MALARIA: Management of Low-Latency Routing Impact on Mix Network Anonymity

Mahdi Rahimi
COSIC, KU Leuven
Leuven, Belgium
mahdi.rahimi@esat.kuleuven.be

***Abstract*—Mix networks (mixnets) offer robust anonymity even against adversaries monitoring all network links; however, they impose high latency on communications. To address this, recent research has explored strategic low-latency routing within mixnets. While these strategies appear to reduce latency, their impact on mixnet anonymity has not been carefully assessed, raising concerns about potential deanonymization of clients. Tackling this challenge, this paper first quantifies the anonymity loss associated with low-latency routing techniques in mixnets. Building on these insights, second, we introduce a novel low-latency routing method that maintains mixnet anonymity while achieving significant latency reductions compared to the state-of-the-art solution LARMix (NDSS, 2024). Our approach also ensures a more balanced load distribution among mixnet nodes. Moreover, under adversarial conditions where parts of the mixnet are compromised, our method does not confer significant advantages to the adversary, unlike LARMix. Thus, our proposal emerges as the optimal choice for low-latency routing in mixnets.**

***Index Terms*—Network Security, Anonymity, Mix Networks**

I. Introduction

Anonymous communication, which conceals the identities of communicating parties, can be achieved through various methods [1]–[3]. Among these, mix networks (mixnets) [2], [4]–[6] stand out as one of the most effective strategies for ensuring anonymity. Mixnets are designed to provide anonymity even in the presence of global passive adversaries (GPA) who monitor all traffic exchanges between network entities. They accomplish this by routing clients' messages through multiple intermediary nodes, known as mixnodes, which reorder incoming messages to prevent linking them to their outgoing counterparts [7]. This process ensures that as long as at least one mixnode along the message route performs its reordering duties without colluding with the GPA, correlating input messages to their outputs becomes impossible, thereby safeguarding the anonymity of the clients' messages [8].

The anonymization achieved by the mixnet is primarily tied to the mixing operations performed by the mixnodes, which can be executed using various methods. The first method is threshold mixing [2], where messages are forwarded once the number of accumulated messages inside the mixnode reaches a certain threshold. The second method, known as pool mixing [9], requires both a threshold number of messages to accumulate and a specific period to elapse before sending out the messages. Finally, stop-and-go mixing [10] flushes each message based on a random delay determined by an exponential distribution. Among these methods, stop-and-go mixing provides a good level of anonymity due to the memoryless property of the exponential distribution and offers a manageable average latency, making it a suitable choice for practical deployment [6].

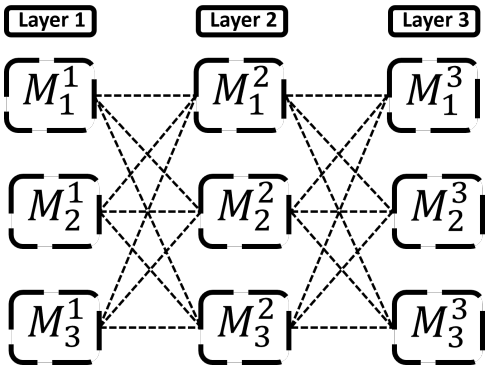


Fig. 1: A stratified mixnet consists of three layers, with each layer containing three mixnodes.

Besides various message mixing types, there are also different ways to construct a mix network [11]. To build a mixnet featuring L hops, one can consider a cascade topology by organizing mixnodes into cascades, each containing L mixnodes. Clients then select one of these cascades for message routing [12], [13]. Alternatively, clients can randomly select L distinct nodes from all available mixnodes to establish a message path, forming what is known as a Free Routes mixnet [14]. Another configuration involves categorizing mixnodes into L different layers, creating a stratified mixnet [4], [6], where clients select one node from each layer to construct a message path.

Among all the mixnet topologies and message mixing types, the stratified topology coupled with Poisson mixing is more favorable for online communications [11]. Specifically, the stratified topology provides a high degree of anonymity due to the numerous message-route options it offers. It achieves the highest anonymity for a mixnet when incorporating cover traffic, which is generated by clients and eventually loops back to the clients themselves to deceive the GPA and increase the degree of anonymity for real messages [11]. Furthermore, the stratified topology becomes particularly practical when

equipped with stop-and-go mixes, forming a mixnet known as Loopix [4]. This combination provides high anonymity and manageable mixing delay, making it suitable for deployment, and has recently been implemented by Nym.[1] To base our analysis on a real-world mixnet, we also consider this design in our paper. Fig. 1 depicts this configuration, with the number of layers L set to 3, and each layer containing 3 stop-and-go mixnodes.

As mentioned, the Loopix-like mixnet provides a high degree of anonymity while maintaining manageable average latency for mixing operations. However, various sources of latency in this mixnet can lead to high end-to-end communication latency, deterring clients from using mixnets for web browsing or instant messaging. Specifically, this high latency occurs because, **first**, messages are forwarded through multiple hops, which can be located in diverse geographical locations. This can add high link delays between mixnodes, contributing to high end-to-end latency. **Second**, these messages are intentionally delayed at each hop to ensure untraceability among their peers, which additionally causes high latency.[2]

To manage the heightened latency in mixnets, one method is to reduce the average mixing delay introduced by mixnodes. For example, in the case of stop-and-go mixes, this can be achieved by tuning the parameter of the exponential distribution. However, reduced mixing delay results in fewer messages being mixed together, thereby decreasing anonymity. Consequently, this approach is not particularly appealing. Alternatively, LARMix [15] proposes an innovative way to reduce end-to-end latency. Instead of reducing the mixing delay, LARMix suggests that clients strategically choose hops in the mixnet to ensure they are not too far from each other. This approach reduces high link delays between nodes, potentially lowering end-to-end latency. More precisely, LARMix proposes a node selection scheme that ensures a diverse assignment of mixnodes across network layers, incorporating a sufficient number of nodes from varied jurisdictions. Additionally, they suggest routing messages based on a formula that prioritizes proximity in node selection along message paths.[3]

