Towards a Theory of Anonymous Networking

J. Ghaderi and R. Srikant
Department of ECE, and Coordinated Science Lab.
University of Illinois at Urbana-Champaign
{jghaderi, rsrikant}@illinois.edu

Abstract—**The problem of anonymous networking when an eavesdropper observes packet timings in a communication network is considered. The goal is to hide the identities of source-destination nodes, and paths of information flow in the network. One way to achieve such an anonymity is to use mixers. Mixers are nodes that receive packets from multiple sources and change the timing of packets, by mixing packets at the output links, to prevent the eavesdropper from finding sources of outgoing packets. In this paper, we consider two simple but fundamental scenarios: double input-single output mixer and double input-double output mixer. For the first case, we use the information-theoretic definition of the anonymity, based on average entropy per packet, and find an optimal mixing strategy under a strict latency constraint. For the second case, perfect anonymity is considered, and a maximal throughput strategy with perfect anonymity is found that minimizes the average delay.**

I. Introduction

Secure communication has become increasingly important. Privacy and anonymity considerations apply to all components of a communication network, such as contents of data packets, identities of source-destination nodes, and paths of information flow in the network. While a data packet's content can be protected by encrypting the payload of the packet, an eavesdropper can still detect the addresses of the source and the destination by *traffic analysis*. For example, observing the header of the packet can still reveal the identities of its corresponding source-destination pair. *Onion Routing* [1] and *Tor* network [2] are well-known solutions that provide protection against both eavesdropping and traffic analysis. The basic idea is to form an overlay network of Tor nodes, and relay packets through several Tor nodes instead of taking the direct path between the source and the destination. To create a private network, links between Tor nodes are encrypted such that each Tor node only knows the node from which it receives a packet and the node to which it forwards the packet. Therefore, any node in the Tor network sees only two hops (the previous and next nodes) but is not aware of the whole path between the source and the destination, Therefore, a compromised node cannot use traffic analysis to identify source-destination pairs. But Tor cannot solve all anonymity problems. If an eavesdropper can observe the traffic in and out of some nodes, it can still correlate the incoming and outgoing packets of relay nodes to identify the source and the destination or, at least, discover parts of the route between the source and the destination. This kind of statistical analysis is known as timing analysis since the eavesdropper only needs packet timings. For example, in Figure 1, if the processing delay is small, there is a high correlation between output and input processes, and the eavesdropper can easily identify the source of each outgoing packet.

To provide protection against the timing analysis attack, Tor nodes need to perform an additional task, known as mixing, before transmitting packets on output links. A Tor node with mixing ability is called a *Mixer*. In this solution, the Mixer receives packets from multiple links, re-encrypts them, and changes timings of packets, by mixing (reordering) packets at the output links, in such a way that the eavesdropper cannot relate an outgoing packet to its corresponding sender.

The original concept of mix was introduced by Chaum [3]. The mix anonymity was improved by random delaying [4] (Stop-and-go MIXes), and dummy packet transmission [5] (ISDN-MIXes), and used for the various Internet applications such as email [6] and WWW [7](Crowds). Other proposed anonymity schemes are JAP [8], MorphMix [9], Mixmaster [10], Mixminion [11], Buses [12], etc.

However, theoretical analysis of the performance of Chaum mixing is very limited. The information-theoretic measure of anonymity, based on Shannon's equivocation [13], was used in [14] to evaluate the performance of a mixing strategy. The approach of [14] does not take into account the delay or traffic statistics; whereas, modifying packet timings to obfuscate the eavesdropper indeed increases the transmission latency. So, the question of interest is: what is the maximum achievable anonymity under a constraint on delay?

Characterizing the anonymity as a function of traffic load and the delay constraint has been considered in [15]. The authors in [15] have considered a mix with two input links and one output link, where arrivals on the input links are two poisson processes with equal rates, and they characterize upper and lower bounds on the maximum achievable anonymity under a strict delay constraint. The basic idea is that the mixer waits for some time, collects packets from two sources, and sends a batch containing the received packets to the output. The implicit assumption in [15] is that there is no constraint on the capacity of the output link, i.e., the batch can be transmitted instantaneously at the output, no matter how many packets are contained in the batch.

The path between any source-destination pair in an anonymous network contains several nodes; each of which has,

This research was supported by NSF Grant 07-21286 and an ARMY MURI subcontract.

possibly, several input links and several output links. At each node, to perform routing, traffic generated by two or more sources can be merged into one outgoing stream, or the merged stream can be decomposed at several output links for different destinations. To expose the main features of mixing strategies, we focus on two fundamental cases: double input-single output mixer, Figure 1, and double input-double output mixer, Figure 2. Compared to [15], our model considers cases with finite link capacities and derives optimal solutions for certain cases. The remainder of the paper is organized as follows. In section II, the double input-single output mixer is considered, and the optimal mixing strategy is found to maximize the anonymity under a strict latency constraint. Section III is devoted to the double input-double output mixer, where the optimal mixing strategy is found under a constraint on packet drop rate, or transmission rate of dummy packets. Finally, we end the paper with some concluding remarks.

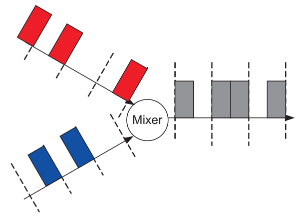


Fig. 1. The double input-single output mixer. The capacity of each input link is 1 packet/time slot and the capacity of the output link is 2 packets/time slot.

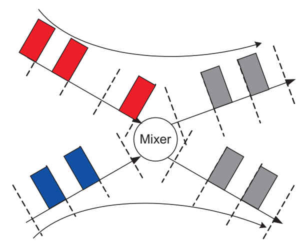


Fig. 2. The double input-double output mixer. The capacity of each link is 1 packet/time slot.

II. DOUBLE INPUT-SINGLE OUTPUT MIXER

Consider Figure 1 where there are two incoming flows, red and blue, and one outgoing link. The capacity of each input link is 1 packet/time-slot, and the capacity of the output link is 2 packets/time-slot. This model ensures that packets do not have to be dropped due to lack of capacity, even when the input links bring in data at maximum rate. Red and blue packets arrive according to i.i.d. Bernouli processes with rates λ_R and λ_B respectively. There is an eavesdropper observing the incoming and outgoing packets. Assume the eavesdropper knows the source of each incoming packet, i.e., its color. This might be made feasible by traffic analysis if the mixer is the first hop of the route or, otherwise, by timing analysis of the previous hop. Given the source of each incoming packet, the eavesdropper aims to identify the source of each outgoing packet, i.e., assign colors, red and blue, to the outgoing stream of packets.

