Fundamental Limits on the Anonymity Provided by the MIX Technique

Dogan Kesdogan
RWTH Aachen
kesdogan@acm.org

Dakshi Agrawal
IBM Watson Res. Ctr.
agrawal@us.ibm.com

Vinh Pham
RWTH Aachen
vinh.pham@gmx.net

Dieter Rautenbach
University of Bonn
rauten@or.uni-bonn.de

Abstract

The MIX technique forms the basis of many popular services that offer anonymity of communication in open and shared networks such as the Internet. In this paper, fundamental limits on the anonymity provided by the MIX technique are found by considering two different settings. First, we consider an information theoretic setting to determine the extent of information inherent in observations of the traffic passing through the MIX. We show that if the size of sender anonymity sets is less than the total user population, the information contained in traffic observations is sufficient to deduce all communication relationships between senders and receivers using the MIX. More importantly, we show that even if every user sends a message in each communication round, it is possible to compromise the anonymity significantly. We precisely characterize the extent of compromised anonymity in each case.

In the second setting, we assume that the attacker has unlimited computational resources and is free to choose any attack algorithm. We derive tight upper and lower bounds on the minimum number of observations required to deduce all recipient peer-partners of a targeted user. The analysis done in these two settings reveals many discrete mathematical structures inherent in anonymity sets, and the intuition gained from these structures can be used when designing or using a MIX based anonymity technique.

1 Introduction

According to the Merriam-Webster's online dictionary [22] privacy is the quality or state of being in retirement from observation of others. However, popular network protocols (e.g. TCP/IP used in the Internet) allow all communications or used services to be observable—at least network providers and intruders have access to traffic information such as who has communicated to whom, for how long, and from which location. Thus, the problem of traffic observability is one of the most fundamental problems in the privacy of digital communication. A number of *anonymity* techniques have been proposed to solve this problem [5, 6, 28, 15, 16, 18, 3, 7, 13, 11].

Our goal in this paper is to investigate the fundamental limits on the protection provided by the anonymity techniques. We carry out our investigation by using the well known MIX technique proposed by Chaum [5]. We use a simplified model of the MIX technique that can be generalized to include other anonymity techniques. In this model, N users want to provide anonymity to each other. The communication happens in rounds, and every user can provide its own (private) message to a central node, the MIX, in each round. The traffic in each round is organized by the participating users and the MIX in such a manner that the traffic information of a particular user is hidden from the attacker due to the additional traffic generated by the other users.

In anonymity literature, the set of N participating users (senders) of the MIX is called the *anonymity set*. In fact, anonymity itself can be defined using the term "anonymity set": anonymity is the state of not being identifiable within a set of subjects, *anonymity set* [27]. In this work, we do not investigate a particular implementation of the MIX protocol. We assume a "perfect" MIX, and investigate the effect of typical user-behavior (e.g. communicating within a circle of friends, visiting a set of websites frequently) on the anonymity set. Specifically, we investigate how repeated communication (to a number of fixed communication partners) can reduce the anonymity.

In recent years, a number of attacks on the MIX technique have been proposed and analyzed [30, 2, 4, 14, 17, 33, 1, 8, 32, 31, 34, 9, 10, 12, 19, 20, 24, 21, 35, 23]. These attacks can be classified by the fact whether the attack exploits a weakness of the MIX protocol or if the attack is independent of the MIX protocol[1]. Since the focus of this paper is on the fundamental limits, we skip all protocol dependent attacks such as attacks involving flooding the network, sending more messages than allowed, or replaying messages. Instead we focus on *traffic analysis attacks* which are more related to our work [2, 4, 17, 33, 1, 8, 32, 31, 34, 9, 10, 12, 19, 20, 21, 24, 35, 23]. Traffic analysis attacks are mostly passive attacks[2], and can

[1] For a nice classification and overview of various attacks, see [30].

[2] An active timing attack is suggested in [20].

be as simple as counting packets on lone connections[3] [31], or as complicated as correlating data flows by any observable patterns [35](e.g. time and frequency statistics). These attacks are harder to thwart since they exploit the information leakage (inference) inherent in the anonymity sets. They model the information leakage and its accumulation over time mostly in terms of probabilities (stochastic models) [8]. To make traffic analysis attacks harder the anonymity set size can be increased[4] or additional measures, such as using dummy messages, can be taken[5]. However, to quantify the benefits of additional measures, there is a need to understand and precisely describe the discrete mathematical structure behind the anonymity sets.

Our goal in this paper is to extend the previous work in this direction [17, 1, 8, 12, 21, 9, 10, 19] by exploring fundamental limits in two settings: the first setting shows to which extent anonymity sets leak information (information theoretic setting), and the second setting computes the minimal number of observations required by an attacker to compromise the anonymity provided by anonymity sets (practical setting).

This paper is organized as follows. In the next section, we will provide basic terminology and overview the MIX technique. In Section 3, we describe the contributions of this paper. In Section 4, we consider an information-theoretic setting in which the attacker has access to all possible observations that the MIX could produce as well as access to unlimited computational resources. Under this setting, we calculate the extent of communication relationships that such an attacker can derive. In Section 5, we consider a practical setting in which the attack does not have access to all possible observations of the MIX. Instead, the attacker watches acts of communication one by one as they are performed by the targeted user and stops when the anonymity of the targeted user has been compromised. We calculate the lower and upper bounds on the minimum number of observations required by the attacker in this setting, and compare these bounds to the results obtained by simulation. In Section 6, we summarize our results.

2 Background

2.1 The basic MIX technique

We consider an *omnipresent passive attacker* who is capable of observing all communication links simultaneously. Against an omnipresent passive attacker[6] the anonymity of a single transmission by a single person can not be protected—the attacker can observe the act of sending a message and follow the message physically to the receiver, thereby detecting the act of communication between the sender and the receiver of a message. Hence, anonymity techniques require additional *cover traffic* to confuse the attacker and conceal communication relationships between senders and recipients of messages.

A well known example from everyday life is the *ballot box* used in electoral procedures. The caster of a specific vote is hidden among a number of other voters (cover traffic) by confidentially collecting a number of votes from N distinct voters in a closed box, and by randomly shuffling (e.g. by shaking) the votes in the ballot box.

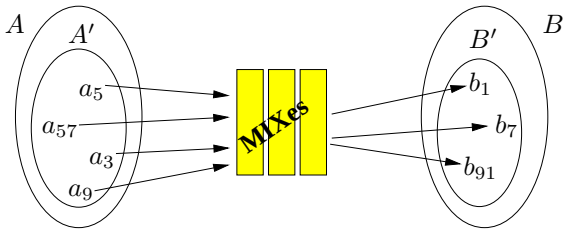

Figure 1. Formal model

In 1981, David Chaum proposed the *MIX* technique that is analogous to the ballot box example [5]. Figure 1 shows the basic ingredients of this technique which consist of a set of senders A, a set of recipients B, and a MIX node. All senders in A are connected to the MIX and the MIX itself is connected to all recipients in B by a communication network with reliable secure channels. A reliable secure channel does not result in loss or duplication of transmitted messages, and guarantees authenticity, integrity, and confidentially of transmitted messages. The users and the MIX transmit messages by using the following protocols:

User Protocol: Users prepare their messages to be of constant length either by splitting long messages or by padding short messages to the specified length. Each message is encrypted twice with one time pads: first the message is encrypted using a shared secret between the sender and the intended recipient, and then it is encrypted using a shared secret between the sender and the MIX. The users send twice encrypted messages to the MIX.

MIX Protocol: A MIX collects n messages (called a *batch*) from distinct users, decrypts the messages, and outputs the decrypted messages in a batch in a different order than the order in which they were received (lexicographically sorted or randomly delayed). The output is broadcasted to all recipients. Furthermore, any incoming packet is compared with formerly received messages

[3] So it is possible to follow anonymized streams.

[4] Increasing anonymity size has a cost: the operation time of the protocol has to be increased to build the appropriate anonymity sets.

[5] Use of dummy messages is prohibitive on large scale networks like the Internet.

[6] In the rest of this paper, the term attacker means omnipresent passive attacker unless qualified otherwise.

(i.e. by locally caching formerly received messages) in order to reject any duplicate messages.

The basic MIX technique described above can perfectly hide the communication relationships between senders and recipients of messages from everybody but the MIX and message senders. Even the act of sending or receiving can be perfectly hidden if the above protocol is applied in fixed time slots, and if every user supplies a fixed number of messages (perhaps some or all of them being dummy messages) to each slot and the whole output batch in a time slot is distributed to every user [25, 5, 26]. Pfitzmann [26] states that the MIX technique provides information-theoretic anonymity and unobservability based on complexity-theoretic secure cryptography.