Problem Statement. The LARMix strategy has shown promise in reducing end-to-end latency in mixnets; nevertheless, their approach has certain drawbacks that can lead to the deanonymization of clients using the mixnet. Specifically, LARMix employs a heuristic model to reduce latency and intuitively describe its impact on anonymity. However, such a heuristic model may significantly underperform in scenarios where latency distributions, for example, follow a Pareto distribution. This type of analysis may prove ineffective in practice, as it does not account for all potential cases that may arise in real-world networks, thereby failing to establish a robust balance between anonymity and latency. Consequently, there is a need to quantify the effects of such strategies to develop more effective routing methods.

Our Contributions. To better understand the anonymity loss when employing low-latency routing, our first contribution is to derive the relationship between strategic routing and anonymity in mixnets, which can be applied to any strategic routing, including low-latency routing. Specifically, we consider a stratified mixnet with L layers and W mixnodes at each layer. We begin by referencing the definition of message anonymity described in [9], which is tied to the probability of a GPA matching a targeted message entering the mixnet to any message exiting the mixnet. Furthermore, the Shannon entropy of this probability distribution [18] is used as the measure of message anonymity.[4] Taking this into account, we show that when mixnodes receive roughly the same amount of traffic, message anonymity can be formalized as a linear combination of routing entropy and mixing process entropy. Routing entropy is defined as the entropy of the probability distribution mapping the entry node of a target message to any of the exit nodes in the last layer, reflecting the adversary's attempt to identify the last node in the target message path. The entropy of the mixing process is defined as the message anonymity of a target message when the path is deterministic, such as in a cascade topology.

Considering such formalization, we realize that low-latency routing will primarily affect the entropy of routing. Based on this observation, analyzing low-latency routing strategies in terms of their impact on anonymity before deployment is crucial for managing the trade-off between latency and anonymity. To this end, we **first** recognize LARMix low-latency routing [15] as heuristic-based routing. LARMix proposes a heuristic formula that considers the latency among nodes in the mixnet and outputs the probability of node selection. However, such heuristics can fall short in scenarios where there is a relatively Pareto distribution of latency in the network. In such cases, traffic will be directed to mixnodes that are only slightly faster, leaving other mixnodes with less traffic and thereby lowering anonymity. Additionally, the heuristic formula may not fully optimize mixnet latency due to its inherent limitations.

To provide a more effective approach to reducing latency, we **second** propose the Managing the Randomness for Selecting Edges (MORSE) approach. In this method, low-latency routing is treated as a linear programming problem, solved with constraints controlling the randomness of the routing via a tunable input parameter to the algorithm. This allows for a more optimized latency achieved by low-latency routing.[5] However, this method may sacrifice anonymity more significantly due to its focus on finding the best-optimized routing.

[1] https://nymtech.net

[2] Note that we assume cryptographic operations, needed for mixing and overlay routing, incur negligible latency.

[3] CLAM [16] is another similar work to LARMix, designed to extend LARMix's results to include low-latency message forwarding from clients to the mixnet. Additionally, LARMix++ [17] provides a low-latency routing solution for the Free Routes topology in mixnets.

[4] For example, if the anonymity of messages is n bits, it means that the adversary is confused among a set of 2^n messages to find the counterpart of the targeted message.

[5] This approach is an optimized routing method that always provides the lowest latency routing, therefore not heuristic.

Lastly, we propose the Leave One Random (LOR) method, which forwards messages to the fastest mixnodes from the first mixnode to the one before the last layer, and then selects the very last mixnode in the message path uniformly at random. This innovative approach ensures no sacrifice in anonymity (routing entropy) while achieving latency reduction as effective as the MORSE approach.[6]

To analyze our proposals, we created a mixnet simulator using *Simpy* [19] in *Python*, featuring all the methods proposed in this paper. In the simulation, we assumed a GPA observing all the links and recorded the latency of messages along with the analysis of anonymity, including routing and mixing process entropies. Simulation results confirm that decomposing message anonymity into the entropy of the mixing process and routing entropy is valid. Furthermore, our analysis specifically demonstrates that MORSE generally provides lower latency compared to the LARMix proposal, potentially reducing latency by up to 50% more than LARMix. However, LARMix exhibits higher anonymity. Additionally, the simulation results reveal that the LOR approach does not sacrifice routing entropy at all, proving to provide the highest anonymity among the strategies. Moreover, it reduces latency as effectively as MORSE, if not slightly better.

Additionally, we measured the overall percentage of nodes in the mixnet that have the potential to become overloaded using the proposed routing strategies. Our analysis shows that LOR keeps the mixnodes more balanced compared to the MORSE and LARMix approaches. We extended our analysis to the computational burdens of these approaches, revealing that using interior-point methods for linear programming, MORSE is the least efficient with a complexity of $O(N^6)$ when N is the number of mixnodes ($W \times L$). LARMix has a complexity of $O(N \log(N))$, and, remarkably, LOR also has a complexity of $O(N \log N)$ and even less complexity compared to LARMix when comparing more accurately. This demonstrates that LOR is the most efficient approach when considering all aspects.

