Anonymity of a Buffer Constrained Chaum Mix: Optimal Strategy and Asymptotics

Abhishek Mishra
Dep. of Electrical & Computer Eng.
Lehigh Univeristy
PA, USA
Email: abm210@lehigh.edu

Parv Venkitasubramaniam
Dep. of Electrical & Electronics Eng.
Lehigh University
PA, USA
Email: parv.v@lehigh.edu

***Abstract*—As networked systems increasingly pervade every facet of life, it is quintessential for users to communicate without revealing their identities or the paths of data flow. Chaum Mixes are intermediate nodes or routers that are used to provide anonymity by using cryptographic and batching methods to hide source identities. The anonymity achievable by batching strategies, are however, severely impacted by limited buffer capacity of the mix node. This paper presents an information theoretic investigation of a buffer constrained mix, and provides the first single letter characterization of the maximum achievable anonymity as a function of buffer size for a mix serving two users with equal arrival rates. For two users with unequal arrival rates the anonymity is expressed as a solution to a series of finite recursive equations. For more than two users and arbitrary arrival rates, a lower bound on the convergence rate of anonymity is derived as buffer size increases and it is shown that under certain arrival configurations the lower bound is tight.**

I. Introduction

Concealing information in datagram networks goes well beyond encrypting communicated data; hiding the identities of communicating parties is equally critical [1], [2]. Using the knowledge of source-destination pairs or routes of information flow obtained through traffic analysis, an adversary can jam a particular flow, deploy black holes or launch other sophisticated attacks. In this paper, our goal is to study the protection against retrieval of such "networking information" from traffic analysis by appropriate design of scheduling policies for intermediate nodes in a datagram network. In particular, we focus on the protection against the analysis of transmission timing to retrieve information.

On the Internet, Chaum Mixes are used to prevent information retrieval from timing [3]. Mixes are proxy servers or special routers that receive packets from different sources and employ packet padding and layered encryption to ensure packets in the outgoing packets are content-wise "indistinguishable", thus making it infeasible for an eavesdropper to retrieve any information about the communicating parties from the contents and structure of the packets. Importantly, mixes reduce the correlation between the timing of incoming and outgoing packets by randomly reordering the arrived packets prior to transmission. In his original design, David Chaum proposed a mixing strategy [3] that requires the mix to wait until at least one packet arrives from u different users before transmitting them all together.

As is expected, memory limitations on a mix reduce the anonymity provided by any mixing strategy. In networks, it is not practical to use high memory intermediate nodes, hence optimizing anonymity under limited memory is essential. Although many batching strategies have been proposed after Chaum's work [4], [5] to mitigate the effect of memory constraints, the optimal mixing strategy for a buffer constrained mix remains unknown. In fact, there are some fundamental questions in this regard that need to be answered. What is the maximum achievable anonymity as a function of the buffer size of a mix? How does the achievable anonymity vary with number of users and arrival rates? As buffer sizes increase, does the anonymity converge to the maximum achievable, if so, at what rate? In this work, we investigate these issues using an information theoretic approach, where Shannon entropy is used to measure the anonymity achieved. The analytical platform was proposed in [6] for a delay constrained system, where the optimal derivative of anonymity was characterized under light traffic conditions. In [7], the model was used to show that a little relaxation of the *fair* First Come First Serve policy can achieve significant improvement in anonymity.

Some aspects of anonymity under buffer constraints was done previously in [8]. Specifically, we proposed an asymptotically optimal mixing strategy and quantified the anonymity of a single destination mix network, however, the optimal mixing strategy and convergence rate of anonymity is as yet unknown. This is the first work to shed light on optimal mixing under buffer constraints. It is important to note that Chaum mixes rely mainly on random reordering of arrived packets that is a different approach for providing secrecy compare to timing channels where secret information is encoded using inter-arrival time of packets [9]. The approach we use in this paper to derive the single letter characterization of the maximum anonymity as a function of the buffer size relies on a reduction of the problem to a Markov Decision Process and solving the resulting recursion. Further, a network extension is used to prove that as the buffer size k increases the maximum achievable anonymity converges to the prior entropy at a rate $\frac{1}{k^2}$ for a range of multi user arrival configurations.

This work was supported in part by the National Science Foundation through grants CNS-1117701 and CCF-1149495.