2.2 The pure MIX technique

The "perfect" anonymity solution discussed above uses dummy messages. Even though this solution can provide perfect anonymity, it is not followed widely in large networks such as the Internet. The reasons are manifold and some of them are listed below:

Feasibility Reason If the protocol is applied in rounds then all senders have to participate with a message. This procedure requires not only that all participants are synchronized, but also that all users are able and willing to send at predetermined times, whether anyone really wants to send or not. This is hard to realize in large networks such as the Internet with millions of users, some of which may be off-line.

Cost Reason Implementation of the dummy messages in a MIX for a large network is very expensive. For instance, consider $10,000$ users sending one message to the MIX. Assuming that the fixed message size is 4 KB, the data sent by users is $40,000$ KB. Assuming that all the messages are broadcasted to all users, the volume of data increases to $400,000$ MB.

As a consequence, most current implementations and solutions use a variant of the perfect MIX solution by neither using dummy messages nor the broadcasting function. In other words, resources are occupied only if real information is transmitted and the cover traffic consists only of the real traffic. We refer to a MIX technique without dummy messages or the broadcasting function as a *pure* MIX technique. Specifically, in the rest of this paper, we will use the following formal model of a MIX and information leakage therein for our analysis.

Formal Model of the Pure MIX Technique

- A communication system consists of a set of senders A, and a set of recipients B, and a MIX node (see Figure 1). If a sender $a \in A$ communicates with a recipient $b \in B$, then we say that a and b are peer partners. If the roles of sender and receiver need to be distinguished, then we say that a is a peer sending partner of b and b is a peer recipient partner of a.
- In each *communication round*[7] a subset $A' \subseteq A$ of all senders A send a message to their peer partners. Let $B' \subseteq B$ be the set of intended recipients. The act of sending or receiving a message is not hidden among dummy messages.
- The size of the sender anonymity set is $|A'| = n$, where $1 < n \leqslant |A| = N$.
- The size of the recipient anonymity set is $|B'| \leqslant n$ since each sender sends exactly one message and several senders may communicate with the same recipient.
- The information leakage X available to an attacker in a communication round consists of the pair (A', B') of peer senders and receivers.

3 Our Contributions

It is evident from the prior research work that anonymity sets provide a limited protection of anonymity in presence of repeated communication. Kesdogan, Agrawal, and Penz proposed the *disclosure attack* to identify all peer recipient partners of a targeted user [17]. They subsequently analyzed the number of observations required by an attacker to mount the disclosure attack, and showed that the disclosure attack is an NP-complete problem [1]. Danezis significantly improved the performance of the disclosure attack by exploiting statistical properties of the observations and proposed the *statistical disclosure attack* [8, 10]. Kesdogan and Pimenidis have recently proposed a deterministic attack, the *hitting set attack*, and a variation that exploits statistical properties of observations, the *statistical hitting set attack*, on the MIXes [19]. By using simulations, they showed that the statistical hitting set attack requires the least number of observations among all known attacks.

The goal of this paper is extend this line of work by identifying the fundamental structure of anonymity sets and by deriving fundamental limits that are independent of attack methods. To that end, we use two different adversarial models

[7] A communication round consists of the following events: The MIX node collects messages from a fixed number of distinct senders, and after applying the "MIX" protocol, it forwards the collected messages to their intended recipients.

to derive the following two different limits on the anonymity provided by a pure MIX technique:

Information Theoretical Limits: We derive information theoretical limits on anonymity by assuming that the attacker has observed all possible pairs of sender and receiver sets (A', B') and that the attacker has unbounded computational resources. These assumptions allow us to compute the absolute extent of information present in the observations of the traffic passing through the MIX.

We show that for open sender groups, that is, for $|A'| = n < |A| = N$, the attacker gains knowledge of all peer recipient partners of a targeted sender. For closed sender groups, $A' = A$, the attacker cannot deduce all peer partners of a targeted sender. However, the attacker can still obtains valuable information such as how many peer senders a particular recipient has, or if two recipients share a common sender.

Practical Limits: Here we assume that the information leakage is theoretically sufficient (i.e. $|A'| = n < |A| = N$) to deduce all peer partners of a targeted user Alice. In contrast to the previous case, we assume that the attacker gathers observations one by one as they happen, and stops when there are sufficient observations to compromise the anonymity of a targeted user.

Under these assumption, we derive tight lower and upper bounds on the minimum number of observations required on the average by an attacker to deduce all peer partners of a targeted user Alice. We first show that the hitting set attack requires the least number of observations among all possible deterministic attacks, and therefore a bound on the number of observations required by the hitting set attack is a fundamental bound. We compare derived bounds to the simulated results and show that bounds computed in the paper provide a good approximation of the simulated results. The process of bound calculation also illuminates the fundamental structure of anonymity sets.

4 Information Theoretic Limits

We model the communication relationships among senders and receivers by a bipartite *relationship graph* $G = (A \cup B, E)$ (see Figure 2). The partite sets A and B represent the sets of senders and receivers respectively, while the set E consists of edges between senders and their peer recipient partners. Thus if Alice communicates with Bob then there is a corresponding edge in E between the vertex representing Alice in A and the vertex representing Bob in B (see Figure 2). Note that even if the real users behind the senders and recipients are the same persons, we distinguish them by their attributes "sender" and "recipient" with respect to a message

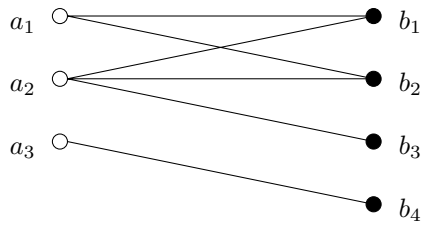

Figure 2. Bipartite relationship graph with three senders and four receivers

and put them in disjoint sets A and B. Each sender has a subset of all recipients as its peer partners. The sets of peer partners of two different senders may be disjoint, be the same, or overlap with each other. We now introduce three useful terms: *multiset*, *multiplicity*, and *neighbor function*.

Definition 1 *A multiset S with cardinality k is an unordered list of k elements $x_1, x_2, ..., x_k$ such that $x_i = x_j$ for $1 \leqslant i < j \leqslant k$ is possible. The multiset S is denoted by $\langle x_1, x_2, \ldots, x_k \rangle$. Since S is unordered, $\langle x_k, x_{k-1}, \ldots, x_1 \rangle$ is also one of the possible notations for S.*

Definition 2 *The multiplicity of x in a multiset $S = \langle x_1, x_2, ..., x_k \rangle$ is denoted by $mult(x, S)$ and is given by the number of times x occurs in S, that is, $\mathrm{mult}(x, S) = |\{i \mid 1 \leqslant i \leqslant k,\, x = x_i\}|$.*

Definition 3 *The neighbor function $N(T)$, $T \subseteq A$ or $T \subseteq B$, gives the set of all peer partners of nodes in T,*

$$N(T) = \{v \mid v \text{ is adjacent to a node } t \in T\}.$$

If the set T is a singleton $T = \{t\}$, then for the convenience of notation, we denote $N(\{t\})$ by $N(t)$. In Figure 2, $N(\{a_1, a_2\}) = \{b_1, b_2, b_3\}$ and $N(a_3) = \{b_4\}$.

Given a bipartite relationship graph G and the batch size n, we can derive a mapping M that maps each *sender set* $A' = \{a_1, a_2, ..., a_n\} \subseteq A$ of size n to a set of all possible *receiver multisets* of A':

$$M(A') = \{B' = \langle b_1, b_2, ..., b_n \rangle | B' \subseteq B \text{ and } b_i \in N(a_i) \text{ for } 1 \leqslant i \leqslant n\}$$

In Figure 2, $M(\{a_1, a_2\}) = \{\langle b_1, b_1 \rangle, \langle b_1, b_2 \rangle, \langle b_2, b_2 \rangle, \langle b_1, b_3 \rangle, \langle b_2, b_3 \rangle\}$.

We assume that the attacker has access to all possible observations that may occur through the MIX. This is equivalent to assuming that the attacker knows the full mapping M defined by G and n. We are interested in asking the following question: Given M and n, what does the attacker know about G? In the next two subsections, we will show that for closed

sender groups, $n = N$, the attacker cannot reconstruct G, however, it can derive important functions that reveal structure of G. For open sender groups, $n < N$, the attacker can reconstruct G and therefore, deduce all peer partners of all senders. The attacker can derive more in the open sender group since in this case, the information leakage is contained in both the sender as well as the receiver sets while in the closed sender group case, the information leakage is contained only in the receiver sets.