Finally, we performed an analysis to ensure that the proposed approach does not substantially enhance the capabilities of mixnode adversaries who corrupt (own) some mixnodes in the mixnet and, in collaboration with the GPA, attempt to deanonymize client connections. This analysis considers both random corruption of nodes and corruption using our proposed greedy corruption strategy, which we developed to allow an adversary to strategically compromise the most critical nodes, quantified in terms of the fraction of fully corrupted paths (FCP) they provide, which incurs full deanonymization. Our experiments indicate that our approach, LOR, is not significantly affected by such adversaries and also performs much better than both LARMix and MORSE in being resilient to such attacks.

Outline. In the remainder of the paper, Section II-A investigates the effect of strategic (low-latency) routing on mixnet anonymity. Section II-B presents the MORSE and LOR strategies and compares them to LARMix. Section III evaluates our proposed approaches, and Section IV provides an analysis of mixnode adversary threats. Finally, Section V concludes the paper.

[6]LOR always provides full anonymity in any situation, making it suitable under any conditions, and therefore not heuristic in this sense.

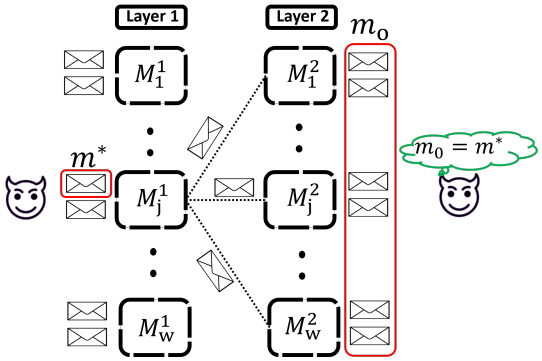


Fig. 2: Mixnet with two layers.

II. Methodology

In this section, we first explore the components shaping message anonymity in mix networks. Building on these findings, we then investigate different low-latency routing strategies in mix networks.

A. Anonymity Components

Mixnets are designed to thwart the attempts of an adversary observing all communication parts of the network (GPA) by performing mixing operations in multiple mixnodes. As a result, it might seem that anonymity in mixnets is solely concerned with the number of messages being mixed in the mixnodes. However, this section delves deeper into this concept, revealing a more nuanced understanding. We consider an adversary who targets an input message to the mixnet and tries to find its counterpart at the exit of the mixnet. By examining the probability distribution of the targeted message over all the exiting messages and computing the entropy of this distribution, we quantify the anonymity of the target message. Through various scenarios, we demonstrate how the anonymity of this message is affected when routing within the mixnet is more deterministic. This analysis reveals that the mixing process alone is not the only contributing factor to message anonymity.

1) Two-Layer Mixnet: We begin our analysis with a two-layer mixnet featuring W mixnodes at each layer, as illustrated in Fig. 2. To assess anonymity as suggested in [9], we consider a GPA targeting messages at the input of the mixnet, aiming to identify their counterparts at the exit of the mixnet, with the goal of deanonymizing the client-destination connections.

More specifically, Fig. 2 describes this scenario where the GPA targets a message m^* entering mixnode j in the first layer. Upon arrival in this mixnode (M_j^1), the targeted message is mixed with its peer messages inside the mixnode and then sent out after an exponential delay to the succeeding mixnodes in the second layer according to a predetermined routing

policy, which is essentially a probability distribution from M_j^1 over the mixnodes in the second layer. More accurately, we denote the probability of routing the message from M_j^1 to M_j^2 as r_j.

In this case, the adversary watching the output stream wants to determine the probability $\mathbb{P}(m_o = m^*)$ for every outgoing message m_o. We aim to analyze how much this probability is affected by the choice of mixnodes in the second layer to further understand how message anonymity is influenced by the routing policy. To do so, we rewrite the probability distribution of the targeted message over the outgoing messages at the mixnet exit as shown in Eq. (1), based on the possibility of having m_o emerge from any mixnode in the second layer.

$$\begin{aligned}\mathbb{P}(m_o = m^*) &= \sum_{j=1}^{W} \mathbb{P}(m_o = m^* | m^* \in M_j^2)\mathbb{P}(m^* \in M_j^2), \\ &= \sum_{j=1}^{W} \mathbb{P}(m_o = m^* | m^* \in M_j^2) r_j. \end{aligned} \tag{1}$$

Eq. (1) describes that the probability of each outgoing message being the targeted message is tied to $\mathbb{P}(m_o = m^* | m^* \in M_j^2)$ and the routing probability r_j. However, the probability $\mathbb{P}(m_o = m^* | m^* \in M_j^2)$ is influenced by the mixing process performed by the mixnodes. Therefore, the distribution of the targeted message over outgoing messages is a joint probability of both the mixing process by mixnodes in the first and second layers and the routing distribution from the first to the second mixnet layers. Since all mixnodes in the mixnet should ensure a consistent level of message mixing, we can approximate that this mixing process is independent of the routing policy, i.e., $\mathbb{P}(m_o = m^* | m^* \in M_j^2) = \mathbb{P}_{Mix}(m_o = m^*)$.[7]

With this analysis, we can formalize the anonymity of the targeted message (the entropy of the target message) as $\mathcal{H}(m) = \mathcal{H}(f_M; f_R) = \mathcal{H}(f_M) + \mathcal{H}(f_R|f_M)$, where f_M refers to the probability distribution of the targeted message m^* over the outgoing message m_o caused by the mixing process, and f_R refers to the probability distribution of the routing. Since the routing policy is nearly independent of the mixing process at each mixnode, we can simplify this to $\mathcal{H}(m) = \mathcal{H}(f_M) + \mathcal{H}(f_R)$.

This reveals that message anonymity has two components: the mixing process conducted by mixnodes in the first and second layers, and the routing policy. The effect of the mixing process is equivalent to the scenario where messages from mixnode j in the first layer are forwarded to mixnode i in the second layer with a probability of 1. In such a case, the GPA only needs to consider the mixing process of both mixnodes for analyzing the probability distributions, with no routing entropy involved. However, when routing is not fully deterministic, the choice of mixnode in the second layer also contributes to message anonymity as $\mathcal{H}(f_R)$, thereby increasing overall anonymity.