II. System Model

Consider a mix that receives packets from u users and transmits them on a single outgoing link. Since packets arriving to the mix are padded and encrypted in layers to prevent information retrieval using packets' size or contents, it is reasonable to assume that in a mix network, all the packets are of equal length [10]. The system model consists of the following.

Chaum Mix: The mix receives packets from u users and reorders them while satisfying the desired memory restriction. The mix can store a maximum of k packets in its buffer. The buffer restriction is critical for the practicality of the system; furthermore, an unlimited buffer would result in the mix waiting indefinitely long to make a decision about transmission [7]. The mix has access to private randomness to reorder the packets, the realization of which is unknown to the eavesdropper. The mix is not allowed to drop any packets; if the buffer of the mix is full and a new packet arrives then the mix necessarily has to transmit at least one packet.

Arrival Process: The arrival of packets to the mix is modelled using a marked temporal point process whose realization consists of arrival times $\{t_j\}$, $t_j \in \mathbb{R}$ and users index $\{U_j\}$, $U_j \in \{1, \cdots, u\}$ of packets where $j = 1, 2, \cdots$. For analytical characterization of anonymity, we assume that arrivals are independent of each other and the probability of multiple arrivals in an infinitesimal time interval is vanishing. The Poisson process is a subclass of these arrival processes.

Departure Process: The departure of packets from the mix can also be viewed as a marked temporal point process. It is important to note that the knowledge of the mixing strategy, memory restriction and arrival process can be used to correlate the outgoing packets with the incoming streams. The primary goal of a mix is to minimize this correlation.

Eve: Eve observes the arrival and departure processes; the packets on the departing stream of mix are indistinguishable to her using their contents or sizes. We assume that Eve knows the mix's strategy of reordering the arrived packets. This assumption may not always be true but it captures the worst case scenario and hence, provides a benchmark for the anonymity. It is worth noting that even though Eve knows the mix's strategy, she does not know the random realization of the strategy which is eventually responsible for the anonymity in the system. Given her knowledge, Eve's goal is to determine the sources of outgoing packets.

Anonymity: Let $\{Y_n\}$ represent a random process corresponding to the sources of the departing packets from Eve's perspective. In other words, Eve's complete observation, denoted by Φ, and her knowledge of the strategy results in an a posterior distribution of $\{Y_n\}$ from Eve's perspective. We define the normalized expected entropy rate of $\{Y_n\}$, given Eve's complete observation, as our metric for measuring the anonymity. Specifically

$$\mathcal{A} = \lim_{n \to \infty} \frac{\mathbb{E}[H(Y_1, \cdots, Y_n | \Phi)]}{n} \tag{1}$$

Here, we take the expectation over arrival process and entropy is calculated for each such realization of the arrival process. The above definition of anonymity captures all the knowledge of Eve. This metric, defined previously in [8], is used in this work to characterize the maximum achievable anonymity.

III. Optimal Anonymity for two user case

In this section, we investigate anonymity for a Chaum mix serving two users, henceforth referred to as *red* and *blue* for convenience, and through the process will obtain the optimal mixing strategy.

Prior to describing the mixing strategy, it is important to note a key reduction that simplifies the class of mixing strategies without losing generality. Specifically, we consider only those strategies where a packet is transmitted only upon a new arrival to a full buffer. To understand that this simplification does not lose generality, consider any strategy that does not belong to this class. Such a strategy would make a decision to transmit a packet when the buffer is not full, or without being triggered by an arrival. Consider a modification of this strategy, wherein the the choice of packet to transmit is made at times identical to the original strategy; however the actual transmission occurs only when the buffer is full and a new packet arrives. Note that for the modified strategy, the departure process is a delayed version of the arrival process, thus providing no information about the packet choices made by the mix. Since anonymity depends on the departure process and the decision making strategy (as known to Eve), the anonymity achieved by the modified strategy can only be better than the original strategy which provides additional information through the departure process (conditioning reduces entropy). Further note that when the memory of a mix is limited, it is sufficient for Eve to know the sequence of arriving packets in place of the complete timing information of the arrival point process; a packet can wait in the buffer until the next packet arrives regardless of what time it arrives, so all that matters to Eve is the state of the buffer and the source of the next arriving packet. Therefore, without loss of generality, we will assume that Eve's observation is restricted to the incoming sequence of packets and we will consider only those mixing strategies in which mix transmits packets *iff* its buffer is full and a new packet arrives. We use the above reduction to obtain the following result:

Theorem 1: The maximum achievable anonymity $\mathcal{A}(k)$ of a Chaum mix with buffer capacity k serving two users