4.1 Closed sender group

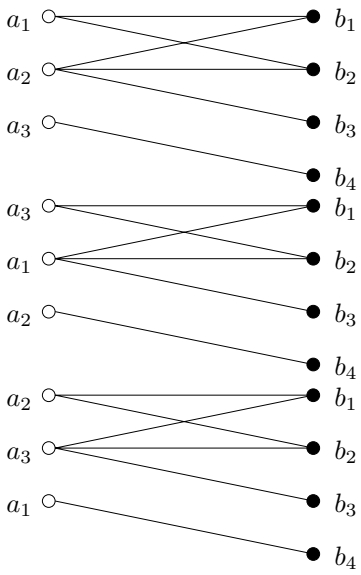

Figure 3. Permutation of nodes in bipartite graph.

This case is characterized by $|A'| = |A|$, and therefore in the information leakage $X = (A', B')$, only the receiver multisets B' provides some information about the underlying relationship graph G.

Consider an anonymity system Ψ defined by (G, n). We can obtain another anonymity system $\Psi' = (G', n)$ by permuting the labels of nodes in the sender partite-set A while leaving the edge set E and recipient partite-set B unchanged (see Figure 3). Since the permutation of the node labels in the sender partite-set A does not change the 'structure' of the graph G, the receiver multisets observed in Ψ' would be exactly the same as those observed in Ψ. *Therefore, the best we can hope in the case of the closed sender group is to learn the 'structure' of the graph G. In the rest of this section, we formally show that indeed it is possible to deduce the 'structure' of graph G by observing receiver multisets.*

4.1.1 Structure of a bipartite graph

To formally capture the notion of the 'structure' of a bipartite graph $G = (A \cup B, E)$, consider an alternative representation in which the edge set E is represented by a matrix Σ of size $2^{|B|} \times |A|$. The rows of Σ are indexed by the subsets of B and the columns are indexed by the vertices in A. The value of $\Sigma[B'][a_j]$, $B' \subseteq B$ and $a_j \in A$, is 1 if B' is the set of all peer recipients partners of a_j, and is 0 otherwise. Thus, each column of Σ has only one non-zero entry. Since $N(a), \forall a \in A$, specify the graph G completely, $(A \cup B, \Sigma)$ is a complete albeit highly redundant representation of the graph G.

For a bipartite graph $G = (A \cup B, E) = (A \cup B, \Sigma)$, permuting the node labels in the partite set A by using a permutation π ($a_j \rightarrow a_{\pi(j)}$) while leaving the edge set E and the recipient set B unchanged to obtain the graph G' has the effect of permuting the columns of Σ to obtain a matrix Σ' such that $\Sigma[\cdot][a_j] = \Sigma'[\cdot][a_{\pi(j)}]$ and $G' = (A \cup B, \Sigma')$. Since the permutation π can be arbitrarily chosen, the only common structure among Σ and Σ' is the weight of rows $\Sigma[B'][\cdot]$ denoted by $w(B')$. Thus formally, the structure of G is captured by the function $w(B'), B' \subseteq B$, the number of nodes in A that have B' as their neighbor.

4.1.2 Deducing the structure of a bipartite graph

In the following we will prove that the function $w(B')$ can be computed by observing receiver multisets in the leaked information $X = (A, B')$. We will compute the function $w(B')$ in two steps. In the first step, we compute an ancillary function $u(B')$ that gives the number of senders in A whose peer partners are a subset of B', that is,

$$u(B') = |\{a \mid a \in A \text{ and } N(a) \subseteq B'\}|$$

Let S^c denote the complement of the set S. Given $B' \subseteq B$, partition the set A in three disjoint subsets, $A_{B'}, A_{B'^c}$, and A_{B',B'^c}, where each sender in $A_{B'}$ has all of its peer partners in B', each sender in $A_{B'^c}$, has all of its peer partners in B'^c, and each sender in A_{B',B'^c} has peer partners in both B' and B'^c. Clearly,

$$u(B') = |A_{B'}| \tag{1}$$

$$N = |A_{B'}| + |A_{B'^c}| + |A_{B,B'^c}| \tag{2}$$

For the given set of receivers B', chose a receiver multiset such that the number of times the receivers in B' occur in the chosen receiver multiset is maximum among all possible receiver multisets. Denote the chosen receiver multiset by $R_{B'}$:

$$R_{B'} = \underset{R}{\text{argmax}} \sum_{b \in B'} \text{mult}(b, R)$$

where $\text{argmax}_x f(x)$ equals to a value of x that maximizes $f(x)$.

In the sender set corresponding to the receiver multiset $R_{B'}$, all senders who could send a message to a receiver in B' did so, and therefore,

$$\sum_{b \in B'} \text{mult}(b, R_{B'}) = |A_{B'}| + |A_{B', B'^c}|. \quad (3)$$

Similarly, pick a receiver multiset $R_{B'^c}$ such that the sum of multiplicity of nodes in B'^c in $R_{B'^c}$ is maximum among all receiver multisets. By a similar reasoning,

$$\sum_{b \in B'^c} \text{mult}(b, R_{B'^c}) = |A_{B'^c}| + |A_{B', B'^c}|. \quad (4)$$

We have four independent linear equations (1)–(4) in four variables, $u(B'), A_{B'}, A_{B'^c}$, and A_{B', B'^c}, and we can solve for $u(B')$.

In the second step, we derive $w(B')$ from $u(B')$ by applying the Möbius inversion formula. Recall that according to the Möbius inversion formula if

$$g(E) = \sum_{F\,:\,F \subseteq E} h(F)$$

then

$$h(E) = \sum_{F\,:\,F \subseteq E} (-1)^{|E|-|F|} g(F)$$

Since

$$u(B') = \sum_{B''\,:\,B'' \subseteq B'} w(B'') \quad (5)$$

by using the Möbius inversion formula, we can calculate the function $w(B')$ from the function $u(B')$.

Knowledge of the function $w(B')$ (and the ancillary function $u(B')$) enables an attacker to compromise anonymity: the value of $w(\{b\})$, $b \in B$, is equal to the number of senders who only communicate with b; $N - u(\{b\}^c)$ is equal to the number of senders who communicate to b; $N - u(\{b_1\}^c) - u(\{b_2\}^c) + u(\{b_1, b_2\}^c)$, $b_1, b_2 \in B$, provides the number of common sender partners of b_1 and b_2 etc.

4.2 Open sender group

In an open environment, the attacker can observe the relationship between proper subsets of $A' \subset A$ to the subsets of $B' \subset B$. We next show that in such environments, an attacker can reconstruct the relationship graph G, and deduce who communicates with whom.

Note that the sets $N(b) = \{a \in A \mid b \in N(a)\}$, $b \in B$, are sufficient to construct G, and it suffices to show that the attacker can determine $N(b)$ for all $b \in B$. Let m_b be the maximum possible multiplicity of b in receiver multisets corresponding to the sender sets of size n, that is,

$$m_b = \max_{\tilde{A} \subseteq A,\, |\tilde{A}| = n} \left(\max_{\tilde{B} \in M(\tilde{A})} \text{mult}(b, \tilde{B}) \right).$$

There are three cases to be considered:

• If $m_b = 0$, then $|N(b)| = 0$, and as a consequence $N(b) = \emptyset$.

• If $m_b = n$, then $|N(b)| \geqslant n$. Let $\mathcal{A}'$ be the set of all sender sets $\tilde{A} \subseteq A$ such that $|\tilde{A}| = n$ and the maximum possible multiplicity of b in receiver multisets corresponding to $\tilde{A}$ is n, that is,

$$\mathcal{A}' = \{\tilde{A} | \tilde{A} \subseteq A \text{ and } \max_{\tilde{B} \in M(\tilde{A})} \text{mult}(b, \tilde{B}) = |\tilde{A}| = n\}.$$

If $\tilde{A} \in \mathcal{A}'$, then all members of $\tilde{A}$ have b as its neighbor, that is, $\tilde{A} \subset N(b)$. If that was not the case, the maximum multiplicity of b in receiver multisets corresponding to $\tilde{A}$ would be less than n, producing a contradiction with the definition of $\mathcal{A}'$.

We claim that $N(b) = \bigcup_{\tilde{A} \in \mathcal{A}'} \tilde{A}$. As $\tilde{A} \in \mathcal{A}'$ implies that $\tilde{A} \subset N(b)$, $\bigcup_{\tilde{A} \in \tilde{\mathcal{A}}} \tilde{A} \subset N(b)$. We will prove $N(b) \subset \bigcup_{\tilde{A} \in \mathcal{A}'} \tilde{A}$ by contradiction. Assume that there exist $a' \in N(b)$ such that a' is not a member of any set $\tilde{A}$ in $\mathcal{A}'$. Now take a set A' in $\mathcal{A}'$, and replace one of its element by a' to obtain the sender set A''. This operation would preserve the cardinality of A'' to n since a' is not a member of $A' \in \mathcal{A}'$. However since all member of thus constructed A'' have b as its neighbor,

$$n = \max_{B'' \in M(A'')} \text{mult}(b, B'').$$

Therefore, $A'' \in \tilde{\mathcal{A}}$ and a' is a member of a set in $\mathcal{A}'$, a contradiction.