First, consider the case where we do not allow for any delay, i.e., the mixer must send packets out in the same slot in which they arrived. Note that this is possible, without any packet drop, since at most two packets arrive in each slot, and the capacity of the output link is 2 packets/slot. Then, the only way to confuse the eavesdropper is to send out a random permutation of received packets in each slot.

By allowing a strict delay T for each packet, the mixer can do better; it can select and permute packets from the current slot and also from the previous slots up to $T-1$ slots before. Let Ψ_T denote the set of all possible mixing strategies that satisfy the strict delay constraint T. Let the random variable I_k denote arrivals in k-th slot, therefore I_k can be $\emptyset$, R, B, or RB, where they respectively denote the cases of no arrivals, red arrival but no blue arrival, blue arrival but no red arrival, and both red and blue arrivals. Similarly define a random variable O_k for the output sequence such that $O_k \in \{\emptyset, R, B, RB, BR\}$ (Note that ordering of packets at the output matters). Next, we define anonymity of a mixing strategy $\psi \in \Psi_T$, based on the average conditional entropy of the output sequence given the input sequence, as follows.

Definition 1. *The anonymity A^{ψ} of a mixing strategy ψ is defined as*

$$A^{\psi} = \lim_{N\to\infty} \frac{1}{N(\lambda_R+\lambda_B)} H(O_1 O_2 \cdots O_N | I_1 I_2 \cdots I_N).$$

Note that in the above definition, the numerator is the entropy of the output sequence given the input sequence of length N, and the denominator is the average number of red and blue arrivals in N slots. So, as $N \to \infty$, anonymity is the amount of uncertainty in each outgoing packet, bits/packet, observed by the eavesdropper.

Remark 1. *By using the Fano's inequality, the anonymity provides a lower bound for the probability of error in detection incurred by the eavesdropper [16].*

We wish to find the optimal strategy $\psi^* \in \Psi_T$ that maximizes the anonymity. The case of $T = 0$ is trivial since, in this case, the output sequence is i.i.d. as well, and therefore

$$\begin{aligned} \frac{1}{N} H(O_1 O_2 \cdots O_N | I_1 I_2 \cdots I_N) &= \frac{1}{N}\sum_{k=1}^{N} H(O_k|I_k) \\ &= H(O_1|I_1) \\ &= \lambda_R \lambda_B H(O_1|I_1 = RB). \end{aligned}$$

Therefore, to maximize the anonymity, the mixer must send a random permutation of the received packets, in the case of both read and blue arrival, with equal probability to get $H(O_1|I_1 = RB) = 1$. Correspondingly, the maximum anonymity is given by

$$A^{\psi^*} = \frac{\lambda_R \lambda_B}{\lambda_R + \lambda_B}.$$

In the rest of this section, we consider the more interesting case of $T = 1$, where each packet has to be sent out in the current slot or in the next slot. By the chain rule [17], the conditional entropy of the output sequence can be written as

$$H(O_1 \cdots O_N | I_1 \cdots I_N) = H(O_1|I_1 \cdots I_N) + H(O_2|I_1 \cdots I_N, O_1) + \cdots + H(O_N|I_1 \cdots I_N, O_1 \cdots O_{N-1}).$$

For the latency constraint $T = 1$, the right hand side of the equality can be simplified as

$$H(O_1|I_1)+H(O_2|I_1,I_2,O_1)+\cdots+H(O_N|I_{N-1},I_N,O_{N-1}).$$

But

$$H(O_k|I_k,I_{k-1},O_{k-1}) = H(O_k|I_k,I_{k-1},O_{k-1},Q_{k-1}),$$

where $Q_{k-1} := I_{k-1} \backslash O_{k-1}$. Note that Q_{k-1} denotes what has been left in the queue for transmission in the next slot. Noting that O_k is conditionally independent of I_{k-1} and O_{k-1}, given both I_k and Q_{k-1}, the conditional entropy of the output sequence can be written as

$$H(O_1 \cdots O_N | I_1 \cdots I_N) = \sum_{k=1}^{N} H(O_k|I_k, Q_{k-1}),$$

where we defined the initial condition as $Q_0 = \emptyset$. Therefore, maximizing the average entropy of the output sequence is equivalent to

$$\max \lim_{N \to \infty} \frac{1}{N} \sum_{k=1}^{N} E_{q_{k-1}} \left[H(O_k|I_k, q_{k-1}) \right],$$

where $q_{k-1} \in \{\emptyset, R, B, RB\}$ denotes a realization of the random variable Q_{k-1}. This can be viewed as an average reward maximization problem where, at each slot k, the state of the system X_k is the queue at the end of the previous slot, i.e., Q_{k-1}, and the reward of action u_k in state $x_k (= q_{k-1})$ is $c(x_k, u_k) = H(O_k|I_k, q_{k-1})$. Roughly speaking, the action u_k is to randomly select some packets from I_k and q_{k-1}, and send the permutation of the selected packets to the output. Let w denote the maximum value of the above average entropy maximization problem, then, by definition,

$$A^{\psi^*} = \frac{w}{\lambda_R + \lambda_B}$$

and the optimal mixing strategy ψ^* is the one that chooses the corresponding optimal policy for the average entropy maximization problem. In order to solve the problem, next we identify the possible actions for different states which will allow us to define the reward function in more detail and provide an explicit solution.

A. Set of possible actions and corresponding rewards for different states

There is a set of possible actions for each state depending on different arrival types. In the following, we identify the set of actions and their corresponding rewards for each case.

1) $\mathbf{Q_{k-1} = \emptyset}$

(i) $I_k = \emptyset$: In this case, obviously, there will be no transmission at the output link and the queue will remain empty as well, i.e., $O_k = \emptyset$ and $Q_k = \emptyset$. The corresponding entropy is $H(O_k|I_k = \emptyset, Q_{k-1} = \emptyset) = 0$.

(ii) $I_k = R$: Two options are possible; the mixer can queue the arrived packet, with probability α_k, or send the packet in the current slot, with probability $1-\alpha_k$. No matter what the mixer does, the entropy in this slot $H(O_k|I_k = R, Q_{k-1} = \emptyset) = 0$. Correspondingly, the queue is updated as $Q_k = R$, with probability of α_k, or $Q_k = \emptyset$, with probability of $1 - \alpha_k$.

(iii) $I_k = B$: This case is similar to the previous case except that we use β_k instead of α_k. Therefore, $Q_k = B$, with probability β_k, or $Q_k = \emptyset$, with probability $1 - \beta_k$, and $H(O_k|I_k = B, Q_{k-1} = \emptyset) = 0$.