[7]This consistent load distribution over mixnodes is guaranteed by clients sending some messages without a specific destination (cover traffic), which enhances anonymity by compensating for the lower volume in less-loaded mixnodes.

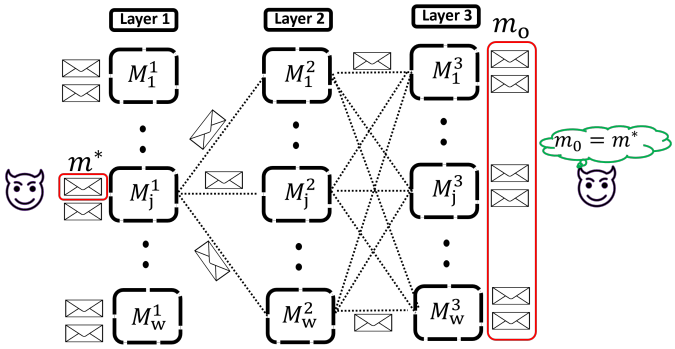


Fig. 3: Mixnet with three layers.

Furthermore, we note that the entropy of the mixing process depends on the number of messages within the mixnodes; the higher the number of messages, the higher the anonymity provided by mixing. However, the anonymity provided by routing depends on the probability distribution over W available mixnodes in the second layer. In the best case, using a uniform distribution, the entropy of this probability can be maximized to $\mathcal{H}(f_R) = \log(W)$, while when the routing is fully deterministic, this entropy is equal to 0.

2) Three-Layer Mixnet: We now extend the scenario mentioned for a mixnet with two layers to a mixnet with three layers, as shown in Fig. 3. In this case, as before, the adversary targets a message m^* in the mixnet input and aims to distinguish the outgoing message m_o corresponding to this target message. In this context, to measure the probability distribution of the target message to the exiting messages $\mathbb{P}(m_o = m^*)$, we first define the routing probability from mixnode i in layer k to mixnode j in layer $k+1$ as r_{ij}^k. Subsequently, considering all the distributions from layer k over layer $k+1$, we define the routing matrix of layer k as $\mathbf{R}^k = [r_{ij}^k]$. Taking this into account, we can derive $\mathbb{P}(m_o = m^*)$ as shown in Eq. (3) and Eq. (4).

$$\mathbb{P}(m_o = m^*) = \sum_{i=1}^{W} \mathbb{P}(m_o = m^* | m_o \in M_i^3)\mathbb{P}(m_o \in M_i^3), \tag{2}$$

$$= \sum_{i=1}^{W} \mathbb{P}(m_o = m^* | m_o \in M_i^3) \times \left(\sum_{k=1}^{W} \left[\mathbb{P}(m_o \in M_i^3 | M_k^2 \to M_i^3)\mathbb{P}(M_k^2 \to M_i^3) \right] \right), \tag{3}$$

$$= \sum_{i=1}^{W} \mathbb{P}(m_o = m^* | m_o \in M_i^3) \left(\sum_{k=1}^{W} \left[r_{ki}^2 r_{jk}^1 \right] \right). \tag{4}$$

To derive Eq. (4), we first use the law of total probability in Eq. (2) to rewrite $\mathbb{P}(m_o = m^*)$ based on the exit nodes in the third layer. Then, considering the probability of messages coming out of the third layer, for example, from the i-th

mixnode in the third layer $\mathbb{P}(m_o \in M_i^3)$, we realize that this probability can be based on the messages received from different mixnodes in the second layer as well, which can be formalized in Eq. (3). Furthermore, we note that any mixnode in the second layer with a specific probability receives messages from the j-th mixnode in the first layer, leading to Eq. (4).

Eq. (4) describes the distribution of m^* over m_o, similar to the two-layer case. This distribution is also tied to two parts: **first**, $\mathbb{P}(m_o = m^* | m_o \in M_i^3)$, which describes the effect of the mixing process on anonymity and includes all three mixnodes from the first layer to the third layer; and **second**, the routing policy part $\sum_{i=1}^{W}\sum_{k=1}^{W}\left[r_{ki}^2 r_{jk}^1\right]$, showing the traveling route from the first to the third mixnodes and how differently a message can be forwarded from M_j^1 to any mixnodes at the mixnet exit. As before, we can approximately conclude that $\mathbb{P}(m_o = m^* | m_o \in M_i^3) = \mathbb{P}_{Mix}(m_o = m^*)$, indicating these two components are almost independent.

Looking more closely at $\sum_{i=1}^{W}\sum_{k=1}^{W}\left[r_{ki}^2 r_{jk}^1\right]$, we can rewrite it based on the definition of the routing matrix as $\mathbf{R}_j^1\mathbf{R}^2$, the multiplication of the j-th row of the first layer's routing matrix by the second layer's routing matrix. This demonstrates the different ways to travel from node j in the first layer to any node in the third layer, forming the probability distribution from node j in the first layer to the third-layer mixnodes, constituting the entropy of routing $\mathcal{H}(f_R)$. Together with the entropy of mixing $\mathcal{H}(f_M)$, this provides the overall anonymity of the message $\mathcal{H}(m)$, showing a direct influence of the routing policy on anonymity.[8]