• If $1 \leqslant m_b \leqslant n-1$, then $|N(b)| = m_b$, that is, $b \in N(a)$ for exactly m_b elements of A. Let A' be a set for which

$$m_b = \max_{B' \in M(A')} \text{mult}(b, B').$$

Clearly $N(b) \subset A'$. We can examine each element a' of A' in turn to see if it belongs to $N(b)$ by using the following procedure. Consider $a' \in A'$ and construct a sender set A'' by substituting a' from A' with an element of $A \setminus A'$, and compute

$$m''_b = \max_{B'' \in M(A'')} \text{mult}(b, B'').$$

Clearly, if $m''_b = m_b$, then $a' \notin N(b)$, and if $m''_b < m_b$, then $a' \in N(b)$.

Hence in each case, we can determine $N(b)$ from the information leakage and the proof is complete. □

4.3 Example

Consider the bipartite relationship graph with three senders a_1, a_2, and a_3, and four receivers b_1, b_2, b_3, and b_4, given in Figure 2. We will illustrate the theoretical limits for the closed and open sender-group cases by using this relationship graph.

4.3.1 Closed sender group calculations

From Figure 2, for the closed sender group, the receiver multisets are given by:

$$M(\{a_1, a_2, a_3\}) = \{\langle b_1, b_1, b_4\rangle, \langle b_1, b_2, b_4\rangle, \langle b_2, b_2, b_4\rangle, \langle b_1, b_3, b_4\rangle, \langle b_2, b_3, b_4\rangle\}$$

Next we will produce the alternative representation of the relationship graph in terms of the matrix Σ as described above. We will assume that the first, second, and third columns of Σ correspond to the senders a_1, a_2, and a_3 respectively, and that the first, second, and third rows correspond to the receiver set $\{b_1, b_2\}, \{b_1, b_2, b_3\}$, and $\{b_4\}$, respectively. The matrix Σ has 2^4 rows, one row for each subset of B, however all of its rows except the first three are zero rows since the peer recipients of a_1, a_2, and a_3 are $\{b_1, b_2\}, \{b_1, b_2, b_3\}$, and $\{b_4\}$, respectively. The matrix Σ is given by

$$\Sigma = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ \cdots & \cdots & \cdots \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \quad (6)$$

The weight function for this matrix is given by:

$$w(B') = \begin{cases} 1 & \text{for } B' = \{b_1, b_2\}, \{b_1, b_2, b_3\}, \{b_4\} \\ 0 & \text{otherwise} \end{cases}$$

The theoretical limit derived for the closed sender group indicates that although we cannot derive the matrix Σ from the receiver multisets, we can deduce the weight function $w(B')$ of the matrix Σ. In order to deduce the weight function $w(B')$, we first need to calculate the ancillary function $u(B')$. Set $B' = \{b_1, b_2\}$. We will illustrate how to calculate $u(\{b_1, b_2\})$. We pick the receiver multiset $\langle b_1, b_1, b_4\rangle$ as $R_{B'}$ since the sum of multiplicity of recipients b_1 and b_2 is maximum in this receiver multiset. Other equally good choices are $\langle b_1, b_2, b_4\rangle, \langle b_2, b_2, b_4\rangle$ since they also contain b_1 and b_2 twice.

Similarly, we pick the receiver multiset $\langle b_1, b_3, b_4\rangle$ as $R_{B'^c}$ since it maximizes the sum of multiplicity of b_3 and b_4 to 2. We could have also chosen $\langle b_2, b_3, b_4\rangle$ as $R_{B'^c}$.

Now as discussed above we obtain the following equations:

$$\sum_{b \in B'^c} \text{mult}(b, R_{B'^c}) = 2 = |A_{B'^c}| + |A_{B', B'^c}| \quad (7)$$

$$\sum_{b \in B'} \text{mult}(b, R_{B'}) = 2 = |A_{B'}| + |A_{B', B'^c}| \quad (8)$$

$$3 = |A_{B'}| + |A_{B', B'^c}| + |A_{B'^c}| \quad (9)$$

Solving above three equations, we obtain $|A_{B'}| = 1, |A_{B'^c}| = 1$, and $|A_{B', B'^c}| = 1$. Using (1), $u(\{b_1, b_2\}) = 1$, that is number of senders whose peer recipients are a subset of $\{b_1, b_2\}$ is 1. One quick look at the graph confirms this calculation as a_1 is the only sender not to have a recipient outside $\{b_1, b_2\}$.

By using the same procedure for other subsets of B, we can deduce the ancillary function u completely as follows:
$u(\{b_1\}) = 0$, $u(\{b_2\}) = 0$, $u(\{b_3\}) = 0$, $u(\{b_4\}) = 1$,
$u(\{b_1, b_2\}) = 1$, $u(\{b_1, b_3\}) = 0$, $u(\{b_1, b_4\}) = 1$,
$u(\{b_2, b_3\}) = 0$, $u(\{b_2, b_4\}) = 1$, $u(\{b_3, b_4\}) = 1$,
$u(\{b_1, b_2, b_3\}) = 2$, $u(\{b_1, b_2, b_4\}) = 2$,
$u(\{b_1, b_3, b_4\}) = 1$, $u(\{b_2, b_3, b_4\}) = 1$,
and $u(\{b_1, b_2, b_3, b_4\}) = 3$.

By using Möbius's inversion formula, we get
$w(\{b_1, b_2, b_3\}) = u(\{b_1, b_2, b_3\}) - u(\{b_1, b_2\}) - u(\{b_1, b_3\})$
$-u(\{b_2, b_3\}) + u(\{b_1\}) + u(\{b_2\}) + u(\{b_3\})$
$= 2 - 1 - 0 - 0 + 0 + 0 + 0 = 1$

Similarly we can derive the value of $w(\cdot)$ for other subsets of B.

4.3.2 Open sender group calculations

We will illustrate the open sender group case by using the bipartite graph in Figure 2 with $n = 2$. In this environment we have the following cases:

1. $A' = \{a_1, a_2\}$ and
   $M(A') = \{\langle b_1, b_1\rangle, \langle b_1, b_2\rangle, \langle b_2, b_2\rangle, \langle b_1, b_3\rangle, \langle b_2, b_3\rangle\}$
2. $A' = \{a_1, a_3\}$ and $M(A') = \{\langle b_1, b_4\rangle, \langle b_2, b_4\rangle\}$
3. $A' = \{a_2, a_3\}$ and
   $M(A') = \{\langle b_1, b_4\rangle, \langle b_2, b_4\rangle\ \langle b_3, b_4\rangle\}$

Now focus on b_1. The maximum possible multiplicity of b_1 in receiver multisets corresponding to the sender sets of size two, represented by m_{b_1} is two and it occurs for the sender set $\{a_1, a_2\}$. As discussed above, since for b_1, $m_{b_1} = n$, $N(b_1) = \{a_1, a_2\}$.

Now consider b_2. The maximum possible multiplicity of b_2 in receiver multisets corresponding to the sender sets of size two is again two and it occurs for the sender set $\{a_1, a_2\}$. Therefore, for b_2 also $N(b_2) = \{a_1, a_2\}$.

The case for b_3 differs from b_1 and b_2 in that m_{b_3} is only 1 and it occurs for sender groups $\{a_1, a_2\}$ and $\{a_2, a_3\}$. As discussed above, we need to check whether a_1, a_2, and a_3 belong to the set $N(b_3)$. To do so, replace a_1 in $\{a_1, a_2\}$ by a_3 to obtain $\{a_2, a_3\}$. Since among the receiver multisets corresponding to $\{a_2, a_3\}$ has the maximum multiplicity of b_3 is 1, $a_1 \notin N(b_3)$. To check whether a_2 is in $N(b_3)$, replace a_2 in $\{a_1, a_2\}$ by a_3 to obtain $\{a_1, a_3\}$. The maximum multiplicity of b_3 in the receiver multisets corresponding to $\{a_1, a_3\}$ is 0, and therefore $a_2 \in N(b_3)$. Finally by the same reasoning as the one given for a_1, $a_3 \notin N(b_3)$. Therefore $N(b_3) = \{a_2\}$.

The case for b_4 is similar to the one for b_3, and we can deduce that $N(b_4) = \{a_2\}$.

Thus we can deduce $N(b)$ for $b = b_1, b_2, b_3$, and b_4, and reconstruct the communication relationship graph completely.