(iv) $I_k = RB$: The mixer has four options; it can queue both packets (with probability $1-s_k$), send both out (with probability $s_k(1-y_k)$), keep only R and send B out (with probability $s_k y_k (1-p_k)$), or keep only B and send R out (with probability $s_k y_k p_k$). Note that the parameters s_k, y_k, and p_k have been used to characterize the probabilities. Intuitively, s_k is the probability that a transmission at the output link happens at all, y_k is the probability of sending only one packet out given a transmission must happen, and p_k is the probability of sending R out given that only one packet is transmitted at the output. Accordingly,

$$H(O_k|I_k = RB, Q_{k-1} = \emptyset) = s_k \left(y_k \mathcal{H}(p_k) + 1 - y_k \right),$$

where $\mathcal{H}$ is the binary entropy function given by

$$\mathcal{H}(p) = -p \log(p) - (1-p) \log(1-p)$$

for $0 < p < 1$.

2) $\mathbf{Q_{k-1} = R}$

(i) $I_k = \emptyset$: The mixer has to send the content of the queue to the output, therefore $O_k = R$, and obviously, $H(O_k|I_k = \emptyset, Q_{k-1} = R) = 0$ and $Q_k = \emptyset$.

(ii) $I_k = R$: The mixer can queue the recent R, with probability γ_k, and send Q_{k-1} to the output, or can send both Q_{k-1} and the recent arrival to the output, with probability $1 - \gamma_k$. Therefore, $Q_k = R$ ($O_k = R$) with probability γ_k, or $Q_k = \emptyset$ ($O_k = RR$) with probability $1 - \gamma_k$. The corresponding entropy will be zero, i.e., $H(O_k|I_k = R, Q_{k-1} = R) = 0$.

(iii) $I_k = B$: Again mixer has two options; it can send a random permutation of R and B to the output, i.e., $Q_k = \emptyset$, with probability a_k, or it can queue the B and send only the R out, i.e., $Q_k = B$, with probability $1 - a_k$. The entropy is $H(O_k|I_k = B, Q_{k-1} = R) = a_k$.

(iv) $I_k = RB$: The mixer has three options; it can queue both arrivals, i.e., $Q_k = RB$, with probability $1-t_k$, keep only the red arrival in the queue, i.e., $Q_k = R$, with probability $t_k(1-d_k)$, or keep only the blue arrival in the queue, i.e., $Q_k = B$, with probability $t_k d_k$. Correspondingly,

$$P(O_k = o_k) = \begin{cases} t_k d_k & ; o_k = RR \\ t_k(1-d_k)/2 & ; o_k = RB \\ t_k(1-d_k)/2 & ; o_k = BR. \end{cases}$$

and

$$H(O_k|I_k = RB, Q_{k-1} = R) = t_k\left(\mathcal{H}(d_k) + 1 - d_k\right).$$

3) $\mathbf{Q_{k-1} = B}$

Since this case is similar to the previous case, the details are omitted for brevity.

(i) $I_k = \emptyset$: Obviously, $H(O_k|I_k = \emptyset, Q_{k-1} = B) = 0$, and $Q_k = \emptyset$.

(ii) $I_k = B$: $H(O_k|I_k = B, Q_{k-1} = B) = 0$. Options are $Q_k = B$, with probability δ_k, or $Q_k = \emptyset$, with probability $1-\delta_k$.

(iii) $I_k = R$: $H(O_k|I_k = R, Q_{k-1} = B) = b_k$. Options are $Q_k = R$, with probability $1-b_k$, or $Q_k = \emptyset$, with probability b_k.

(iv) $I_k = RB$: The mixer can keep both arrivals in the queue, i.e., $Q_k = RB$, with probability $1-z_k$, keep only the red arrival in the queue, i.e., $Q_k = R$, with probability $z_k r_k$, or keep only the blue arrival in the queue, i.e., $Q_k = B$, with probability $z_k(1-r_k)$. The entropy is

$$H(O_k|I_k = RB, Q_{k-1} = B) = z_k\left(\mathcal{H}(r_k) + 1 - r_k\right).$$

4) $\mathbf{Q_{k-1} = RB}$

The mixer has to send the contents of the queue to the output, i.e., $O_k = RB$ or BR with equal probabilities, and queue all the recent arrivals, i.e., $Q_k = I_k$. The entropy is simply $H(O_k|I_k, Q_{k-1} = RB) = 1$.

Recall that the reward function is

$$C(x_k, u_k) = H(O_k|I_k, q_{k-1}) = E_{I_k}\left[H(O_k|i_k, q_{k-1})\right]$$

where i_k denotes a realization of I_k. Therefore, the reward function, and queue updates, for each state are the following.

1) $Q_{k-1} = \emptyset$:

The reward function is given by

$$C(\emptyset, u_k) = \lambda_R\lambda_B s_k\left(y_k\mathcal{H}(p_k) + 1 - y_k\right)$$

and the queue is updated as

$$P(Q_k = q|Q_{k-1} = \emptyset, u_k) = \begin{cases} \lambda_R(1-\lambda_B)\alpha_k + \lambda_R\lambda_B s_k y_k(1-p_k) & ; q = R \\ \lambda_B(1-\lambda_R)\beta_k + \lambda_R\lambda_B s_k y_k p_k & ; q = B \\ \lambda_R\lambda_B(1-s_k) & ; q = RB \\ --- & ; q = \emptyset \end{cases}$$

where we used the notation "$---$" for the probability of having an empty queue, since we will not need the explicit expression of this probability, although, it can be, obviously, derived from the other three probabilities.

2) $Q_{k-1} = R$:

The reward function is given by

$$C(R, u_k) = \lambda_B(1-\lambda_R)a_k + \lambda_R\lambda_B t_k\left(\mathcal{H}(d_k) + 1 - d_k\right)$$

and the queue is updated as

$$P(Q_k = q|Q_{k-1} = R, u_k) = \begin{cases} \lambda_R(1-\lambda_B)\gamma_k + \lambda_R\lambda_B t_k(1-d_k) & ; q = R \\ \lambda_B(1-\lambda_R)(1-a_k) + \lambda_R\lambda_B t_k d_k & ; q = B \\ \lambda_R\lambda_B(1-t_k) & ; q = RB \\ --- & ; q = \emptyset \end{cases}$$

3) $Q_{k-1} = B$:

The reward function is given by

$$C(B, u_k) = \lambda_R(1-\lambda_B)b_k + \lambda_R\lambda_B z_k\left(\mathcal{H}(r_k) + 1 - r_k\right)$$

and the queue is updated as

$$P(Q_k = q|Q_{k-1} = B, u_k) = \begin{cases} \lambda_R(1-\lambda_B)(1-b_k) + \lambda_R\lambda_B r_k z_k & ; q = R \\ \lambda_B(1-\lambda_R)\delta_k + \lambda_R\lambda_B z_k(1-r_k) & ; q = B \\ \lambda_R\lambda_B(1-z_k) & ; q = RB \\ --- & ; q = \emptyset \end{cases}$$

4) $Q_{k-1} = RB$:

The reward function is given by

$$C(RB, u_k) = 1$$

and the queue is updated as

$$P(Q_k = q|Q_{k-1} = RB, u_k) = \begin{cases} \lambda_R(1-\lambda_B) & ; q = R \\ \lambda_B(1-\lambda_R) & ; q = B \\ \lambda_R\lambda_B & ; q = RB \\ --- & ; q = \emptyset \end{cases}$$

B. Optimal stationary mixing strategy

The following Theorem states one of our main results.