1. when arrival rate of both users are equal is
$$\mathcal{A}(k) = 1 + \log_2\left(\cos\left(\frac{\pi}{k+3}\right)\right) \tag{2}$$
2. when the arrival rate of users are $q\lambda$ and $(1-q)\lambda$ for $q \in [0, 1]$ is
$$\mathcal{A}(k) = t^{k+2} \log_2\left(\frac{\psi_k}{\psi_{k+1}}\right)$$

where ψ_ns are obtained by solving following equations

$$\begin{aligned} \psi_0 &= 1 \\ \psi_n^{t^n} + \psi_{n-2}^{t^n} &= \left(\frac{\psi_k}{\psi_{k+1}}\right)^{t^{k+2}} \psi_{n-1}^{t^n} \;\; \forall 2 \leq n \leq k+1 \\ \psi_1^{\frac{1}{t^{k+1}}} \psi_{k+1} &= \psi_k \end{aligned} \tag{3}$$

where $t = \frac{q}{1-q}$.

Proof: The mix is defined to be in state r if it has r red packets in its buffer including the new arrival. Due to the reduction in strategies wherein the mix transmits only upon a new arrival, the state of the mix transitions only upon a new arrival. Let S_n represent the mix's state at the n^{th} arrival. This definition of state assumes that it is sufficient to combine a new arrival into the buffer of the mix, although, the eavesdropper observes a new arrival separately. This reduction is justified (and does not lose optimality) since whichever packet the mix chooses to transmit (red or blue) the transition to the subsequent state depends only on the total composition of the buffer including the new arrival, and not each individual component. This is by virtue of the independence in arrival processes.

Recall that anonymity is defined as

$$\begin{aligned} \mathcal{A}(k) &= \lim_{n\to\infty} \frac{\mathbb{E}[H(Y_1, \cdots, Y_n|\Phi)}{n} \quad (4) \\ &\overset{(a)}{=} \lim_{n\to\infty} \frac{\mathbb{E}[H(Y_1|\Phi)] + \cdots + \mathbb{E}[H(Y_n|Y_1^{n-1}, \Phi)]}{n} \quad (5) \\ &\overset{(b)}{=} \lim_{n\to\infty} \frac{\mathbb{E}[H(Y_1|S_{k+1})] + \cdots + \mathbb{E}[H(Y_n|S_{k+n})]}{n} \quad (6) \end{aligned}$$

(a) follows from the chain rule of entropy.

(b) follows from the observation that $(Y_1, \cdots, Y_{n-1})$ and Eve's observation completely determine the state S_{n+k}, and $Y_n - S_{n+k} - Y_1, \cdots, Y_{n-1}, \phi$ is a Markov string.

From (6), it is clear that anonymity of any strategy for a mix with buffer size k can be written as the average sum of expected entropy of outgoing packets at each state transition. Let $p(r, n)$ denote the probability of transmitting a red packet if the state upon arrival of the $(n+k)^{th}$ packet is r. Then, $H(Y_n|S_{k+n} = r) = h(p(r, n))$. Under any mixing strategy, since each arrival is independent of previous arrivals and actions of the mix, the state transitions follow a Markov chain such as that depicted in Fig 1. Consequently, the optimal actions of the mix in each state, the probability $p(r, n)$, are determined by solving a Markov Decision Process. The following well known result about MDP proves that the optimal strategy is stationary ($p(n, r) = p_r$) and reduces the optimization to solving the corresponding Bellman equation [11].

Lemma 1: If s represent the present state of a Markov process, s_1 represent the next future state of the Markov process, U represent the decision space, $\phi(s, u)$ represent the cost incurred by the decision $u \in U$ at state s, and there exists a constant f and a bounded function ϕ, unique up to an additive constant, satisfying the following optimality equation

$$f + \phi(s) = \max_U \{\phi(s, U) + \mathbb{E}[\phi(s_1)|S_0 = s, U_0 = u]\}$$

Then f is the maximal average-cost and optimal stationary policy is the one that chooses the optimizing u. □

