutual information between channel v_2's input and channel v_3's output. We can create a channel v_4 whose input is channel v_2's input and whose output is channel v_3's output. By construction, the capacity C_{v_4} of channel v_4 will be no greater than C_{v_2}, the capacity of the channel v_2.

So far, we have

$$C_A \geq C_{v_1} \geq C_{v_2} \geq C_{v_4} \tag{24}$$

and

$$C_{v_4} = \max_{p([s_i,r_j]_a)=1} I\Big([s_i, r_j]_a; [s_u, r_v]_s\Big), \tag{25}$$

where $s_u \in S_A \bigcup S_B - \{M_A, M_B\}$, $r_v \in R_A \bigcup R_B - \{M_A, M_B\}$, and $S_B = \{s'_1, s'_2, \ldots, s'_{m_B}, M_A\}$.

Clearly, the output symbols of channel v_4 are the same as the output symbols of channel $cc(\{M_A, M_B\})$, which is built on the supermix. The input symbols of channel v_4 contain a part of the input symbols of channels $cc(\{M_A, M_B\})$ and $[M_B, M_B]_a$.

Similarly, we can get

$$C_B \geq \max_{p([s_i,r_j]_a)=1} I\Big([s_i, r_j]_a; [s_u, r_v]_s\Big), \tag{26}$$

where $I'_2 = \{[s, r]_a : s \in S_B - M_A, r \in R_A \bigcup R_B - \{M_A, M_B\}\}$, and $S_B = \{s'_1, s'_2, \cdots, s'_{m_B}, M_A\}$. The other part of the input symbols $cc(\{M_A, M_B\})$ is included in I'_2.

The capacity of channel $cc(\{M_A, M_B\})$ is

$$\begin{aligned} C_s = & \max_{p([s_l,r_j]_a)=1} I\Big([s_i, r_j]_a; [s_u, r_v]_s\Big) \\ & + \max_{p([s_i,r_j]_a)=1} I\Big([s_i, r_j]_a; [s_u, r_v]_s\Big) \end{aligned} \tag{27}$$

$$+ \max_{p([s'_i,r_j]_a)=1} I\Big([s'_i, r_j]_a; [s_u, r_v]_s\Big) \tag{28}$$

$$C_A + C_B. \tag{29}$$

The inequality between (27) and (28) holds because of two reasons: First, the maximization range comprising of $\sum_{[s_i,r_j]_a \in I_2 \bigcup \{[M_B,M_B]\}} p([s_i, r_j]_a) = 1$ and $\sum_{[s'_i,r_j]_a \in I'_2 \bigcup \{[M_A,M_A]\}} p([s'_i, r_j]_a) = 1$ includes the maximization range $\sum_{[s_l,r_j]_a \in I_2 \bigcup I'_2} p([s_l, r_j]_a) = 1$. Second, according to the *log sum inequality* [27],

$$\sum_{[s_i,r_j]_s \in O_2} p([M_B, M_B]_a, [s_i, r_j]_s) \cdot \log \frac{p([M_B, M_B]_a, [s_i, r_j]_s)}{p([M_B, M_B]_a) p([s_i, r_j]_s)} \tag{30}$$

$$\geq \sum_{[s_i,r_j]_s \in O_2} \Big(p([M_B, M_B]_a, [s_i, r_j]_s) \cdot \log \frac{\sum_{[s_i,r_j]_s \in O_2} p([M_B, M_B]_a, [s_i, r_j]_s)}{\sum_{[s_i,r_j]_s \in O_2} p([M_B, M_B]_a) p([s_i, r_j]_s)} \tag{31}$$

$$= 0, \tag{32}$$

where O_2 is the set of output symbols of channels v_4 and $cc(\{M_A, M_B\})$. Adding nonnegative terms will not change the direction of the inequality. From (28) to (29), inequalities (24) and (26) and (25) are used. □

By extending the two-mix case in Lemma 3, we obtain the following lemma:

Lemma 4. *For two mixes connected with more than one link, the capacity of the covert channel built on the supermix $cc(\{M_A, M_B\})$ will be no greater than $C_A + C_B$.*

The proof is similar to that of Lemma 3. Instead of only one path between M_A and M_B, there are several such paths. This will not affect the validity of the inequalities employed in the proof of Lemma 3.