Theorem 1. *For the double input-single output mixer, the optimal mixing strategy is the following. At each time k, given Q_{k-1} and I_k, if*

1) $Q_{k-1} = \emptyset$
- $I_k = \emptyset$, R, B: $Q_k = I_k$, $O_k = \emptyset$.
- $I_k = RB$: *send R out with probability p^* or B with probability $1-p^*$*, $Q_k = I_k \backslash O_k$.

2) $Q_{k-1} = R$
- $I_k = \emptyset, R$: $Q_k = I_k$, $O_k = Q_{k-1}$.
- $I_k = B$: *transmit a random permutation of R and B*, $Q_k = \emptyset$.
- $I_k = RB$: *transmit RR with probability d^*, or transmit a random permutation of R and B with probability $1-d^*$*, $Q_k = I_k \backslash O_k$.

3) $Q_{k-1} = B$
- $I_k = \emptyset, B$: $Q_k = I_k$, $O_k = Q_{k-1}$.
- $I_k = R$: *transmit a random permutation of R and B*, $Q_k = \emptyset$.

- $I_k = RB$: *transmit BB with probability r^*, or transmit a random permutation of R and B with probability $1 - r^*$, $Q_k = I_k \backslash O_k$,*

where probabilities p^*, d^*, and r^* depend on arrival rates λ_R and λ_B.

In the special case $\lambda_R = \lambda_B$, $p^* = \frac{1}{2}$, $d^* = \frac{1}{3}$, and $r^* = \frac{1}{3}$.

Proof of Theorem 1: Having formally defined the reward function and the dynamics of the system in subsection II-A, we use the following well-known result to solve the average reward maximization problem [19].

Lemma 1. *Suppose there exists a constant w and a bounded function ϕ, unique up to an additive constant, satisfying the following optimality equation*

$$w + \phi(x) = \max_u \{C(x,u) + E[\phi(x_1)|x_0 = x, u_0 = u\}$$

Then w is the maximal average-reward and the optimal stationary policy is the one that chooses the optimizing u.

Since ϕ is unique up to an additive constant, without loss of generality, assume $\phi(\emptyset) = 0$. Then, for $x = \emptyset$, the optimality equation can be written as

$$\begin{aligned} w \quad = \quad & \max_{s,p,y,\alpha,\beta} \{\lambda_R \lambda_B s \left(y \mathcal{H}(p) + 1 - y\right) \\ & +[\lambda_R(1-\lambda_B)\alpha + \lambda_R\lambda_B sy(1-p)]\phi(R) \\ & +[\lambda_B(1-\lambda_R)\beta + \lambda_R\lambda_B syp]\phi(B) \\ & + [\lambda_R\lambda_B(1-s)]\phi(RB)\}. \end{aligned}$$

Obviously, $\alpha = 1$ and $\beta = 1$ maximize the right hand side if $\phi(R)$ and $\phi(B)$ are nonnegative. We will later see that $\phi(R)$ and $\phi(B)$ are indeed nonnegative. Therefore, the right hand side of the optimality equation can be written as

$$\begin{aligned} & \lambda_R\lambda_B s \left[y\left(\mathcal{H}(p) - 1 + (1-p)\phi(R) + p\phi(B)\right) + 1 - \phi(RB)\right] \\ & \quad +\lambda_R(1-\lambda_B)\phi(R) + \lambda_B(1-\lambda_R)\phi(B) + \lambda_R\lambda_B\phi(RB). \end{aligned}$$

First, consider the term $\mathcal{H}(p) - 1 + (1-p)\phi(R) + p\phi(B)$. This term is maximized by choosing

$$p^* = \frac{1}{1 + 2^{\phi(R)-\phi(B)}}. \tag{1}$$

We will later show that

$$\mathcal{H}(p^*) - 1 + (1-p^*)\phi(R) + p^*\phi(B) \geq 0, \tag{2}$$

and therefore $y^* = 1$. Furthermore, for $y^* = 1$, we will see that the term inside the brackets is always nonnegative, i.e.,

$$\mathcal{H}(p^*) + (1-p^*)\phi(R) + p^*\phi(B) - \phi(RB) \geq 0, \tag{3}$$

and therefore $s^* = 1$. Finally, w is given by

$$\begin{aligned} w \quad = \quad & \lambda_R\lambda_B\mathcal{H}(p^*) + \lambda_R(1-\lambda_B p^*)\phi(R) \\ + \quad & \lambda_B(1-\lambda_R(1-p^*))\phi(B). \end{aligned} \tag{4}$$

Next, consider the optimality equation for $x = R$. It can be written as

$$\begin{aligned} w + \phi(R) \quad = \quad & \max_{\gamma,d,t,a} \{\lambda_B(1-\lambda_R)a + \lambda_R\lambda_B t \left(\mathcal{H}(d) + 1 - d\right) \\ & +[\lambda_R(1-\lambda_B)\gamma + \lambda_R\lambda_B(1-d)]\phi(R) \\ & +[\lambda_B(1-\lambda_R)(1-a) + \lambda_R\lambda_B td]\phi(B) \\ & +\lambda_R\lambda_R(1-t)\phi(RB)\}. \end{aligned}$$

Similar to the argument for $x = \emptyset$, $\gamma^* = 1$, if $\phi(R) > 0$, and $a^* = 1$ if $\phi(B) < 1$. Furthermore, taking the derivative respect to d, setting it to zero, and solving it for d^* yields

$$d^* = \frac{1}{1 + 2^{1+\phi(R)-\phi(B)}}. \tag{5}$$

Finally, $t^* = 1$ if

$$\mathcal{H}(d^*) + 1 - d^* + (1-d^*)\phi(R) + d^*\phi(B) - \phi(RB) \geq 0, \tag{6}$$

and the optimality condition is simplified to

$$\begin{aligned} w + \phi(R) \quad = \quad & \lambda_B(1-\lambda_R) + \lambda_R\lambda_B \left(\mathcal{H}(d^*) + 1 - d^*\right) \\ & +[\lambda_R(1-\lambda_B) + \lambda_R\lambda_B(1-d^*)]\phi(R) \\ & +\lambda_R\lambda_B d^*\phi(B) \end{aligned} \tag{7}$$