5 Practical Limits

Information theoretic limits answer the general question of whether the anonymity sets leak sufficient information to deduce peer partners of a targeted user. These results show that information leaked in all possible observations of the system is sufficient to deduce peer partners if and only if $A' \subset A$. In practice, an attacker does not have access to all possible observations. Instead, the attacker starts at a given time, observes communication rounds in which the targeted user Alice participates, and stops when all peer partners of Alice have been deduced. Important questions from the view point of a practical attacker are the following. What is the minimum number of observations required to deduce all peer partners of a targeted user? Is there a feasible attack that can deduce all peer partners using only the minimum number of observations? In this section, we will answer these questions.

Our first observation is that the hitting set attack proposed by Kesdogan and Pimenidis [19] requires the least number of observations among all possible attacks, known or unknown, that deterministically find all peer partners of Alice. This observation can be easily made by considering the following description of the hitting set attack[8].

Hitting Set Attack To mount the hitting set attack, the attacker starts with the set $\mathcal{S}_0$ that contains all $\binom{N}{m}$ possible subsets of cardinality m of N recipients[9]. Since Alice has m peer partners, exactly one subset in $\mathcal{S}_0$ is the set of all peer partners of Alice. Let $\{B_1, B_2, B_3, \ldots, \}$ be the recipient sets in the successive communication rounds in which Alice participates. Since Alice has a peer partner in B_1, a set in $\mathcal{S}_0$ that has an empty intersection with B_1 cannot be the set of all peer partners of Alice. Thus upon observing B_1, the attacker obtains a new solution set $\mathcal{S}_1$ by discarding all recipients sets in $\mathcal{S}_0$ that have an empty intersection with B_1. The attacker repeats this process to generate solution sets $\mathcal{S}_2, \mathcal{S}_3, \ldots$ after observing recipient sets $B_2, B_3, \ldots$ respectively, until the solution set $\mathcal{S}_T$ has only one subset in it. The last remaining subset in the solution set $\mathcal{S}_T$ has to be the set of all peer partners of Alice. Note that the hitting set attack finds the *unique minimal* hitting set of all observations. Also note that under our assumptions, the above procedure will stop in finite number of observations.

Corollary 5.1 *The hitting set attack requires the minimum number of observations to deterministically find all peer partners of Alice.*

Proof In the above construction, all subsets in $\mathcal{S}_i$ are consistent with the observations $B_1, B_2, \ldots, B_i$. Thus as long as there are more than one subset in $\mathcal{S}_i$, all peer partners of Alice cannot be deduced deterministically. As a corollary, the hitting set attack requires the minimum number of observations to deterministically find all peer partners of Alice. □

5.1 Lower bound on the number of observations required for the hitting set attack

Claim 1 *Let Bob be one of the peer partners of Alice. We claim that before finishing the hitting set attack, either Bob occurs as the only recipient in a recipient set or Bob occur as the only peer partner of Alice in at least two recipient sets.*

Proof We will prove this claim by contradiction. Assume that the attacker has finished the hitting set attack, that is, it has found $N(\text{Alice})$ which equals to the *unique minimal* hitting set of cardinality m of all observed recipient sets. Also assume that neither Bob occurs as the only recipient in a recipient set nor does he occur as the only peer partner of Alice in at least two recipient sets. There are two cases in which this can happen. In the first case, all recipient sets observed by the attacker that include Bob also include at least one another peer partner of Alice[10]. In this case,

[8] The description of the hitting set attack given here is for deriving our observation. In practice, the attacker would use efficient implementations to circumvent computational complexity of the hitting set attack [29].

[9] Here we assume that the attacker knows the value of m. See [17] for a justification of this assumption.

[10] This includes the case when Bob has not been observed by the attacker in any of the recipient sets.

$H' = N(\text{Alice})/\{\text{Bob}\}$ is a hitting set of cardinality $m-1$ contradicting our assumption that the attacker has found a unique hitting set of cardinality m.

In the second case, Bob occurs as the only peer partner of Alice in exactly one observed recipient set B', and all other observed recipient sets that include Bob also include at least one other peer partner of Alice. Since according to our assumptions, Bob does not occur as the only recipient in B', there is another recipient, say Carolyn, in B' who is not a peer partner of Alice. In this case, $H' = N(\text{Alice})/\{\text{Bob}\} \cup \{\text{Carolyn}\}$ will also be a hitting set of cardinality m, contradicting our assumption that the attacker has found a unique hitting set of cardinality m. □

If a recipient set contains only one peer partner of Alice, we will say that the peer partner of Alice occurs *exclusively*. We can obtain a lower bound on the number of observations required by the hitting-set attack by counting the number of observations required to see each peer partner of Alice either occur as the only recipient in an observed recipient set or occur exclusively at least twice in observed recipient sets. We loosely call it "two-exclusivity" observation of Alice's peer partners [11]. We compute this number by constructing a Markov process X_t, where t denotes the index of B_t. The state-space of X_t is a triplet $\{m_0, m_1, m_2\}$, where m_0 is the number of peer partners of Alice that have not been observed, m_1 is the number of peer partners of Alice that have been observed exclusively exactly once, and m_2 is the number of peer partners of Alice that have been observed either at least twice exclusively in recipient sets or at least once as the only recipient in a recipient set.

Clearly $m_0 + m_1 + m_2 = m$. Label the state $\{m_0, m_1, m_2\}$ by index $i = m_0(m+1) - m_0(m_0-1)/2 + m_1 + 1$. This assign each state a unique label i, $1 \leqslant i \leqslant (m+1)(m+2)/2$[12].

In the beginning of the hitting set attack, the state is given by $\{m, 0, 0\}$ and is indexed by $(m+1)(m+2)/2$. The state $\{0, 0, m\}$, indexed by 1, corresponds to the point before which hitting set attack cannot be concluded. Let M be the transition matrix[13] of X_t with the M_{ij} denoting the probability of going from the i-th state to the j-th state after making an observation. Let $e(j)$ be the column matrix of size $(m+1)(m+2)/2 \times 1$ with j-th entry equal to 1 and other entries zero. The probability of ending in state $\{0, 0, m\}$ after T observations is given by

$$P(T) = e'((m+1)(m+2)/2)M^T e(1) \tag{10}$$

where x' denotes the transpose of the matrix x.

[11] The term "two-exclusivity" is loose since if a peer partner occurs as the only recipient in a recipient set, then it is not necessary to observe that peer partner twice.

[12] The state-space is of size $\sum_{m_0=0}^{m} \sum_{m_1=0}^{m-m_0} 1 = (m+1)(m+2)/2$ since for each value of m_0, $m_0 = 0, 1, \ldots, m$, $m_1 = 0, 1, \ldots, m - m_0$.

[13] Appendix A contains explicit formulas for the transition matrix M.

It follows that with probability $P(T) - P(T-1)$ the attacker will take at least T observations to finish the hitting set attack, and a lower bound on the expected number of observations is given by:

$$L = \sum_{k=1}^{k=\infty} k(P(k) - P(k-1)) \tag{11}$$

$$= \sum_{k=1}^{k=\infty} ke'((m+1)(m+2)/2)[M^k - M^{k-1}]e(1) \tag{12}$$

The time and space complexity of computing the lower bound L is mainly determined by the size of the matrix M which only depends on m and is given by $O(m^4)$. With an appropriate mathematical representation, we avoided explicit computation of the whole matrix M, and managed to determine L in the time and space complexity of $O(m^2)$. On a Pentium 4 PC with 2.4GHz processor and 512MB RAM, the highest runtime and space consuming computation ($N = 200000, n = 100, m = 65$) in our plots required less than 1 second and 1MB of memory. For more implementation details, the reader is referred to [29].

5.2 Upper bound on the number of observations required for the hitting set attack

Let $p(l)$ be the probability that l independent recipient sets have at least one non-peer partner of Alice in common. Clearly, $p(l)$ decreases monotonically with l. For our parameters of interests, we found that $p(l) \approx 0$ for $l \geqslant 4$. Thus if a peer partner of Alice occurs exclusively in three recipient sets, then with high probability, the attacker can find identify it as a peer partner. In Appendix C, we formally show that the effect of $p(l)$, for $l \geqslant 4$ is negligible for typical values of MIX parameters, and the number of observations required to see each peer partner exclusively in at least three recipient sets provides a good upper bound on the number of observations required for the hitting set attack.

We can compute the average number of observations required to see each peer partner occur exclusively in at least three recipient sets or as the only recipient in a recipient set by using similar arguments as used for the lower bound. We refer to this as "three-exclusivity" observation [14]. In this case, the state of the Markov process X_t, is a quartet $\{m_0, m_1, m_2, m_3\}$, where m_0 is the number of peer partners of Alice that have not been observed, m_1 is the number of peer partners of Alice that have been observed exactly once

[14] The term "three-exclusivity" is loose since if a peer partner occurs as the only recipient in a recipient set, then it is not necessary to observe that peer partner thrice.