Theorem 2. *In a mix network of K mixes, the sum of the capacities of all the covert channels in the mix network will be no greater than $\sum_{h=1}^{K} C_h$.*

Proof. This theorem can be proved by induction on K mixes with the help of Lemma 4, as any set of $K+1$ mixes can be partitioned into a supermix of K mixes and a single mix. □

7.2 Relationship

Similar to the single-mix case in Section 5, we are interested in how bounds on the achievable anonymity degree are affected by the covert-channel capacity of the system and vice versa. For example, it is obvious that an upper bound on the anonymity degree will result in a lower bound on the total covert-channel capacity, following the observation that anonymity attacks are more effective in less anonymous mixes.

The upper bound D_{upper} on the anonymity gives rise to a lower bound C_{lower} on the sum of the local channel capacities:

$$C_{lower} = min \Big(\sum_{n=1}^{K} C_h \Big). \tag{33}$$

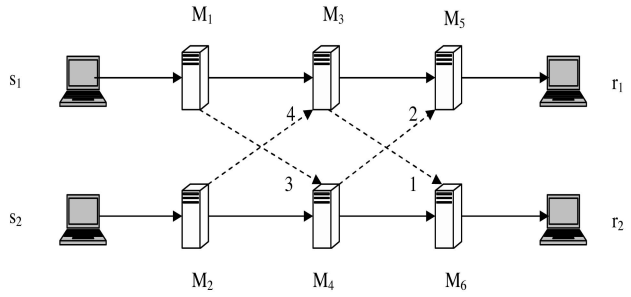

Fig. 9. An example mix network.

Equation (33) gives rise to a minimization problem over anonymity attack results $p_h([s_u, r_v]_s|[s_i, r_j]_a)$, with the following three constrains. First, the local a priori probabilities for communications at each mix M_h must sum to one:

$$\sum_{i=1}^{m_h}\sum_{j=1}^{n_h} p_h([s_i, r_j]_a) = 1. \quad (34)$$

Second, the transition probability from each input symbol $[s_i, r_j]_a$ of each mix should sum up to one:

$$\sum_{u=1}^{m_h}\sum_{v=1}^{n_h} p_h([s_u, r_v]_s|[s_i, r_j]_a) = 1. \quad (35)$$

Third, the anonymity of the system, as computed in Section 6.1, should not exceed D_{upper}.

We can solve this constrained optimization problem analytically by using Lagrange multipliers and Kuhn-Tucker conditions or by using numerical methods such as Monte Carlo.

Similarly, given upper bound C_{upper} on the total covert-channel capacity of the mix network, we would like to find out a lower bound D_{lower} for the anonymity degree of the mix network.

The objective function becomes

$$D_{lower} = \min\left[1 - \frac{I([S_M, R_M]_a; [S_M, R_M]_s)}{\log(m \cdot n)}\right]. \quad (36)$$

This optimization problem is over all possible anonymity attack results $p_h([s_u, r_v]_s|[s_i, r_j]_a)$. Constraints (34) and (35) still hold in this case. The new constraint is

$$C_{upper} \geq \sum_{h=1}^{K} C_h. \quad (37)$$

8 Evaluation

We use the mix network shown in Fig. 9 as an example to illustrate the relationships established in the previous section. We choose six mixes because it is a nontrivial topology, and both a mix cascade and a stratified network [28] can be established on the six mixes.

We assume that communications between each sender-receiver pair have the same a priori probability (alternatively, the same share of total traffic volume). Since there are two senders and two receivers, we have four sender-receiver pairs. The actual path for communication $[s_i, r_j]_a$ is shown in Table 1 if the actual path is not specified and the path is possible in the topology. For our examples, we use adaptive simulated annealing to solve the optimization problem to establish D_{lower} from a known bound on the mix network capacity.