Next, consider the optimality equation for $x = B$

$$\begin{aligned} w + \phi(B) \quad = \quad & \max_{\delta,r,z,b} \{\lambda_R(1-\lambda_B)b + \lambda_R\lambda_B z \left(\mathcal{H}(r) + 1 - r\right) \\ & +[\lambda_B(1-\lambda_R)\delta + \lambda_R\lambda_B z(1-r)]\phi(B) \\ & +[\lambda_R(1-\lambda_B)(1-b) + \lambda_R\lambda_B zr]\phi(R) \\ & +\lambda_R\lambda_R(1-z)\phi(RB)\} \end{aligned}$$

In parallel with the argument for $x = R$, $\delta^* = 1$ if $\phi(B) \geq 0$, and $b^* = 1$ if $\phi(R) \leq 1$. Moreover, $z^* = 1$ if

$$\mathcal{H}(r^*) + 1 - r^* + (1-r^*)\phi(B) + r^*\phi(R) - \phi(RB) \geq 0 \tag{8}$$

where

$$r^* = \frac{1}{1 + 2^{1+\phi(B)-\phi(R)}}. \tag{9}$$

The optimality condition is simplified to

$$\begin{aligned} w + \phi(B) \quad = \quad & \lambda_R(1-\lambda_B) + \lambda_R\lambda_B \left(\mathcal{H}(r^*) + 1 - r^*\right) \\ & +[\lambda_B(1-\lambda_R) + \lambda_R\lambda_B(1-r^*)]\phi(B) \\ & +\lambda_R\lambda_B r^*\phi(R) \end{aligned} \tag{10}$$

Finally, the optimality equation for $x = RB$ is given by

$$\begin{aligned} w + \phi(RB) \quad = \quad & 1 + \lambda_R(1-\lambda_B)\phi(R) \\ & +\lambda_B(1-\lambda_R)\phi(B) + \lambda_R\lambda_B\phi(RB) \end{aligned} \tag{11}$$

Therefore, we need to solve equations (4), (7), and (10) to find w, $\phi(R)$, and $\phi(B)$. Then, (11) can be used to find $\phi(RB)$. Eventually, what remains to be shown is that $0 \leq \phi(R), \phi(B) \leq 1$, and, in addition, $\phi(R)$, $\phi(B)$, and $\phi(RB)$ satisfy inequalities (2), (3), (6), and (8).

First, consider the special case of $\lambda_R = \lambda_B = \lambda$. By symmetry, $\phi(R) = \phi(B)$ which yields $p^* = 1/2$ and

$d^* = r^* = 1/3$. Then, by solving equations (4) and (7), we have

$$\phi(R) = \phi(B) = \frac{\lambda^2(\log 3 - 2) + \lambda}{-\lambda^2 + \lambda + 1}$$

and

$$w = \frac{\lambda^2}{-\lambda^2 + \lambda + 1}\left[-\lambda^2(\log 3 - 1) + 2(\log 3 - 2)\lambda + 3\right].$$

Then, the anonymity is $A^{\phi^*} = w/2\lambda$, and it is easy to check that the solutions satisfy all the inequalities.

Next, consider the general case with, probably, unequal arrival rates. We prove that the solutions indeed exist and they satisfy the required conditions. Using (4) to replace w in (7) and (10) yields

$$\phi(R) = \lambda_B[1 - \phi(B)(1 - \lambda_R)] + \lambda_R\lambda_B g(\xi) \quad (12)$$

$$\phi(B) = \lambda_R[1 - \phi(R)(1 - \lambda_B)] + \lambda_R\lambda_B f(\xi) \quad (13)$$

where

$$g(\xi) = (d^* - p^*)(-\xi) + \mathcal{H}(d^*) - \mathcal{H}(p^*) - d^*,$$

$$f(\xi) = (r^* + p^*)\xi + \mathcal{H}(r^*) - \mathcal{H}(p^*) - r^* - \xi,$$

and $\xi = \phi(R) - \phi(B)$.

Therefore, the optimal probabilities can be expressed as functions of ξ by

$$p^* = \frac{1}{1 + 2^{\xi}},\ d^* = \frac{1}{1 + 2^{1+\xi}},\ r^* = \frac{1}{1 + 2^{1-\xi}}.$$

Lemma 2. *The function $g(\xi)$ is an increasing function of ξ and $f(\xi)$ is a decreasing function of ξ (see [18] for the proof).*

For any pair $(\phi(R), \phi(B))$ chosen from $[0, 1]\times[0, 1]$, $-1 \le \xi \le 1$, and therefore, by Lemma 2, functions f and g can be bounded from below and above by

$$g(-1) \le g(\xi) \le g(1),$$

and

$$f(1) \le f(\xi) \le f(-1).$$

but it is easy to check that

$$g(1) = f(-1) = \log(5/3) - 1,\ g(-1) = f(1) = 1 - \log 3,$$

and therefore,

$$-1 < f(\xi), g(\xi) < 0.$$

Consequently, the right-hand sides of (12) and (13) form a continuous mapping from $[0, 1] \times [0, 1]$ to $[0, 1] \times [0, 1]$, and therefore, by the Brouwer fixed point theorem ([20], p. 72), the system of nonlinear equations, (12), (13), has a solution $(\phi(R), \phi(B)) \in [0, 1] \times [0, 1]$.

Next, we show that the solutions indeed satisfy the inequalities. First, we prove that (2) holds. Define

$$\begin{aligned}\psi_1(\xi) &= \mathcal{H}(p^*) + (1 - p^*)\phi(R) + p^*\phi(B)\\ &= \mathcal{H}(p^*) - p^*\xi + \phi(R).\end{aligned}$$

First, consider the case that $-1 \le \xi \le 0$, then

$$\begin{aligned}\frac{d}{d\xi}(\mathcal{H}(p^*) - p^*\xi) &= p^{*\prime}\log\frac{1 - p^*}{p^*} - p^* - p^{*\prime}\xi\\ &= -p^* \le 0.\end{aligned}$$

Hence,

$$\psi_1(\xi) \ge \psi_1(0) = 1 + \phi(R) \ge 1.$$

For the case that $0 \le \xi \le 1$, rewrite $\psi_1(\xi)$ as the following

$$\psi_1(\xi) = \mathcal{H}(p^*) + (1 - p^*)\xi + \phi(B).$$

Then,

$$\frac{d}{d\xi}(\mathcal{H}(p^*) + (1 - p^*)\xi) = 1 - p^* \ge 0,$$

and hence,

$$\psi_1(\xi) \ge \psi_1(0) = 1 + \phi(B) \ge 1.$$

Therefore, for $-1 \le \xi \le 1$, $\psi_1(\xi) \ge 1$, and (2) holds.

Note that from (11), we have

$$\phi(RB) = \frac{1 - \lambda_R\lambda_B\psi_1(\xi)}{1 - \lambda_R\lambda_B} \quad (14)$$

and since (2) holds, we have

$$\phi(RB) \le 1,$$

and consequently (3) will be satisfied as well.