3) Generalizations: Finally, in this section, we generalize the derived results to a stratified mixnet with L layers, each featuring W mixnodes, where layer K has its routing matrix $\mathbf{R}^k$. In such a case, a GPA attempting to deanonymize a connection starting from mixnode j in the first layer faces a challenging task modeled by the entropy of the targeted message ($\mathcal{H}(m)$). This entropy can be further broken down into two parts: the entropy of the mixing process ($\mathcal{H}(f_M)$) and the entropy of routing ($\mathcal{H}(f_R)$). The entropy of mixing, as before, depends on the mixing process done by mixnodes and is independent of the routing distributions. The entropy of routing can similarly be derived as the entropy of f_R, which can be formulated, as seen in the two- and three-layer cases, in Eq. (5):

$$f_R = \mathbf{R}_j^1 \prod_{k=2}^{L-1} \mathbf{R}^k. \tag{5}$$

In other words, Eq. (5) describes that to find the entropy of routing given the targeted message starting from mixnode j in the first layer, one needs to consider the routing matrix for the first layer and then multiply the j-th row of this matrix by the products of subsequent layers' matrices. To compact this into a concise formula, we define $\mathbf{T} = \prod_{k=1}^{L-1}\mathbf{R}^k$. We call matrix $\mathbf{T}$ the transition matrix that maps the input mixnodes of the mixnet to the last layer mixnodes. To derive the targeted message's routing anonymity, one needs to calculate the entropy of the j-th row of this matrix. For an overall evaluation, we can consider the average entropy of all rows or the minimum entropy in this matrix.[9] Another interpretation of the entropy of the transition matrix is the confusion an adversary faces in identifying the last mixnode in the message route given its first mixnode in the route. This, together with the mixing process entropy, culminates in the overall anonymity of the message.

Lastly, we note that the mixing process generally depends on the number of messages inside the mixnodes, which should be controlled with the exponential distribution parameter and the introduction of more cover traffic if needed, while the entropy of routing depends on routing policies. One needs to ensure that biased routing does not significantly sacrifice routing anonymity, as it can vary from 0 to $\log(W)$.

B. Low-Latency Routing Strategies

Our analysis so far highlights two primary sources of anonymity for mixnodes with a stratified topology: the mixing process performed by each mixnode and the routing policy. The effect of these sources can be summarized as the anonymity of the target messages ($\mathcal{H}(m) = \mathcal{H}(f_M) + \mathcal{H}(f_R)$). This formulation holds under the condition that mixnodes provide a relatively uniform level of security (mixing) for messages and becomes particularly useful when considering biased routing in mixnets, allowing us to measure the effect of routing distributions directly on the anonymity provided by the mixnet. As an example of such a scenario, we consider low-latency routing.

Low-latency routing becomes relevant as mixnets achieve their anonymity by forwarding client traffic through multiple intermediary mixnodes, altering the flow of traffic. However, this process is accompanied by heightened latency resulting from link delays between mixnodes and the intentional delays introduced by mixnodes. To extend the application of mixnets to low-latency environments, routing might need to be biased towards low-latency paths to reduce end-to-end latency. In this section, we consider such a scenario, first investigating the existing method LARMix [15] for low-latency routing. Based on the results of the previous sections, we propose methods that outperform this state-of-the-art approach, optimizing latency and achieving both reduced latency and maintained anonymity.

1) Heuristics-Based Routing (LARMix): We begin our analysis of low-latency routing by introducing one of the proposed heuristic methods to prioritize low-latency links over high-latency ones in the mixnet. While heuristic formulas can be helpful, they might not always provide the best performance.

[8] In this case, the mixing process's effect on anonymity is similar to having a cascade of three mixnodes for forwarding the message, where the routing between mixnodes is deterministic. In such a scenario, the confusion an adversary faces when mapping the outgoing message to the targeted message is determined by $\mathcal{H}(f_M)$ without involving the entropy of routing, as the routing is deterministic. If the routing is not deterministic, the entropy of routing further contributes to increasing the adversary's confusion (anonymity of the message) by $\mathcal{H}(f_R)$, which can vary from 0 to $\log(W)$.

[9] A similar definition is given to this matrix in LARMix. However, the LARMix paper intuitively highlights this as important, in contrast to our methodology, which naturally derives this matrix.

However, they can be valuable for understanding the impact of routing distribution effects. To this end, we explore the state-of-the-art LARMix low-latency routing formula [15].

The LARMix routing heuristic is described in Eq. (6), showing r_{ij}^k for choosing mixnode j at layer $k+1$ from layer k tied to l_{ij}^k, the corresponding latency between these two nodes. LARMix prioritizes mixnodes with lower latency by considering the inverse of latency raised to the power of $1-\tau$. Additionally, LARMix introduces a ranking function f_{ij}, which assigns a rank indicating the closeness of j to i, starting from 0 for the closest mixnode to $W-1$ for the farthest mixnode. This ranking mechanism ensures that as τ approaches 0, the selection of the closest mixnode occurs with a probability approaching 1, tuning the routing from a fully deterministic choice of the closest mixnode to a uniform routing when $\tau = 1$. The parameter τ helps provide low latency while controlling the level of anonymity.

$$r_{ij}^k = \frac{\left(\frac{1}{e}\right)^{f_{ij}\cdot\frac{(1-\tau)}{\tau}} \cdot \left(\frac{1}{l_{ij}^k}\right)^{(1-\tau)}}{\sum_{j=1}^{W}\left(\frac{1}{e}\right)^{f_{ij}\cdot\frac{(1-\tau)}{\tau}} \cdot \left(\frac{1}{l_{ij}^k}\right)^{(1-\tau)}}. \tag{6}$$

$$\text{minimize} \left(\sum_{k=1}^{L-1}\sum_{i,j=1}^{W} r_{ij}^k \cdot l_{ij}^k\right), \tag{7}$$