TABLE 1
Path of the Actual Communications

Communication	Actual Path
$[s1, r1]_a$	$s_1 \rightarrow M_1 \rightarrow M_3 \rightarrow M_5 \rightarrow r_1$
$[s1, r2]_a$	$s_1 \rightarrow M_1 \rightarrow M_3 \rightarrow M_6 \rightarrow r_2$
$[s2, r1]_a$	$s_2 \rightarrow M_2 \rightarrow M_4 \rightarrow M_5 \rightarrow r_1$
$[s2, r2]_a$	$s_2 \rightarrow M_2 \rightarrow M_4 \rightarrow M_6 \rightarrow r_2$

The following measurements focus on the impact of parameters of the anonymity network. Parameters that are extraneous to the network proper (for example, the number of concurrent users) are not measured.

Impact of the connectivity. Obviously, the connectivity will affect the anonymity degree in a mix network. In our first set of examples, the base topology contains only the solid lines in Fig. 9. Then, edges are incrementally added to the base topology in the order of the label assigned to each edge. The average degrees of the topologies, including the base topology, are 2, $\frac{14}{6}$, $\frac{16}{6}$, 3, and $\frac{20}{6}$, respectively.

For every mix in the base topology, there is only one input link and one output link. Therefore, there is only one sender receiver pair for the mix in the base topology. A channel with only one input symbol and one output symbol has capacity zero. Therefore, the capacity C_{sum} is zero for the base topology.

In Fig. 10a, we first observe that the lower bound of the anonymity degree decreases with increasing bound on the capacity, just as expected. In addition, the capacity C_{sum} increases with increasing connectivity. For a given upper bound of the capacity C_{sum}, increasing connectivity will increase the anonymity degree. Third, we can observe that there is large gap between the base topology and the topology of the next higher average degree. This is because adding the edge of label 1 will connect s_1 and r_2, and communication $[s_1, r_2]_a$ can be suspected as $[s_1, r_1]_s$. Therefore, the initial edge added to the topology can increase the anonymity degree significantly. In comparison, the effect of adding an edge with label 4 is marginal.

Effect of adding different edges. In the second set of examples, we use the solid lines and edge with label 1 as the base topology. Then, we add one more edge, with label 2, 3, or 4, to the base topology. We label the new topologies as A, B, and C, respectively. Clearly, these topologies are of the same average degree. In Fig. 10b, we observe that the anonymity degree increase caused by the addition of the edge with label 3 is smaller than that for the other two edges. This is because adding the other two edges can make communication $[s_2, r_1]_a$ possible, and communication $[s_2, r_1]_a$ can be suspected as some other communication.

Effect of path selection. In this set of examples, we focus on the topology containing all the solid and dashed lines except the edge with label 3. We consider two cases. In one case, the actual path for communication $[s_2, r_1]_a$ follows path A, as in Table 1. In the other case, the actual path B for communication $[s_2, r_1]_a$ is $s_1 \rightarrow M_2 \rightarrow M_3 \rightarrow M_5 \rightarrow r_1$.

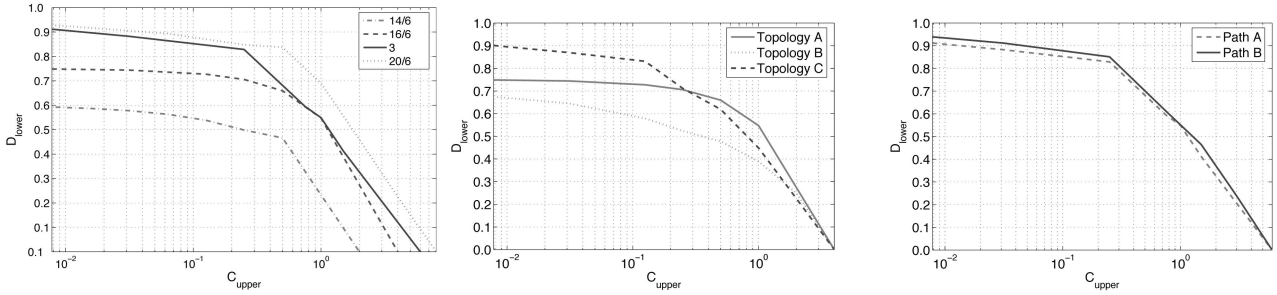

Fig. 10. Effect of topology. (a) Impact of the connectivity. (b) Effect of adding different edges. (c) Effect of path selection.