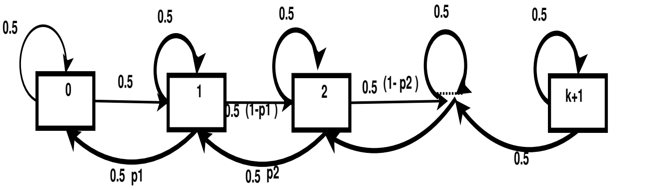

Fig. 1: Markov Chain for the strategy Ψ^* for equal arrival rate

Since ϕ is unique up to an additive constant, without loss of generality, let $\phi(0) = 0$. The optimality equation in Lemma 1 when applied to the MDP for the two user mixing strategy gives rise to the following recursive equations

$$\begin{aligned} w + \phi_0 &= q\phi_1 + (1-q)\phi_0 \quad (7) \\ w + \phi_r &= \max_{p_r}\{h(p_r) + p_r(q\phi_r + (1-q)\phi_{r-1}) + \cdots \\ &\quad (1-p_r)(q\phi_{r+1} + (1-q)\phi_r)\} 1 \leq r \leq k \quad (8) \\ w + \phi_{k+1} &= q\phi_{k+1} + (1-q)\phi_k \quad (9) \end{aligned}$$

By solving the above maximization problem and replacing $\frac{2^{\phi_r}}{2^{\phi_{r-1}}} = \left(\frac{\psi_r}{\psi_{r-1}}\right)^{\frac{1}{q}t^r}$ the result for the general two user system is obtained.

When arrival rates are equal, *i.e.* $q = \frac{1}{2}$, the optimality equation after the maximization and transformation results in the form of following recursion:

$$\psi_{r+1} = c\psi_r - \psi_{r-1} \forall r = 1, \cdots, k$$

with initial conditions $\psi_0 = 1$ and $\psi_1 = \frac{\psi_k}{\psi_{k+1}} \overset{\triangle}{=} c$. The recursion can be solved resulting in $\psi_r = \frac{\sin((r+1)\theta)}{\sin(\theta)}$ where $\theta = \frac{\pi}{k+3}$. The solution to the recursive equation also provides the optimal value of p_r (See equation (8) when $q = \frac{1}{2}$: $p_r = \frac{\psi_{r-1}}{\psi_{r-1}+\psi_{r+1}}$. Consequently, the optimal mixing strategy (Ψ^*) to achieve the maximum anonymity is as follows.

1. If the mix has r red packets in its buffer and a red packet arrives, then the mix transmits a red packet with probability $p_r = \frac{\sin((r+1)\theta)}{\sin((r+1)\theta)+\sin((r+3)\theta)}$
2. if the mix has r red packets in its buffer and a blue packet arrives, then the mix transmits a red packet with probability $q_r = \frac{\sin(r\theta)}{\sin(r\theta)+\sin((r+2)\theta)}$

□

It is important to note that Lemma 1 does not depend on the initial state of a Markov chain, therefore, the anonymity of the optimal strategy does not depend on the initial composition of the mix's buffer.

The above theorem represents the first closed form characterization of the maximum anonymity of a buffer limited Chaum mix, and provides the optimal strategy to achieve the maximum anonymity. When $q \neq 0.5$, the resulting series of nonlinear equations are hard to solve, and consequently a single letter characterization is as yet unknown. However, numerical analysis and algebraic reduction can be used in some special cases. For instance, when $k = 1$, and $q = \frac{1}{n}$ where $n \in \mathbb{N}$, then anonymity $\mathcal{A}(1) = -q\log_2(1-t)$ where t is the unique solution of polynomial $x^{n-1} + x - 1 = 0$ which lies in $(0, 1)$.

While the analytical method can be extended to systems with more than two users, the resulting Markov chain would be multidimensional and the optimizing variables would increase exponentially, consequently limiting the analytical tractability of the problem. Although the closed form characterization for the optimum anonymity is as yet unknown for a general multi user configuration, in the subsequent sections we study the asymptotic behavior of anonymity with buffer size, and demonstrate that for several multi user systems, the optimal convergence rate of the anonymity is $O(1/k^2)$.

We know from Theorem 1, that for a two user system with equal arrival rates, $\mathcal{A}(k) = 1 + \log_2(\cos \frac{\pi}{k+3}) \geq 1 + \log_2 \left(1 - \frac{\pi^2}{2(k+3)^2}\right) \geq 1 - \frac{\pi^2}{2\ln(2)(k+3)^2}$. Therefore, the optimal mixing strategy Ψ^* achieves a convergence rate of $\frac{1}{k^2}$ (to the prior source entropy of 1). In the next section, we study the optimal convergence rate of anonymity, and show that for any number of users and any proportion of arrival rates the convergence rate of anonymity can be no better than $O\left(\frac{1}{k^2}\right)$.