S.t:

$$\vee\, i, \quad \sum_{j=1}^{W} r_{ij}^k = 1, \tag{8}$$

$$\left(\frac{1}{W}\right)^{(1+\theta)} \leq r_{ij}^k \leq \left(\frac{1}{W}\right)^{\exp(-\theta)}. \tag{9}$$

2) Managing the Randomness for Selecting Edges (MORSE): Heuristic methods like LARMix have a primary drawback in that they do not always provide the most optimized routing in terms of latency. To overcome this limitation, we propose the MORSE method, which minimizes the average link latency in the mixnet using linear programming methods. This approach guarantees the best settings under any conditions, making it non-heuristic in terms of latency. Furthermore, MORSE manages the randomness of routing by constraining the routing variations. Specifically, consider r_{ij}^k, the routing probability between mixnode i in layer k and mixnode j in layer $k+1$. In uniform routing, this probability would be equal to $\frac{1}{W}$; however, in low-latency routing, it can vary from 0 to 1. MORSE ensures that r_{ij}^k is controlled to minimize the average latency.

To formalize the MORSE approach, consider the latency between mixnode M_j^K and mixnode M_i^{K+1}, defined as l_{ij}^K. For a mixnet with L layers, we aim to minimize the average link delays by solving a linear programming problem specified in Eq. (7), subject to the constraints in Eq. (8), ensuring that the routing distributions are valid probability distributions, and Eq. (9), where we control the level of bias for r_{ij}^k with parameter θ, which varies from 0 to infinity. If θ is 0, we have minimum variation (zero variation), so r_{ij}^k will be $\frac{1}{W}$. However, as θ increases, the variation also increases, and when θ is infinite, the variation ranges from 0 to 1, representing the maximum variation. This approach provides flexibility for the network designer to find the best trade-offs while minimizing latency under such constraints.

3) Leave One Random (LOR): Although the MORSE method perfectly minimizes latency compared to the LARMix heuristic, it still has the same shortcoming as LARMix concerning sacrificing anonymity to achieve low latency. To overcome this limitation, we propose the Leave One Random (LOR) method, which does not sacrifice anonymity at all and minimizes latency to the lowest possible level.

To understand LOR, consider the transition matrix $\mathbf{T}$, which describes the probability of messages entering input mixnodes being routed to any of the output mixnodes. This matrix is equal to the product of all routing matrices. Using the example shown in Fig. 3 with a mixnet featuring three layers, imagine the first layer's routing matrix is biased with a low-latency routing approach, while the second layer's routing matrix is uniform (i.e., each entry of $\mathbf{R}^2$ is $\frac{1}{W}$). To derive the entry $\mathbf{T}_{ij}$, we compute $\mathbf{T}_{ij} = \sum_{k=1}^{W} r_{ik}^1 \times r_{kj}^2$. Given that $r_{kj}^2 = \frac{1}{W}$, we can conclude that $\mathbf{T}_{ij} = \frac{1}{W}\sum_{k=1}^{W} r_{ik}^1$. Since $\sum_{k=1}^{W} r_{ik}^1$ is the summation of a probability distribution (each row of $\mathbf{R}^1$ is a probability distribution and should sum to 1), we have $\mathbf{T}_{ij} = \frac{1}{W}$.

This means that we can bias the routing for the first layer's routing matrix and still achieve maximum rouitng anonymity by having the second layer's routing matrix set to uniform. We can similarly extend this to a mixnet with L layers. If the last routing matrix is uniform, we can achieve maximum routing anonymity as the transition matrix will be uniform independently of how the other routing matrices are designed. This reveals an interesting feature of the transition matrix. We define this feature as Leave One Random, specifically referring to the last routing matrix that should be uniform.[10]

To provide low latency using the LOR strategy, consider the link latency between mixnode M_j^K and mixnode M_i^{K+1}, defined as l_{ij}^K. For a mixnet with L layers, our goal is to minimize the average link delays while maintaining anonymity. For all layers except the last, routing will be performed by assessing all possible succeeding links and choosing the fastest link (almost fully deterministic routing up to the last layer). Finally, by adding a uniform distribution for $\mathbf{R}^{L-1}$, we obtain routing policies for all layers. This approach ensures that we achieve low-latency routing without compromising anonymity.

III. Evaluation

In this section, we evaluate the analysis developed for the anonymity components and the low-latency approaches proposed in this paper, comparing them to the state-of-the-art LARMix. Our analysis includes both analytical evaluation and

[10]Note that when the routing matrices are biased but balanced, meaning that the summation of each column is equal to one, we can have one of the routing matrices set to uniform random at any layer and still achieve a uniform random transition matrix.

simulation. To align our evaluation with real-world scenarios, we used the RIPE Atlas dataset [20] to emulate the link latency between nodes in the mixnet. We considered a stratified mixnet with L layers, each featuring W mixnodes. To ensure robust results, we repeated each experiment 400 times, using a new set of nodes to configure the mixnets each time. After configuring the mixnet, we computed the average link delays using the formula $\sum_{k=1}^{L-1}\sum_{i,j=1}^{W} r_{ij}^{k} l_{ij}^{k}$, which represents the average influence of routing on the latency experienced by messages traveling between mixnodes in the mixnet. Additionally, with all routing matrices at our disposal, we quantified $\mathcal{H}(f_R)$.

Furthermore, to measure the anonymity of messages ($\mathcal{H}(m)$) and the anonymity provided by the mixing process ($\mathcal{H}(f_M)$), we simulated a mix network using the *SimPy* [19] environment in *Python*, which is essentially a discrete event simulator. In this simulation, we sent actual messages to a layered network of stop-and-go mixes, each equipped with a mixing delay following an exponential distribution with an average delay of 50 ms (based on a real deployed network, NYM). On average, we sent 1000 messages per second, based on a Poisson distribution, to each mixnode at the entry layer of the mixnet. We recorded the entry and exit times of messages to measure end-to-end latency and randomly targeted some messages at the network input, using methods explained in [21] to measure the entropy of messages ($\mathcal{H}(m)$).