We observe that the use of mix M_3 will slightly increase the anonymity in Fig. 10c. This is because mix M_3 has more input and output links than the other mixes. Communications that use mix M_3 are thus easier to hide. In general, the rather minor effect of path selection measured in this experiment reflects the fact that the underlying mix network is rather homogeneous, in terms of both symmetry of topology and "quality" of the nodes.

The interested reader may be missing an evaluation that measures the effect of the size of the user population at this point. As we indicated in the introduction, in this paper, we attempt to capture the inherent quality of the anonymity network and leave the effect of user population size reflected in the per-mix information leakage capacity.

In practice, in order to better construct an anonymity network, first, it is important to have a rich connectivity in the network. Second, given similar connectivity in anonymity networks, the topology will affect the "quality" of the anonymity network. Finally, the effect of path selection is minor when the network is homogeneous in terms of symmetry of topology and "quality" of the nodes.

9 Summary and Future Work

Nodes in anonymity networks are inherently heterogeneous: platforms can vary. The expertise of their operators varies widely as well, and so does the exposure to misuse. Similarly, the connectivity of nodes is also not uniform. In addition, the types of attacks and their effects on the anonymity of the system cannot be predicted. This all makes it very difficult for designers of anonymity protocols and operators and users of anonymity systems to assess the effectiveness of the system. In this paper, we proposed a generic information-theoretic measure for the anonymity degree. This simple measure ranges between zero and one to indicate the overall effectiveness of the mix network as a whole. Our work is the first to develop a relationship between the anonymity degree and the capacity of anonymity-based covert channels. In the mix-network case, this relationship is described in a scenario-oriented fashion. What is needed is a set of rules to map and cluster arbitrary networks into supermixes and clouds. Further research will focus on multicast or broadcast channels in the anonymity network and their effect on the anonymity degree. Finally, we need to extend the work from anonymity-based covert channels to general covert channels in mix networks, such as the nonanonymity-based covert channels described in [13], [21], and [22] or other formalizations of information-leakage-based attacks. Eventually, a conclusion is needed that allows the aggregation attacks and the formulation of the level of anonymity provided by systems with less-than-perfect components.

Acknowledgments

This work was supported in part by the Texas Information Technology and Telecommunication Task Force (TITF).

References

[1] M. Reiter and A. Rubin, "Crowds: Anonymity for Web Transactions," *ACM Trans. Information and System Security,* vol. 1, no. 1, June 1998.

[2] I. Clarke, O. Sandberg, B. Wiley, and T.W. Hong, "Freenet: A Distributed Anonymous Information Storage and Retrieval System," *Proc. Designing Privacy Enhancing Technologies: Workshop Design Issues in Anonymity and Unobservability,* pp. 46-66, July 2000.

[3] M.J. Freedman and R. Morris, "Tarzan: A Peer-to-Peer Anonymizing Network Layer," *Proc. Ninth ACM Conf. Computer and Comm. Security (CCS '02),* Nov. 2002.

[4] R. Dingledine, N. Mathewson, and P. Syverson, "Tor: The Second-Generation Onion Router," *Proc. 13th Usenix Security Symp.,* Aug. 2004.

[5] D. Kesdogan, J. Egner, and R. Büschkes, "Stop-and-Go MIXes: Providing Probabilistic Anonymity in an Open System," *Proc. Information Hiding Workshop (IH),* 1998.

[6] A. Serjantov and G. Danezis, "Towards an Information Theoretic Metric for Anonymity," *Proc. Privacy Enhancing Technologies Workshop (PET '02),* R. Dingledine and P. Syverson, eds., Apr. 2002.

[7] C. Diaz, S. Seys, J. Claessens, and B. Preneel, "Towards Measuring Anonymity," *Proc. Privacy Enhancing Technologies Workshop (PET '02),* R. Dingledine and P. Syverson, eds., Apr. 2002.

[8] Y. Zhu, X. Fu, B. Graham, R. Bettati, and W. Zhao, "On Flow Correlation Attacks and Countermeasures in Mix Networks," *Proc. Privacy Enhancing Technologies Workshop (PET '04),* May 2004.