To show (6), note that $\phi(R)+1-\phi(RB) \ge 0$, and therefore, it suffices to prove that

$$\begin{aligned}\psi_2(\xi) &= \mathcal{H}(d^*) - d^* - d^*\phi(R) + d^*\phi(B)\\ &= \mathcal{H}(d^*) - d^*\xi - d^*\end{aligned}$$

is nonnegative. But $\psi_2(\xi)$ is a decreasing function since

$$\begin{aligned}\frac{d}{d\xi}\psi_2 &= d^{*\prime}\log\frac{1 - d}{d} - d^{*\prime}\xi - d^* - d^{*\prime}\\ &= d^{*\prime}(1 + \xi) - d^{*\prime}\xi - d^* - d^{*\prime}\\ &= -d^* \le 0\end{aligned}$$

So $\psi_2(\xi) \ge \psi_2(1) = \mathcal{H}(1/5) - 2/5 = \log 5 - 2 \ge 0$, and consequently (6) follows. (8) is also proved by a similar argument. This concludes the proof of Theorem 1. ■

C. Numerical results

Equations (4), (7), and (10) form a system of nonlinear equations which can be solved numerically, for different values of λ_R and λ_B, by using our algorithm in [18]. Figure 3 shows the maximum anonymity, found by running the algorithm, for different arrival rates. The probabilities p^*, d^*, and r^* of the optimal mixing strategy have been evaluated in figure 4 for different arrival rates λ_R and λ_B.

Remark 2. *The stationary policy does not exist for $\lambda_R = \lambda_B = 1$ since as $\lambda_R \to 1$ and $\lambda_B \to 1$, $\phi(RB) \to -\infty$ (see (14)). This makes sense since, in this case, if we start with initial condition $Q_0 = \emptyset$ and use the strategy specified in Theorem 1, we get an anonymity of $A^{\psi^*} = \log(3)/2$; whereas if the initial condition is $Q_0 = RB$, the only possible strategy*

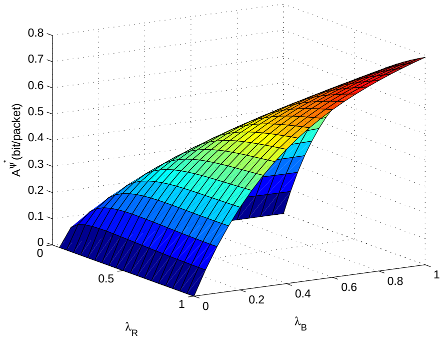
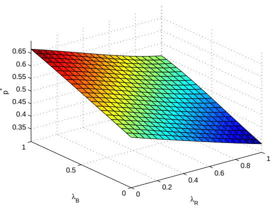
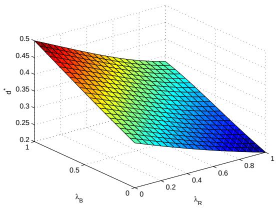


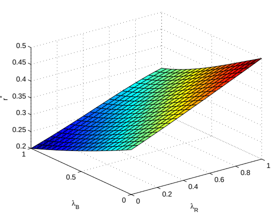
Fig. 3. Anonymity for different values of λ_R and λ_B.



(a) p^*



(b) d^*



(c) r^*

Fig. 4. Probabilities p^*, d^*, and r^* for different arrival rates λ_R and λ_B.

will be to transmit the contents of the queue, and queue the arrived RB in each time slot. This yields an anonymity of $1/2$ bit/packet. Therefore, the optimal strategy depends on the initial condition for $\lambda_R = \lambda_B = 1$.

Remark 3. *Note that the mixing strategy does not change the sequence numbers of packets from the same flow, and therefore it is compatible with network protocols such as TCP.*

III. Double input-double output mixer

Figure 2 shows the double input-double output mixer. The capacity of each link is 1 packet/time slot. Compared to the mixer with one output link, i.e., Figure 1, the flows of outgoing packets are separate. Note that, in this case, the eavesdropper does not need to detect the sender for each outgoing packet; instead, it aims to find the corresponding source of each flow, by observing a sequence of outgoing packets with enough length. Without loss of generality, assume that $\lambda_R > \lambda_B$ (the singular case of $\lambda_R = \lambda_B$ will be discussed later). Then, by calculating the long-run average rates of outgoing flows, the eavesdropper can identify the corresponding source-destination pairs. Therefore, it is not possible to get any anonymity without dropping some packets from the red flow. Hence, the maximum achievable throughput for each flow cannot be more than $\min\{\lambda_R, \lambda_B\}(= \lambda_B)$, and, at least, the packets of the flow with higher rate, which is the red flow here, must be dropped at an average rate of $\lambda_R - \lambda_B$. Note that, in contrast with the double input-single output mixer, a strict delay T for each packet does not make sense, since it might happen that there is a blue arrival but there are no red arrivals for a time duration of T, in which case transmitting the blue packet will immediately reveals the corresponding destination of the blue source. Therefore, instead, we consider the average delay as the QoS metric. We can model the mixer with two queues for red and blue arrivals. We only consider strategies that achieve maximum throughput with perfect anonymity. Perfect anonymity means that, by observing the output sequence, the eavesdropper cannot obtain any information and each outgoing flow is equally likely to belong to one of sources. Therefore, red and blue packets must be transmitted simultaneously on output links, i.e., red packets are only transmitted when there is a blue packet in the second queue, and similarly, the blue packets are served when there is a red packet in the first queue. Therefore, the question of interest is: how to drop red packets at an average rate of $\lambda_R - \lambda_B$ to minimize the average delay?

Since the average delay is proportional to the average queue length by Little's law, we can equivalently consider the problem of minimizing the mean queue length. This problem can be posed as an infinite-state Markov decision problem with unbounded cost. It follows from checking standard conditions, e.g., [21], [22], that a stationary optimal policy exists for our problem, however, the average-cost optimality equation may not hold. Therefore, we follow a different approach.