Finally, when using low-latency routing, some mixnodes can become underloaded with messages. To ensure our analysis reflects a real mixnet, we introduced additional messages generated by clients that have no destination and loop back to themselves. This approach ensures each mixnode receives a nearly uniform amount of traffic, counterbalancing the biased routing traffic resulting from low-latency routing.[11]

A. Varying Mixnet Size

We begin the first set of experiments by modifying the number of mixnodes per layer (W) while keeping the number of layers fixed at $L = 4$. We report the entropy of routing ($\mathcal{H}(f_R)$), the entropy of targeted messages ($\mathcal{H}(m)$), and the link delay latency, along with the level of imbalance in message distribution among mixnodes caused by low-latency routing. Additionally, we set the parameter τ in LARMix routing to 0.6, corresponding to the LARMix proposal [15], and the parameter θ in the MORSE approach to 5 to ensure optimized routing while providing some level of anonymity.

Fig. 4a illustrates the anonymity of routing ($\mathcal{H}(f_R)$) while varying the number of mixnodes per layer from 8 to 64 for all proposed low-latency routing approaches. As shown in this figure, increasing the number of mixnodes enhances the anonymity of the routing. This is because having more nodes in the mixnet increases the potential exit points for messages at the last layer, thereby enhancing the entropy of routing.[12] Furthermore, we observed that LOR provides the highest entropy of routing. A closer examination shows that the entropy of routing provided in this case is equal to the maximum possible entropy ($\log(W)$) that the mixnet can offer under uniform conditions, indicating that LOR does not compromise anonymity at all. In contrast, both LARMix and MORSE reduce the routing anonymity. Specifically, LARMix routing can reduce the anonymity by up to 2 bits when $W = 64$, while MORSE provides anonymity no better than 1 bit.

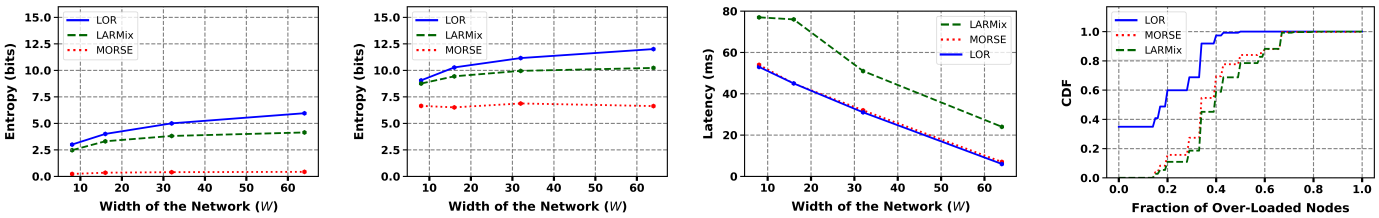
Fig. 4b, on the other hand, represents the anonymity of messages ($\mathcal{H}(m)$) while varying the number of mixnodes per layer. A first look at this figure reveals that the message anonymity curves for all low-latency routing approaches resemble an upward shift in the routing anonymity curves. This observation confirms that message anonymity is a linear combination of routing entropy and mixing process entropy. Therefore, we see an increase in message anonymity with an increase in network size, with LOR providing the highest anonymity. However, these experiments show that although MORSE had low anonymity provided by routing, the anonymity gained from the mixing process compensates to some extent, resulting in an average level of 7 bits of message anonymity.

Furthermore, Fig. 4c describes the average latency caused by link delays between nodes using our proposed low-latency routing while varying the number of nodes per layer. Fig. 4c exhibits an interesting trend of latency reduction by adding more mixnodes per layer for all the approaches. Although it might not be intuitive, the reason for lowering latency by adding more mixnodes stems from the inherent nature of low-latency routing, which always prefers the fastest links. So, when adding more nodes to the layers, the chances of having faster nodes increase, consequently reducing the end-to-end communication latency. More interestingly, Fig. 4c reveals that, generally, MORSE and LOR provide greater reductions in latency. Specifically, MORSE and LOR showed at least a 20 ms larger reduction compared to the latency reduction provided by LARMix. This observation is noteworthy as it shows that although MORSE sacrifices a lot in terms of routing entropy (thus reducing message anonymity), it reduces latency more efficiently than LARMix due to the systematic nature of its linear programming approach. More impressively, LOR reduces latency as effectively as MORSE, if not slightly better, achieving the lowest latency possible in the network without sacrificing routing entropy, leading to the highest anonymity with the lowest latency.

As mentioned earlier, after applying low-latency routing, we add additional cover traffic generated by clients to balance mixnodes and maintain network equilibrium. However, this process increases the power consumption for users. Therefore, to assess the situation where we do not apply cover traffic methods in the mixnet, we performed an experiment to measure the average number of overloaded mixnodes under

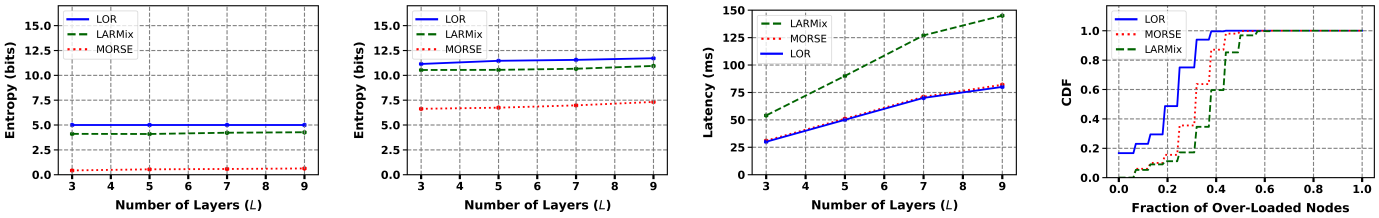
[11] As our evaluation setting resembles that of LARMix, please refer to pages 6, 7, and 8 of LARMix [15] for further clarity on the simulation settings.