[9] G. Danezis and A. Serjantov, "Statistical Disclosure or Intersection Attacks on Anonymity Systems," *Proc. Sixth Information Hiding Workshop (IH '04),* May 2004.

[10] A. Serjantov, R. Dingledine, and P. Syverson, "From a Trickle to a Flood: Active Attacks on Several Mix Types," *Proc. Information Hiding Workshop (IH '02),* F. Petitcolas, ed., Oct. 2002.

[11] S.J. Murdoch, "Hot or Not: Revealing Hidden Services by Their Clock Skew," *Proc. 13th ACM Conf. Computer and Comm. Security (CCS '06),* pp. 27-36, 2006.

[12] Y. Zhu and R. Bettati, "Un-Mixing Mix Traffic," *Proc. Privacy Enhancing Technologies Workshop (PET '05),* May 2005.

[13] I.S. Moskowitz, R.E. Newman, and P.F. Syverson, "Quasi-Anonymous Channels," *Proc. IASTED Int'l Conf. Comm., Network, and Information Security (CNIS '03),* Dec. 2003.

[14] D. Chaum, "Untraceable Electronic Mail, Return Addresses, and Digital Pseudonyms," *Comm. ACM,* vol. 4, no. 2, Feb. 1981.

[15] J. Helsingius, *Press Release: Johan Helsingius Closes His Internet Remailer,* http://www.penet.fi/press-english.html, 1996.
[16] C. Gülcü and G. Tsudik, "Mixing E-Mail with Babel," *Proc. Network and Distributed Security Symp. (NDSS '96),* pp. 2-16, Feb. 1996.
[17] U. Möller and L. Cottrell, *Mixmaster Protocol—Version 2,* unfinished draft, Jan. 2000.
[18] G. Danezis, R. Dingledine, and N. Mathewson, "Mixminion: Design of a Type III Anonymous Remailer Protocol," *Proc. IEEE Symp. Security and Privacy,* May 2003.
[19] O. Berthold, A. Pfitzmann, and R. Standtke, "The Disadvantages of Free MIX Routes and How to Overcome Them," *Proc. Designing Privacy Enhancing Technologies: Workshop Design Issues in Anonymity and Unobservability,* H. Federrath, ed., pp. 30-45, July 2000.
[20] D. Chaum, "The Dining Cryptographers Problem: Unconditional Sender and Recipient Untraceability," *J. Cryptology,* vol. 1, pp. 65-75, 1988.
[21] I.S. Moskowitz, R.E. Newman, D.P. Crepeau, and A.R. Miller, "Covert Channels and Anonymizing Networks," *Proc. Workshop Privacy in the Electronic Soc. (WPES '03),* Oct. 2003.
[22] R.E. Newman, V.R. Nalla, and I.S. Moskowitz, "Anonymity and Covert Channels in Simple Timed Mix-Firewalls," *Proc. Privacy Enhancing Technologies Workshop (PET '04),* May 2004.
[23] B.N. Levine, M.K. Reiter, C. Wang, and M.K. Wright, "Timing Attacks in Low-Latency Mix-Based Systems," *Proc. Eighth Int'l Conf. Financial Cryptography (FC '04),* A. Juels, ed., Feb. 2004.
[24] G. Danezis, "The Traffic Analysis of Continuous-Time Mixes," *Proc. Privacy Enhancing Technologies Workshop (PET '04),* May 2004.
[25] A. Serjantov and P. Sewell, "Passive Attack Analysis for Connection-Based Anonymity Systems," *Proc. Ninth European Symp. Research in Computer Security (ESORICS '03),* Oct. 2003.
[26] I.S. Moskowitz and M.H. Kang, "Covert Channels—Here to Stay," *Proc. IEEE Ninth Annual Conference on Computer Assurance (COMPASS '94),* June 1994.
[27] T.M. Cover and J.A. Thomas, *Elements of Information Theory.* Wiley-Interscience, 1991.
[28] R. Dingledine, V. Shmatikov, and P. Syverson, "Synchronous Batching: From Cascades to Free Routes," *Proc. Privacy Enhancing Technologies Workshop (PET '04),* May 2004.