IV. Asymptotic Anonymity: Lower Bound on Convergence Rate

From the theory of typical sequences, it is easy to see that $\mathcal{A} \leq h(p_1, \ldots, p_i, \ldots, p_u)$ where p_i denotes the probability of an incoming packet from the i^{th} user and $h(p_1, \ldots, p_i, \ldots, p_u)$ represents the Shanon entropy:

$$h(p_1, \ldots, p_i, \ldots, p_u) = -\sum_{i=1}^{u} p_i \log_2(p_i).$$

Equality is possible only if $\{Y_n\}$'s are independent of each other from Eve's perspective which is achievable only if the mix is unconstrained in buffer size (A formal proof of this is omitted here due to paucity of space). In [12], we demonstrate that the proportional method of scheduling achieves this limit asymptotically in the buffer size, with a rate of convergence $O(1/k)$. This convergence rate, as we observed in the equal rate two user system in Section III, is not optimal. In the following theorem, we prove that for two user system with unequal arrival rates, the convergence rate of anonymity of Chaum mix is lower bounded by $O(1/k^2)$.

Theorem 2: For a Chaum mix with buffer capacity k, serving two users with arrival rates $q\lambda$ and $(1-q)\lambda$, the anonymity $\mathcal{A}(k)$ is

$$\mathcal{A}(k) \leq h(q) - O\left(\frac{1}{k^2}\right)$$

Proof: Without loss in generality let $q \leq 1 - q$. Consider a sequence of $k^2 + k$ departures. Due to the buffer limitation of the mix, at least k^2 of those departures would have been chosen from the $k^2 + k$ arrivals that triggered the departures.

If among these $k^2 + k$ arrivals, the number of red packets is less than $(k^2 + k)q - 3k$, then regardless of the sources of previous arrivals and actions of the mix, the $k^2 + k$ departures can have no greater than $(k^2 + k)q - k$ red packets. The mix can transmit $(k^2 + k)q - k$. red packets, only if it has k red packets in its buffer before the k^2+k arrivals and k red packets arrive after the k^2+k arrivals. In all other cases, the departures have less than $(k^2 + k)q - k$ red packets.. Using the Gaussian approximation of Binomial distribution, we can show that the probability of such an event is lower bounded by a constant c independent of k.

As Eve observes all arrivals, she is aware, regardless of the mixing strategy, that the number of red packets in the $k^2 + k$ departures is less than $(k^2 + k)q - k$. Consequently, for any mixing strategy, the entropy of the $k^2 + k$ departures is upper bounded by the logarithm of the total possible sequences of sources given the knowledge. Specifically,

$$H(X_1, \ldots, X_{k^2+k}) \leq \log_2 \left(\sum_{t=0}^{k} \binom{k^2 + k}{(k^2 + k)q - t} \right) \tag{10}$$

Using Chernoff bound for Binomial distribution, we can prove that

$$\log_2 \left(\sum_{t=0}^{k} \binom{k^2 + k}{(k^2 + k)q - t} \right) \leq (k^2 + k)h(q) - O(1)$$

Therefore, even if all other $k^2 + k$ departures have optimal anonymity $h(q)$, we know that

$$\begin{aligned} \mathcal{A}(k) &\leq \tfrac{(1-c)(k^2+k)h(q) + c((k^2+k)h(q) - O(1))}{k^2+k} \\ &= h(q) - O\left(\tfrac{1}{k^2}\right) \end{aligned}$$

□

We use the above theorem to prove that the convergence rate of anonymity for any multiuser system is also lower bounded by $O(1/k^2)$.

Theorem 3: For a Chaum mix serving u users with arrival rates $\lambda_1, \cdots, \lambda_u$, the upper bound on anonymity $\mathcal{A}(k)$ is

$$\mathcal{A}(k) \leq h(p_1, \ldots, p_u) - O\left(\frac{1}{k^2}\right)$$

where $p_i = \frac{\lambda_i}{\sum_i \lambda_i} \; \forall \; i = 1, \ldots, u$

Proof: Without loss of generality assume that $p_1 \leq \sum_{i \neq 1} p_i$. Let Z_i denote the indicator random variable (from Eve's perspective) that determines weather i^{th} departure belongs to source 1 or not. Specifically,

$$Z_i = \begin{cases} 1 & Y_i = 1 \\ 0 & Y_i \neq 1 \end{cases} \tag{11}$$

Since Z_i is a deterministic function of Y_i,

$$\begin{aligned} H(Y_1^{k^2+k}) &= H(Z_1^{k^2+k}) + H(Y_1^{k^2+k}|Z_1^{k^2+k}) - H(Z_1^{k^2+k}|Y_1^{k^2+k}) \\ &= H(Z_1^{k^2+k}) + H(Y_1^{k^2+k}|Z_1^{k^2+k}) \end{aligned} \tag{12}$$