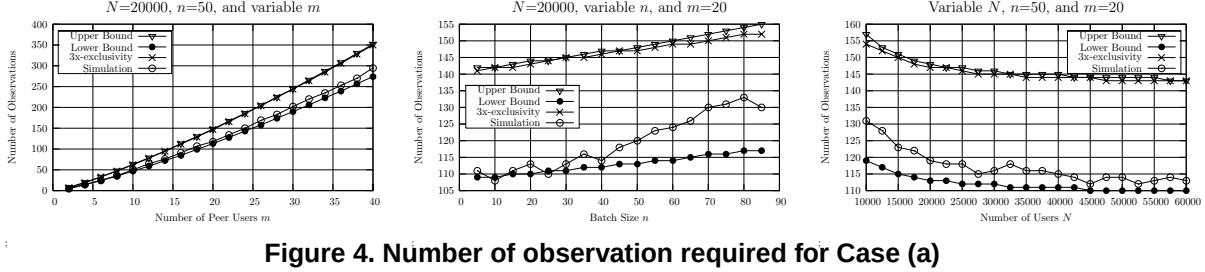

Figure 4. Number of observation required for Case (a)

exclusively, and m_2 is the number of peer partners of Alice that have been observed exactly twice exclusively, and m_3 is the number of peer partners of Alice that have been observed either at least thrice exclusively in recipient sets or at least once as the only recipient in a recipient set.

Clearly $m_0 + m_1 + m_2 + m_3 = m$. Label the state $\{m_0, m_1, m_2, m_3\}$ by index

$$i = \frac{(m-m_3)^2(m-m_3+1) + 2(m-m_3)(m-m_3+1)}{6} + m_0(m-m_3+1) - \frac{m_0(m_0-1)}{2} + m_1 + 1.$$

This assign each state a unique label i, $1 \leqslant i \leqslant \frac{(m+1)^2(m+2)+(m+1)(m+2)}{6}$. Let M be the transition matrix of this Markov process[15], and let $e(j)$ be the column matrix of size $\frac{(m+1)^2(m+2)+(m+1)(m+2)}{6} \times 1$ with j-th entry equal to 1 and other entries zero. It follows that the upper bound is given by:

$$U = \sum_{k=1}^{k=\infty} k \cdot e' \left(\frac{(m+1)^2(m+2) + (m+1)(m+2)}{6} \right) \cdot [M^k - M^{k-1}] \cdot e(1) \quad (13)$$

Similar to the lower bound, the time and space complexity of computing the upper bound U is mainly determined by the size of the matrix M which only depends on m, and in this case, is given by $O(m^6)$. Using the same strategy as used in the computation of the lower bound, we avoided explicit computation of the whole matrix M, and managed to determine U in the time and space complexity of $O(m^3)$. On a Pentium 4 PC with 2.4GHz processor and 512MB RAM, the highest runtime and space consuming computation ($N = 200000, n = 100, m = 65$) in our plots required less than 50 second and 2.1MB of memory. For more implementation details, the reader is referred to [29].

[15] Details of transition matrix are given in Appendix B.

5.3 Comparison of lower and upper bounds with simulated results

In this section, we will compare the lower and upper bounds derived in this paper to the number of observations required by a simulation of the hitting set attack. Note that the upper bound shown here is strict, that is, it includes the additional term ϵ calculated in the Appendix C in equation 21.

We will consider three cases with the nominal parameters given by: (a) N = 20000, n = 50, m=20, (b) N = 400, n = 10, m = 10, (c) N = 200,000, n = 100, m = 40. Case (a) represents a typical case, while cases (b) and (c) represent two extremes of an anonymity providing system working in an open environment. Note that the runtime complexity of the hitting set attack simulation is $O(n^m t \log_2 n)$, where t is the number of observations before the termination of the attack. Its space complexity is linearly bounded by $O(tnm)$. Nevertheless the simulation is much faster in practice than the worst case, e.g. a 2.4GHz Pentium 4 PC with 512 MB of RAM needs 1 to 3 hours for each simulation for hard instances like ($N = 200000, n = 100, m = 65$) or ($N = 400, n = 10, m = 23$) and requires about 80MB of RAM. Thus upper and lower bounds computed in the paper provide a significant advantage over simulating the results. The simulation algorithm as well as a complete derivation of its complexities can be found in [29].

For case (a), Figure 4 shows the number of observations required for the hitting set attack as a function of m, n, and N as the two other parameters are kept fixed. For all three parameters, the lower bound is a fairly good approximation (within 20 observations) of the simulation results, however, the upper bound deviates from the simulation results by as much as 40 observations. The figure also shows the number of observations required for "three-exclusivity". The difference between three exclusivity and the strict upper bound is less than 5. This supports our assumption, made in Section 5.2, of ignoring $p(l)$ for $l \geqslant 4$. Also note that since the simulation results are closer to the lower bound, the attacker would need two exclusive observations to finish the attack in most cases and would need three or more exclusive observations only in a few cases.

For case (b), with $N = 400$, both the upper bound and the lower bound provide a fairly good approximation of the sim-

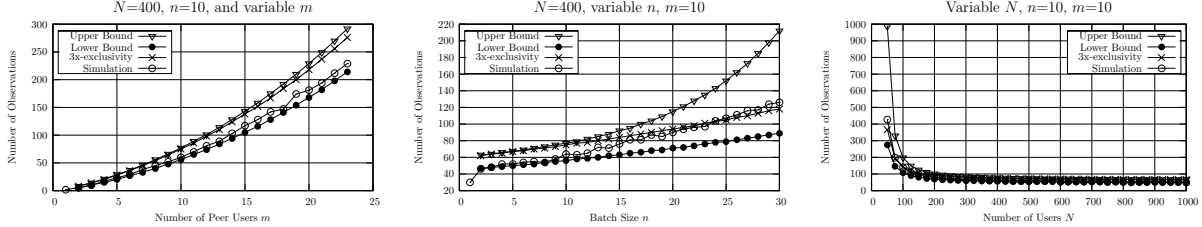

Figure 5. Number of observation required for Case (b)

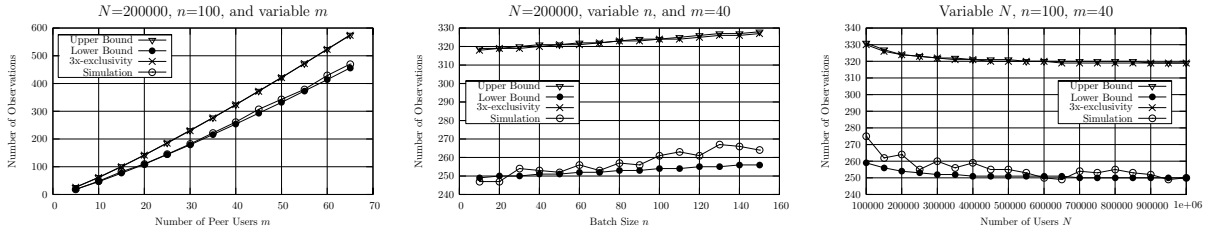

Figure 6. Number of observation required for Case (c)

ulation results except in the case where the batch size n is variable and is large. The loose upper bound in this case is due to the additional term to make the upper bound strict. The three exclusivity calculation in this case provides a good approximation to the simulation results, though it is not a strict upper bound.

For case (c), the lower bound provides a fairly good approximation (within 20 observations), however, the upper bound deviates from the simulation results by as much as 75 observations. In this case, the additional term to make the upper bound strict does not contribute significantly. Therefore, in this case too, the attacker would need two exclusive observations to finish the attack in most cases, and would rarely need three or more exclusive observations.

6 Conclusions

In our work we investigated fundamental protection limits provided by the anonymity sets under certain assumptions (uniform distribution on the recipient sets, static behavior of senders, etc.) on the user behavior. These assumptions are made in such a way that the analysis provides a conservative estimate of the anonymity (i.e. lower bound on the anonymity) provided by the MIXes in the real world. We used an information theoretic setting in which all possible observations of the MIX communication are available to an attacker. Using this setting, we precisely characterized the extent of information inherently contained in anonymity sets. In particular, we showed that if the size of the sender anonymity set is less than the total user population then all communication relationships can be deduced by making a sufficiently large number of observations. We showed that even if the size of the sender anonymity set is equal to the total user population, it is possible to compromise anonymity and deduce important information such as the number of senders that two recipients share by making a sufficiently large number of observations.