We note that when a red packet and a blue packet are both available, to minimize queue length, it is best to transmit them immediately. Therefore, when one of the queues (blue or red) hits zero, from that point onwards, only one of the queues can be non-empty. Thus in steady-state, we can assume that one queue can be non-empty (see [13] for more details). As a result, we have the Markov decision process described next. Figure 5 shows the state transition diagram, where (i, j) represents the state of the system where there are i packets in the red queue and j packets in the blue queue. The transition probabilities are given by

$$\begin{aligned} P[(0,y)|(0,y)] &= \lambda_R\lambda_B + (1-\lambda_R)(1-\lambda_B) \\ P[(0,y-1)|(0,y)] &= \lambda_R(1-\lambda_B) \\ P[(0,y+1)|(0,y)] &= \lambda_B(1-\lambda_R), \end{aligned}$$

and

$$\begin{aligned} P[(x,0)|(x,0)] &= \lambda_R\lambda_B + (1-\lambda_R)(1-\lambda_B) \\ &\quad +\lambda_R(1-\lambda_B)\delta_x \\ P[(x-1,0)|(x,0)] &= \lambda_B(1-\lambda_R) \\ P[(x+1,0)|(x,0)] &= \lambda_R(1-\lambda_B)(1-\delta_x), \end{aligned}$$

where δ_x denotes probability of dropping the red packet in state $(x, 0)$, if there is a red arrival but no blue arrival. So our

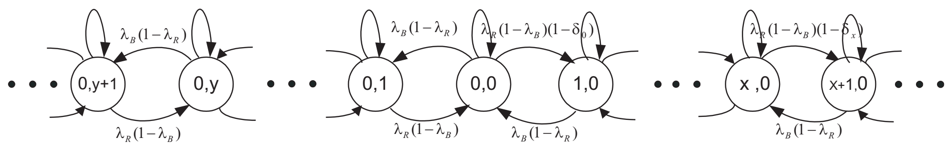


Fig. 5. Markov chain representing the double input-double output mixer

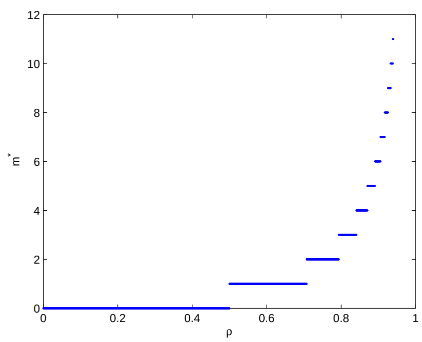


Fig. 6. The optimal threshold to minimize the average delay

problem is to determine δ_x for each x. We will show that the optimal policy is a threshold policy, which is defined below.

Definition 2. *A threshold policy, with threshold m, is a policy that has the following properties: $\delta_x = 0$ for all $0 \le x \le m-1$, and $\delta_m = 1$, where m is a nonnegative integer number.*

The following theorem presents the main result regarding the optimal strategy.

Theorem 2. *For the double input-double output mixer, the threshold policy is optimal, in the sense that it minimizes the average delay among all maximum throughput policies with perfect anonymity. Moreover, the threshold is given by*

$$m^* = \begin{cases} \lceil -\frac{1}{\log \rho} \rceil - 1 & ; \frac{1}{2} < \rho < 1 \\ 0 & ; 0 \le \rho \le \frac{1}{2}, \end{cases} \tag{15}$$

In other words, no buffer is needed for $\lambda_R \ge \frac{2\lambda_B}{1+\lambda_B}$, but, as rates get closer, for $\lambda_B < \lambda_R < \frac{2\lambda_B}{1+\lambda_B}$, a buffer of size m^* for the red flow is needed. The optimal threshold m^* is depicted in Figure 6. Note that the singular case of $\lambda_R = \lambda_B = \lambda$ ($\rho = 1$) is not stable. By allowing a small drop rate of $\epsilon\lambda$ for each flow, where $0 < \epsilon \ll 1$, one buffer for each flow can be considered, and the thresholds and the delay can be expressed as functions of ϵ.

Proof of Theorem 2: The steady state distribution for the Markov chain of Figure 5 is given by

$$\begin{aligned} \pi_{0,y} &= \pi_{0,0}\rho^y, \ y = 1, 2, \cdots \\ \pi_{x,0} &= \pi_{0,0}\rho^{-x}\prod_{i=0}^{x-1}(1-\delta_i), \ x = 1, 2, \cdots \end{aligned}$$

where

$$\pi_{0,0} = \left(\frac{1}{1-\rho} + \sum_{x=1}^{\infty}\rho^{-x}\prod_{i=0}^{x-1}(1-\delta_i)\right)^{-1},$$

and

$$\rho = \frac{\lambda_B(1-\lambda_R)}{\lambda_R(1-\lambda_B)}.$$

Recall that, by assumption, $\lambda_R > \lambda_B$, and therefore $0 \le \rho < 1$. The average queue length is

$$\begin{aligned} \bar{L} &= \sum_{y=0}^{\infty} y\pi_{0,y} + \sum_{x=1}^{\infty} x\pi_{x,0} \\ &= \pi_{0,0}\left[\frac{\rho}{(1-\rho)^2} + \sum_{x=1}^{\infty} x\rho^{-x}\prod_{i=0}^{x-1}(1-\delta_i)\right]. \end{aligned}$$

Note that for any nonnegative integer j, and for fixed values of δ_is, $i \ne j$, $\bar{L}$ is a linear fractional function of δ_j. More formally,

$$\bar{L}(\delta_j) = \frac{A_j + (1-\delta_j)B_j}{A'_j + (1-\delta_j)B'_j},$$

where

$$\begin{aligned} A'_j &= \frac{1}{1-\rho} + \sum_{x=1}^{j}\rho^{-x}\prod_{i=0}^{x-1}(1-\delta_i), \\ A_j &= \frac{\rho}{(1-\rho)^2} + \sum_{x=1}^{j} x\rho^{-x}\prod_{i=0}^{x-1}(1-\delta_i), \\ B'_j &= \frac{\prod_{i=0}^{j-1}(1-\delta_i)}{\rho^{j+1}}\left[1 + \sum_{x=1}^{\infty}\rho^{-x}\prod_{i=j+1}^{x+j}(1-\delta_i)\right], \end{aligned}$$

and

$$B_j = \frac{\prod_{i=0}^{j-1}(1-\delta_i)}{\rho^{j+1}}\left[j + 1 + \sum_{x=1}^{\infty}(j+x+1)\rho^{-x}\prod_{i=j+1}^{x+j}(1-\delta_i)\right].$$

Therefore, $\partial\bar{L}/\partial\delta_j$ is either positive or negative, independent of δ_j, and consequently, the optimal δ_j to minimize $\bar{L}$ is either 0 or 1, i.e., $\delta_j^* \in \{0, 1\}$ for all j. But, all of the δ_js cannot be zero, otherwise the system will not be stable. Define m to be the smallest j such that $\delta_j^* = 1$. Then $\delta_x = 0$ for all $0 \le x \le m-1$, and $\delta_m = 1$ which yields a threshold policy with threshold m. Therefore the threshold policy is the optimal policy.