Given $Z_1^{k^2+k}$, the remaining entropy is equivalent to the achievable anonymity of a $u-1$ user system. Therefore,

$$H(Y_1^{k^2+k}|Z_1^{k^2+k}) \leq (k^2 + k)(1 - p_1)h(\frac{p_2}{1 - p_1}, \ldots, \frac{p_u}{1 - p_1}) \tag{13}$$

From Theorem 2, we know that

$$H(Z_1^{k^2+k}) \leq (k^2 + k)h(p_1) - O(1) \tag{14}$$

Combining (12)-(14), the theorem is proved. □

V. Optimal Asymptotic Behaviour for Countably Many Configurations

In this section, we show that the $O\left(\frac{1}{k^2}\right)$ convergence in anonymity can be achieved in countably many multi user systems. The main idea behind the approach is to let the mix simulate a network of mixes by dividing the mix's buffer into multiple parts and treat each part as an independent mix. The packets from the users are rerouted into each simulated mix such that a pair of equal rate arrival processes flow into the mix. Each mix employs strategy Ψ^* derived in Theorem 1 for the equal rate two user system, which achieves an anonymity of $1-O(1/k^2)$. We then rely on an extension of the anonymity of a single mix to a single destination network of mixes to prove the convergence rate of the overall simulated system.

Theorem 4: For a Chaum mix, serving u users having arrival rate $\lambda, \lambda, 2\lambda, 2^2\lambda, \ldots, 2^{u-2}\lambda$, anonymity $\mathcal{A}(k)$ is lower bounded by

$$\mathcal{A}(k) \geq 2 - \frac{1}{2^{u-2}} - O\left(\frac{1}{k^2}\right) \tag{15}$$

Proof: Consider a 3 user system with arrival rates in the ratio $1:1:2$ respectively. Let $X_i(t)$ denote the arrival rate of user i. The mix simulates a 2 mix system as shown in Figure 2.

Specifically, the buffer of the mix is divided into two parts with equal capacity $\frac{k}{2}$. Simulated mix M_1 mixes the packets from arrival processes $X_1(t)$ and $X_2(t)$. Since $X_1(t)$ and $X_2(t)$ have equal arrival rates, M_1 employs strategy Ψ^*. This ensures that the outgoing stream $Z(t)$ of the first half has a rate 2λ and achieves an anonymity $\mathcal{A}_1 = 1 - O\left(\frac{1}{k^2}\right)$.

The simulated mix M_2 stores and schedules packets from the arrival process $X_3(t)$ and the outgoing stream of M_1, $Z(t)$. Both incoming streams to M_2 have arrival rate 2λ, therefore M_2 employs strategy Ψ^* and ensures that the outgoing stream of M_2 achieves an anonymity $\mathcal{A}_2 = 1 - O\left(\frac{1}{k^2}\right)$. Note that the anonymity $\mathcal{A}_2$ achieved by M_2 is with respect to a two source system where $X_3(t)$ and $Z(t)$ denote the arrivals-$X_3(t)$ arrives from a source, whereas $Z(t)$ arrives from a "simulated" source.

For a network of mixes as shown in Figure 2, the anonymity of the network of mixes can be written as a linear functional of the anonymities achieved by each individual mix [8]

$$\mathcal{A}(k) \geq \frac{s_1}{s}\mathcal{A}_1 + \frac{s_2}{s}\mathcal{A}_2 \tag{16}$$

where s_i is the total arrival rate to simulated mix M_i and s is the total arrival rate into the network. For the three user system, using the values of $\mathcal{A}_1$ and $\mathcal{A}_2$ in the above equation results in a net achievable anonymity given by:

$$\mathcal{A}(k) \geq \frac{3}{2} - O\left(\frac{1}{k^2}\right) \tag{17}$$

It is to be noted when dividing the operation of the mix as a network of mixes, that the theorem in [8] assumes that Eve can observe the intermediate stream $Z(t)$. Since this is an internal process of the single mix, the true anonymity achieved would be more than that determined by the linear functional in eq. (17). For purposes of this theorem, this reduced anonymity is sufficient to prove the optimality of the convergence rate.