In the second setting, we assumed that the attacker makes one observation at a time, and stops when anonymity of a targeted user is compromised. We calculated upper and lower bounds on the number of observations required by the attacker. These bounds are in good agreement with the simulation results for a variety of MIX parameters. In the process of deriving the bounds, we illuminated various structures of anonymity sets. For example, we showed that to compromise anonymity the attacker needs to observe each recipient of a targeted user at least twice, but rarely more than three times for the typical range of MIX parameters.

The knowledge of inherent structures in the anonymity sets can be used for designing MIXes or while using MIXes. Alice can use our analysis (and her software), so that she will never communicate too frequently to a set of peer partners. Either she stops sending messages to the mentioned set or she sends dummy messages to other users outside the set and increases m. The MIX designer or operator can choose appropriate values of anonymity set sizes n according to N and the observed usual user patterns. Our analysis also shows that the (in)security is directly related to the occurrence of the "two time exclusivity" as proven in the lower bound analysis in Claim 1. Thus, the protection of Alice's anonymity highly depends on other users traffic. If the system is able to hinder the occurrence of two time exclusivity, then the system is secure against the hitting set attack, which requires the least number of observations to disclose all peer partners of Alice.

In the future, we plan to repeat the cycle from theory to application and back to theory to refine our models to include our experience with the real traffic and to apply our theoretic models in other environments, e.g. companies, districts of a city, different applications (email).

References

[1] D. Agrawal, D. Kesdogan, S. Penz: "Probabilistic Treatment of MIXes to Hamper Traffic Analysis", *Proceedings of the 2003 IEEE Symposium on Security and Privacy*, Oakland, USA, May, 2003.

[2] A. Back, U. Möller, A. Stiglic: "Traffic Analysis Attacks and Trade-Offs in Anonymity Providing Systems", *In the Proceedings of Information Hiding Workshop* (IH 2001), Springer-Verlag, LNCS 2137, April 2001, pages 245-257.

[3] O. Berthold, H. Federrath, and S. Köpsell, "Web MIXes: A System for Anonymous and Unobservable Internet Access", *International Workshop on Design Issues in Anonymity and Unobservability*, Springer-Verlag, LNCS 2009, 2001.

[4] O. Berthold, A. Pfitzmann, R. Standtke: "The disadvantages of free MIX routes and how to overcome them", *In the Proceedings of Designing Privacy Enhancing Technologies: Workshop on Design Issues in Anonymity and Unobservability*, Springer-Verlag, LNCS 2009, 2001.

[5] D. Chaum, "Untraceable electronic mail, return addresses and digital pseudonyms", *Communications of the A.C.M.*, 24(2):84-88, February 1981.

[6] D. Chaum, "The dining cryptographers problem: Unconditional sender and recipient untraceability", *Journal of Cryptology*, 1:65-75, 1988.

[7] L. Cottrell: "Mixmaster", http://www.obscura.com/ loki/.

[8] G. Danezis: "Statistical Disclosure Attacks: Traffic Confirmation in Open Environments", *Proceedings of Security and Privacy in the Age of Uncertainty, (SEC2003)*, Athens, May, IFIP TC 11, Kluwer 2003.

[9] G. Danezis: "Better Anonymous Communications", *Ph.D. thesis*, University of Cambridge, July 2004.

[10] G. Danezis: "The Traffic Analysis of Continuous-Time Mixes", *Proceedings of Privacy Enhancing Technologies workshop (PET 2004)* Toronto, Canada, May, 2004, Springer-Verlag, LNCS.

[11] G. Danezis, R. Dingledine, N. Mathewson: "Mixminion: Design of a Type III Anonymous Remailer Protocol", *Proceedings of the 2003 IEEE Symposium on Security and Privacy*, Oakland, CA, May 2003.

[12] G. Danezis, A. Serjantov: "Statistical Disclosure or Intersection Attacks on Anonymity Systems", *Proceedings of 6th Information Hiding Workshop (IH 2004)*, Toronto, Canada, May, 2004, Springer-Verlag, LNCS.

[13] C. Díaz, A. Serjantov: "Generalising Mixes", *Proceedings of Privacy Enhancing Technologies workshop (PET 2003)*, Berkeley, CA, March 2003, Springer-Verlag, LNCS 2760.

[14] J. Douceur: "The Sybil Attack", *Proceedings of the 1st International Peer To Peer Systems Workshop (IPTPS 2002)*, Cambridge, MA, USA, March 7-8, 2002, Springer-Verlag, LNCS 2429.

[15] C. Gülcü and G. Tsudik: "Mixing E-mail with BABEL", *In Symposium on Network and Distributed Systems Security (NDSS '96)*, San Diego, California, February 1996.

[16] A. Jerichow, J. Müller, A. Pfitzmann, B. Pfitzmann, M. Waidner: "Real-Time Mixes: A Bandwidth-Efficient Anonymity Protocol", *IEEE Journal on Selected Areas in Communications*, 1998.

[17] D. Kesdogan, D. Agrawal and S. Penz: "Limits of Anonymity in Open Environments", *IH 2002, 5th international workshop on information hiding*, Noordwijkerhout, The Netherlands, 7–9 October 2002. Lecture Notes in Computer Science, Springer-Verlag, 2002

[18] D. Kesdogan, J. Egner, and R. Büschkes, "Stop-and-go mixes providing probabilistic security in an open system", In David Aucsmith, editor, *Information Hiding: Second International Workshop*, Springer-Verlag, LNCS 1525, pages 83-98, Germany, 1998.

[19] D. Kesdogan, L. Pimenidis: "The Hitting Set Attack on Anonymity Protocols", *Proceedings of 6th Information Hiding Workshop (IH 2004)*, Toronto, Canada, May, 2004, Springer-Verlag, LNCS.

[20] B. N. Levine, M. K. Reiter, C. Wang, M. K. Wright: "Timing Attacks in Low-Latency Mix-Based Systems", *In the Proceedings of Financial Cryptography (FC '04)*, Springer-Verlag, LNCS 3110, February 2004.

[21] N. Mathewson, R. Dingledine: "Practical Traffic Analysis: Extending and Resisting Statistical Disclosure", *Proceedings of Privacy Enhancing Technologies workshop (PET 2004)*, Toronto, Canada, May, 2004, Springer-Verlag, LNCS.

[22] Merriam-Webster's Online Dictionary, http://www.m-w.com/dictionary.htm

[23] S. J. Murdoch, G. Danezis: "Low-Cost Traffic Analysis of Tor", *In the Proceedings of the 2005 IEEE Symposium on Security and Privacy*, May 2005.

[24] R. E. Newman, V. R. Nalla, I. S. Moskowitz: "Anonymity and Covert Channels in Simple Timed Mix-firewalls", *Proceedings of Privacy Enhancing Technologies workshop (PET 2004)* Toronto, Canada, May, 2004, Springer-Verlag, LNCS.

[25] M. A. Padlipsky, D. W. Snow, P. A. Karger: "Limitations of End-to-End Encryption in Secure Computer Networks", *ESD-TR-78-158*, August 1978, The MITRE Corporation: Bedford MA, HQ Electronic Systems Division: Hanscom AFB, MA.

[26] A. Pfitzmann: "Diensteintegrierende Kommunikationsnetze mit teilnehmerberprfbarem Datenschutz", *Springer*, ISBN: 3540523278, (in German).

[27] A. Pfitzmann and M. Köhntopp, "Anonymity, Unobservability, and Pseudonymity - A Proposal for Terminology", *Designing Privacy Enhancing Technologies*, Springer-Verlag, LNCS 2009, USA, July 2000.

[28] A. Pfitzmann and M. Waidner, "Networks without user observability, design options. In Advances in Cryptology", *Eurocrypt '85, volume 219 of Lecture Notes in Computer Science*, Spinger-Verlag, 1985.

[29] V. Pham. "Analysis of Attacks on Chaumian MIXes", Master's Thesis, RWTH-Aachen, Germany (in preperation), 2005.

[30] J.-F. Raymond: "Traffic Analysis: Protocols, Attacks, Design Issues, and Open Problems", *Proceedings of Designing Privacy Enhancing Technologies*, Berkley, CA, July 2000, Springer-Verlag, LNCS 2009.

[31] A. Serjantov, P. Sewell: "Passive Attack Analysis for Connection-Based Anonymity Systems", *In the Proceedings of ESORICS 2003*, October 2003.

[32] S. Steinbrecher, S. Köpsell: "Modelling Unlinkability", *Proceedings of Privacy Enhancing Technologies workshop (PET 2003)*, Dresden, Germany, March, 2003, Springer-Verlag, LNCS 2760.

[33] M. Wright, M. Adler, B. N. Levine, C. Shields: "An Analysis of the Degradation of Anonymous Protocols", *In the Proceedings of the Network and Distributed Security Symposium*, IEEE NDSS '02, February 2002.