Next, we find the optimal threshold m^*. The stationary distribution of a threshold policy with threshold m is given by

$$\begin{aligned} \pi_{0,y} &= \pi_{0,0}\rho^y, \ y = 1, 2, \cdots \\ \pi_{x,0} &= \pi_{0,0}(1/\rho)^x, \ x = 1, 2, \cdots, m \end{aligned}$$

where $\pi_{0,0} = (1-\rho)\rho^m$. Therefore, $\pi_{m,0} = 1-\rho$, and the average packet-drop rate, P_{drop}, is given by

$$P_{drop} = \pi_{m,0}\lambda_R(1-\lambda_B) = \lambda_R - \lambda_B$$

which is independent of the threshold m. The average queue length is given by

$$\begin{aligned} \bar{L}(m) &= \sum_{y=1}^{\infty} y\pi_{0,y} + \sum_{x=0}^{m} x\pi_{x,0} \\ &= \left(2\rho^{m+1} + m(1-\rho) - \rho\right)/(1-\rho). \quad (16) \end{aligned}$$

Note that $\bar{L}(m)$, as a continuous function of m, is strictly convex over $m \in [0,\infty)$ for any fixed $0 \le \rho < 1$; therefore, it has a unique minimizer m^* which is either zero or the solution of $\frac{\partial \bar{L}}{\partial m} = 0$. Since we seek the smallest integer-valued m^*, the convexity implies that m^* is zero if

$$\bar{L}(0) \le \bar{L}(1),$$

or it's a positive integer m^* satisfying

$$\bar{L}(m^*) < \bar{L}(m^*-1),$$

and

$$\bar{L}(m^*) \le \bar{L}(m^*+1).$$

Then by using (16), it follows that $m^* = 0$ if $\rho \le \frac{1}{2}$, and for $\rho > \frac{1}{2}$, it satisfies

$$2\rho^{m^*} > 1,$$

and

$$2\rho^{m^*+1} \le 1,$$

which yields

$$m^* = \lceil -\frac{1}{\log \rho} \rceil - 1.$$

This concludes the proof. ∎

Remark 4. *Instead of dropping the packets, the mixer can send dummy packets, at an average rate of $\lambda_R - \lambda_B$, as follows. In the optimal threshold strategy, if a red packet arrives when the red queue is full and there are no blue packets in the blue queue, a red packet is sent out along with a dummy packet on the other link, and the received red packet is accepted to the queue.*

IV. Conclusions

The definition of anonymity and the optimal mixing strategy for a router in an anonymous network depend on its functionality. In the case of a double input-single output mixer, an eavesdropper knows the next hop of every packet but the router attempts to hide the identity of the packet at the output link so as to make it harder for the eavesdropper to follow the path of a flow further downstream. On the other hand, when there are two inputs, two outputs and only two flows, even revealing the identity of one packet at the output compromises that portion of both flow's route. For the first case, the optimal mixing strategy was found to achieve the maximum anonymity under a per-packet latency constraint, and for the second case, the maximum throughput strategy with perfect anonymity that achieves minimum average delay was found. Our results in this paper represent a first attempt at theoretically characterizing optimal mixing strategies in two fundamental cases. Further research is needed to find optimal mixing strategies under more general constraints or for the multiple input-multiple output mixer.

References

[1] http://www.onion-router.net

[2] http://www.torproject.org

[3] D. Chaum, Untraceable electronic mail, return addresses and digital pseudonyms, *Communications of the ACM*, vol. 24, no. 2, pp. 84-88, February 1981.

[4] D. Kesdogan, J. Egner, and R. Buschkes, Stop-and-go MIXes providing probabilistic security in an open system, *Second International Workshop on Information Hiding*, Lecture Notes in Computer Science, Springer-Verlag, pp. 83-98, April 1998.

[5] A. Pfitzmann, B. Pfitzmann, and M. Waidner, ISDN-MIXes: Untraceable communication with very small bandwidth overhead, *Proceedings of the GI/ITG Conference: Communication in Distributed Systems, Informatik-Fachberichte*, vol. 267, (Mannheim, Germany), pp. 451-463, February 1991.

[6] C. Gulcu and G. Tsudik, Mixing e-mail with babel, *Proceedings of the Symposium on Network and Distributed System Security*, pp. 2-19, February 1996.

[7] M. K. Reiter and A. D. Rubin, Crowds: anonymity for Web transactions, *ACM Transactions on Information and System Security*, vol. 1, no. 1, pp. 66-92, 1998.

[8] O. Berthold, H. Federrath, and S. Kopsell, Web MIXes: A system for anonymous and unobservable Internet access, *Proceedings of Designing Privacy Enhancing Technologies: Workshop on Design Issues in Anonymity and Unobservability*, pp. 115-129. Springer-Verlag, LNCS 2009, July 2000.

[9] M. Rennhard and B. Plattner, Introducing morphmix: Peer-to-peer based anonymous internet usage with collusion detection, *Proceedings of the Workshop on Privacy in the Electronic Society (WPES 2002)*, November 2002.

[10] U. Moller, L. Cottrell, P. Palfrader, and L. Sassaman, Mixmaster Protocol, Version 2. Draft, July 2003.

[11] G. Danezis, R. Dingledine, and N. Mathewson, Mixminion: Design of a type III anonymous remailer protocol, *Proceedings of the 2003 IEEE Symposium on Security and Privacy*, May 2003.

[12] A. Beimel and S. Dolev, Buses for anonymous message delivery, *Journal of Cryptology*, vol. 16, no. 1, pp. 25- 39, 2003.

[13] C. E. Shannon, Communication theory of secrecy systems, *Bell System Technical Journal*, 1949.

[14] C. Diaz, S. Seys, J. Claessens, and B. Preneel, Towards measuring anonymity, *Proceedings of Privacy Enhancing Technologies Workshop (PET 2002)*, Springer-Verlag, April 2002.

[15] P. Venkitasubramaniam and V. Anantharam, On the Anonymity of Chaum Mixes, *2008 IEEE International Symposium on Information Theory*, Toronto, Canada, July 2008.

[16] P. Venkitasubramaniam, T. He and L. Tong, Anonymous Networking amidst Eavesdroppers, *IEEE Transactions on Information Theory: Special Issue on Information-Theoretic Security*, Vol. 54, No 6, pp. 2770-2784, June 2008.

[17] T. Cover and J. Thomas, Elements of Information Theory, *New York: Wiley*, 1991.

[18] J. Ghaderi and R. Srikant, Towards a theory of anonymous networking, *Technical Report*, avaiable online at http://www.ifp.illinois.edu/~jghaderi/anonymity.pdf

[19] D. P. Bertsekas, Dynamic programming and optimal control (1), *Athena Scientific*, 1995.

[20] V. Vasilevich, Elements of combinatorial and differential topology, *Graduate Studies in Mathematics*, V. 74, 2006.

[21] S. A. Lippman, Semi-Markov decision processes with unbounded rewards, *Management Science*, vol. 19, No. 7, pp. 717-731, March 1973.

[22] L. Sennott, Average cost optimal stationary policies in infitie state markov decision processes with unbounded costs, *Operations Research*, Vol. 37, No. 4, pp. 626-633, August 1989.