This idea is easily extended when more than 2 users transmit at rates in the ratio $1:1:2:4:\cdots:2^{u-1}$ by simulating a network of $u-1$ mixes wherein simulated mix M_1 receives arrivals from users 1 and 2 and simulated mix M_i for $i > 2$ receives packets from two streams, one from user $i+1$ and the other from the output of simulated mix M_{i-1}. □

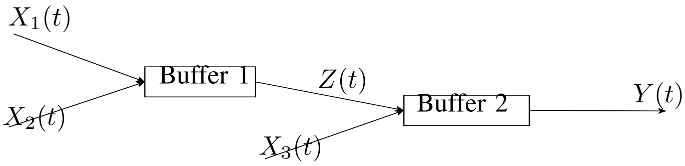

Fig. 2: Representation of strategy for showing the asymptotic of anonymity for three users case

The above theorem represents a particular class of multiuser configurations, where by virtue of dividing the mix's operation as a network of mixes, we proved that the best convergence rate is achievable. This is a general idea that can be used to prove the optimal convergence rate of any user configuration where every mix in the single destination network has equal arrival rates. For instance, in a system where 2^m users transmit at equal rates, a binary tree single destination network can be used to demonstrate optimal convergence. While this approach does not establish the optimality of the convergence rate for any general set of arrival rates, we do believe that an $O(1/k^2)$ convergence is achievable in general.

VI. Conclusion

One of our key contributions in this work is to find the optimal mixing strategy for a Chaum mix under memory restrictions. To the best of our knowledge, this is the first such characterization. Moreover, for some special cases we show that the convergence rate of the optimal anonymity is $O\left(\frac{1}{k^2}\right)$. The extension of this result to a network of mixes, while already investigated in [8], still ignores the possibility of shared randomness across mixes and is a useful area for further study. In addition, the achievable anonymity of a system under dynamic routing configurations is yet another important topic for further investigation.

References

[1] A. Back, U. Moller, and A. Stiglic, "Traffic analysis attacks and tradeoffs in anonymity providing systems," in *Proceedings of 4th International Information Hiding Workshop*, (Pittsburg, PA), April 2001.

[2] N. West, *The SIGINT Secrets: The Signal Intelligence War: 1900 to Today*. New York: William Morrow, 1988.

[3] D. Chaum, "Untraceable electronic mail, return addresses and digital pseudonyms," *Communications of the ACM*, vol. 24, pp. 84–88, February 1981.

[4] L. Cottrell, "Mixmaster and Remailer Attacks," ¡http://www.obscura.com/loki/remailer/remailer-essay.html¿.

[5] C. Díaz and A. Serjantov, "Generalizing mixes," in *Proceedings of Privacy Enhancing Technologies Workshop (PET 2003)*, Springer-Verlag, LNCS 2760, April 2003.

[6] P. Venkitasubramaniam and V. Anantharam, "Anonymity of Mix Networks under Light Traffic Conditions," in *Proceedings of the 36th Allerton Conf. on Communications, Control, and Computing*, (Monticello, IL), October 2008.

[7] A. Mishra and P. Venkitasubramaniam, "Anonymity of an Almost Fair Chaum Mix," in *Proceedings of the 49th Allerton Conf. on Communications, Control, and Computing*, October 2011.

[8] P. Venkitasubramaniam, "Anonymity under Buffer Constraints," in *IEEE International Conference on Communications*, (Cape Town, South Africa), May 2010.

[9] B. Dunn, M. Bloch, and J. Laneman, "Secure bits through queues," in *Networking and Information Theory, 2009. ITW 2009. IEEE Information Theory Workshop on*, pp. 37 –41, june 2009.

[10] M. Reed, P. Syverson, and D. Goldschlag, "Anonymous connections and onion routing," *Selected Areas in Communications, IEEE Journal on*, vol. 16, pp. 482 –494, may 1998.

[11] D. P. Bertsekas, *Dynamic Programming and Optimal Control*, vol. 1. Athena Scientific, 2nd ed., 2001.

[12] A. Mishra and P. Venkitasubramaniam, "Source anonymity in fair scheduling: A case for the proportional method," in *Communications (ICC), 2012 IEEE International Conference on*, pp. 1118 –1122, june 2012.