[34] M. Wright, M. Adler, B. Levine, C. Shields: "Defending Anonymous Communication Against Passive Logging Attacks", *Proceedings of the 2003 IEEE Symposium on Security and Privacy*, Oakland, USA, May, 2003.

[35] Y. Zhu, X. Fu, B. Graham, R. Bettati, W. Zhao: "On Flow Correlation Attacks and Countermeasures in Mix Networks", *In the Proceedings of Privacy Enhancing Technologies workshop*, (PET 2004), LNCS 3424, Springer-Verlag, May 2004.

A Transition Matrix for the Calculation of Lower Bound

Let $p = (1/N)^{n-1}$ be the probability that the peer partner of Alice is the only recipient in a recipient set. Let $q = (1-(m-1)/N)^{n-1}$ be the probability that a recipient set contains only one peer partner of Alice. Thus $q - p$ is the probability that a recipient set contains exactly one peer partner of Alice along with a non-peer partner and $1 - q$ is the probability that a recipient set contains more than one peer partner of Alice. The state-transition probability of $X(t)$ are given by

$$P\big(X_t = \{m_0', m_1', m_2'\} | X_{(t-1)} = \{m_0, m_1, m_2\}\big)$$
$$= \begin{cases} \frac{m_0}{m}(q-p) & m_0' = m_0 - 1, m_1' = m_1 + 1, m_2' = m_2 \\ \frac{m_0}{m}p & m_0' = m_0 - 1, m_1' = m_1, m_2' = m_2 + 1 \\ \frac{m_1}{m}q & m_0' = m_0, m_1' = m_1 - 1, m_2' = m_2 + 1 \\ \frac{(m_0+m_1)(1-q)+m_2}{m} & m_0' = m_0, m_1' = m_1, m_2' = m_2 \\ 0 & \text{otherwise} \end{cases} \tag{14}$$

B Transition Matrix for the Calculation of Upper Bound

For upper bound, let p and q be the same probabilities as calculated to derive the lower bound. The state transition probability of $X(t)$ is given by:

$$P\big(X_t = \{m_0', m_1', m_2', m_3'\} | X_{(t-1)} = \{m_0, m_1, m_2, m_3\}\big) =$$
$$\begin{cases} \frac{m_0}{m}(q-p) & m_0' = m_0 - 1, \quad m_1' = m_1 + 1, \\ & m_2' = m_2, \qquad m_3' = m_3 \\ \frac{m_0}{m}p & m_0' = m_0 - 1, \quad m_1' = m_1, \\ & m_2' = m_2, \qquad m_3' = m_3 + 1 \\ \frac{m_1}{m}(q-p) & m_0' = m_0, \qquad m_1' = m_1 - 1, \\ & m_2' = m_2 + 1, \quad m_3' = m_3 \\ \frac{m_1}{m}p & m_0' = m_0, \quad m_1' = m_1 - 1, \\ & m_2' = m_2, \quad m_3' = m_3 + 1 \\ \frac{m_2}{m}q & m_0' = m_0, \qquad m_1' = m_1, \\ & m_2' = m_2 - 1, \quad m_3' = m_3 + 1 \\ \frac{m_0+m_1+m_2}{m}(1-q) + \frac{m_3}{m} & m_0' = m_0, \quad m_1' = m_1, \\ & m_2' = m_2, \quad m_3' = m_3 \\ 0 & \text{otherwise} \end{cases} \tag{15}$$

C Justification of the Upper Bound

There is a small probability that even after each peer partner has occurred exclusively at least three times, the hitting set attack could not be completed. In this appendix, we will justify our observation that the expected number of additional observations needed to cover such cases is negligible for most values of parameters N, n, and m.

Let $E(i)$ be the event that the hitting set attack cannot be concluded after observing each peer partner exclusively at least i times. Let $\Omega(i)$ be the expected number of observations required to observe each peer partner at least i times exclusively. Let $\Delta(i) = \Omega(i) - \Omega(i-1)$ be the expected number of additional observations required to observe each peer partners exclusively at least i times after each peer partner has been observed exclusively at least $i-1$ times. It follows that if $E(i)$ occurs, then the attacker needs to observe on the average $\Delta(i+1)$ additional observations, and an upper bound on the hitting set attack is given by:

$$U = \Omega(3) + \Delta(4)\text{Prob}(E(3)) + \Delta(5)\text{Prob}(E(4)) + \Delta(6)\text{Prob}(E(5)) + \ldots \tag{16}$$

Thus, the additional term ϵ omitted in the previous calculations is given by:

$$\epsilon = \sum_{i=3}^{\infty} \Delta(i+1)\text{Prob}(E(i)) \tag{17}$$

We note that $\Delta(i)$ is less than the number of observations required to see each peer-partner exclusively at least once, that is $\Delta(i) \leqslant \Delta(1) = \frac{m}{q}\sum_{j=1}^{m}\frac{1}{j}$, where q is the probability that a recipient set contains only one peer partner of Alice.

Let $I(i)$ be the event that i recipient sets, each with only one peer partner, do not have a common non-peer partner. The event $E(i)$ is likely[16] to occur only if there exists a *bottleneck* peer partner, say

[16]There are pathological cases where k peer partners, $k > 1$ occur in synchronization with k non-peer partners in such a manner that the assertion made here is not true. However, the probability of such pathological cases is extremely small [29].

Bob, such that i or more recipient sets in which Bob occurs exclusively have at least one common non-peer partner. Without loss of generality, assume that the bottleneck peer-partner Bob occurs exclusively in the recipient sets $B_1, B_2, \ldots, B_{i+j}, j \geqslant 0$. Since the probability of not having an intersection increases as the number of sets increases, Prob$(I(i))$ is smaller than the probability that the sets $B_1, B_2, \ldots, B_{i+j}$ do not have a common non-peer partner. In other words, Prob$(I(i))$ is smaller than the probability of Bob not being a bottleneck. A similar reasoning can be done for each of the peer partners of Alice to conclude that $(\text{Prob}(I(i)))^m$ is smaller than the probability that none of the peer partners of Alice is a bottleneck. As a result,

$$\text{Prob}(E(i)) \leqslant 1 - \Big(\text{Prob}(I(i))\Big)^m \tag{18}$$

By using union bound and independence of sets $B_1, \ldots, B_i$, we have,

$$\begin{aligned}\text{Prob}(I(i)) &\geqslant 1 - \sum_{r \in B_1 : r \neq \text{Bob}} \text{Prob}(r \in B_2, \ldots, r \in B_i | r \in B_1) \\ &= 1 - (n-1)\Big(\text{Prob}(r \in B_2 | r \in B_1)\Big)^{i-1} \end{aligned} \tag{19}$$

Probability of a particular recipient occurring in a recipient set is given by:

$$\text{Prob}(r \in B_2 | r \in B_1) = 1 - \left(1 - \frac{1}{N}\right)^{n-1} \tag{20}$$

Putting (17)-(20) together, we get

$$\begin{aligned}\epsilon &\leqslant \Delta(1) \sum_{i=3}^{\infty} \text{Prob}(E(i)) \\ &\leqslant \Delta(1) \sum_{i=3}^{\infty} 1 - \Big(1 - (n-1)\big(1 - (1 - \frac{1}{N})^{n-1}\big)^{i-1}\Big)^m \\ &\approx \Delta(1) \sum_{i=3}^{\infty} 1 - 1 + m(n-1)\big(1 - (1 - \frac{1}{N})^{n-1}\big)^{i-1} \\ &= \Delta(1) m(n-1) \sum_{i=3}^{\infty} \big(1 - (1 - \frac{1}{N})^{n-1}\big)^{i-1} \\ &= \Delta(1) m(n-1) \frac{\big(1 - (1 - \frac{1}{N})^{n-1}\big)^2}{1 - 1 + (1 - \frac{1}{N})^{n-1}} \\ &= \Delta(1) m(n-1) \frac{\big(1 - (1 - \frac{1}{N})^{n-1}\big)^2}{(1 - \frac{1}{N})^{n-1}} \\ &\approx \Delta(1) m(n-1) \frac{\big(\frac{n-1}{N}\big)^2}{(1 - \frac{n-1}{N})} \\ &\approx \Delta(1) \frac{m(n-1)^3}{N(N-n+1)} \end{aligned} \tag{21}$$

For the value of N relatively larger than the values of n and m, the value of ϵ is small. For typical parameters N, m, and n, the value of ϵ is less than 1 and it can be ignored. However, for low values of N (e.g. $N = 400$), and high values of m and n (e.g. $n = 30$, and $m = 10$), the value of ϵ is not negligible and needs to be added to get a provable upper bound. In Figure 5, the extra term ϵ contributes to the looseness of upper bound. It is possible to calculate tighter, more precise estimates of ϵ [29], however, details of such calculations are outside the scope of this paper.