[12] However, it should be noted that increasing the size of the network requires more mixnodes, implying a larger network needs a greater message influx. Otherwise, this can be detrimental to anonymity.



(a) Routing anonymity ($\mathcal{H}(f_R)$) (b) Message anonymity ($\mathcal{H}(m)$) (c) Average latency (d) Percentage of imbalance

Fig. 4



(a) Routing anonymity ($\mathcal{H}(f_R)$) (b) Message anonymity ($\mathcal{H}(m)$) (c) Average latency (d) Percentage of imbalance

Fig. 5

different low-latency routing strategies. The results of this experiment, shown in Fig. 4d, are described as the cumulative distribution function (CDF) of the fraction of overloaded mixnodes in the mixnet as a result of low-latency routing with different numbers of nodes per layer.

Fig. 4d indicates that LOR consistently makes the network less imbalanced, as its CDF curve is always higher or equal to the other approaches. Specifically, the surge around 0 for LOR indicates that, in many cases, the number of overloaded mixnodes is negligible, highlighting a strong advantage of LOR compared to MORSE and LARMix, which do not exhibit the same feature. More specifically, in all cases, at most about 40% of the mixnodes will be overloaded using LOR. However, this is different for MORSE and LARMix, as both guarantee that at most 67% of the mixnodes will be overloaded in all scenarios, which is a much weaker guarantee than that of LOR. Moreover, MORSE is slightly better at keeping mixnodes balanced compared to LARMix.

In conclusion, we find that LOR is the most optimal setting for dramatically reducing latency with its intelligent approach. It also provides the highest anonymity with uniform routing, while causing the least load imbalance in the mix network.

B. Varying Number of Layers

The second set of experiments investigates the impact of the number of layers on mixnet performance while keeping the number of mixnodes per layer fixed at $W = 32$. We specifically analyze its impact on the anonymity of routing, the anonymity of messages, the average link delay latency within the mixnet, and the degree of imbalance in the mixnet.

Fig. 5a describes the effect of increasing the number of layers on the anonymity of routing ($\mathcal{H}(f_R)$). The figure shows that for the LOR methodology, which already maximizes anonymity to $\log(W)$, adding more layers does not affect the level of anonymity. However, for LARMix and MORSE methods, an increase in layers slightly enhances the anonymity. The anonymity of MORSE remains the lowest, around 1 bit, while LARMix provides anonymity approximately 1 bit lower than that of LOR.

Fig. 5b highlights the changes in the anonymity of messages when the number of layers changes from 3 to 9. As previously explored, since the anonymity of messages is a linear combination of mixing process entropy and routing entropy, we observe the same trends in the anonymity of messages as seen in routing anonymity. Additionally, adding more layers, resulting in mixing messages with slightly more traffic (increasing the entropy of the mixing process), slightly increases the anonymity of messages.

Fig. 5c represents the average link delay latency caused by routing messages through the mixnodes versus different numbers of layers. This figure shows that for all approaches, increasing the number of layers increases the latency due to messages traveling through more mixnodes, thus encountering higher link delays. The LARMix approach provides the least reduction in latency compared to LOR and MORSE approaches, which can provide up to a 50% larger reduction in latency, especially when $L = 9$. We also observe that both MORSE and LOR present similar levels of latency, with LOR exhibiting a slightly larger reduction in latency.

Fig. 5d finally reveals the cumulative distribution function (CDF) of the percentage of mixnodes that are overloaded in scenarios where cover traffic is not considered. Similar to before, we see a surge around 0 for LOR, indicating a high likelihood of most nodes being only slightly imbalanced when using LOR compared to LARMix and MORSE. Generally, LOR provides the best guarantee for the worst-case percentage of imbalanced mixnodes compared to other approaches. However, MORSE slightly outperforms LARMix in terms of

providing fewer imbalanced mixnodes.

C. Deployment Efficiency

TABLE I: Complexity of Low-Latency Routing Approaches

Low-Latency Strategy	Complexity
LOR	$O(N\log(N))$
LARMix	$O(N\log(N))$
MORSE	$O(N^6)$

In this section, we investigate the efficiency of different low-latency routing approaches in terms of their complexity, as summarized in Tab. I.

We start with the LOR approach, where a client who wants to route their messages through the mixnet must determine the routing for $L-1$ layers. For each layer, based on the previously chosen mixnode, the client should sort all the W mixnodes available at the succeeding layer. Using merge sort, this costs asymptotically $O(W\log(W))$, which should be applied for $L-1$ layers (except the last layer where the routing is uniform), resulting in a complexity of $O(LW\log(W)) = O(N\log(N))$. LARMix routing, on the other hand, requires $O(N\log(N))$ to sort the latency of all the nodes relative to each other, which serves as the basis for LARMix routing operations. Additionally, it involves an $O(N)$ operation to derive the probability distribution over succeeding mixnodes, leading to a total complexity of $O(N\log(N))$. MORSE routing, in contrast to other approaches, is based on solving a linear programming problem, which in the best case can be solved using Interior Point Methods. These methods operate in $O(n^3b)$, where b is the number of bits required for the representation of the numbers and n is the number of constraints. Based on Eq. (8) and Eq. (9), $n = N^2 + N$, leading to a complexity of $O(N^6b)$.

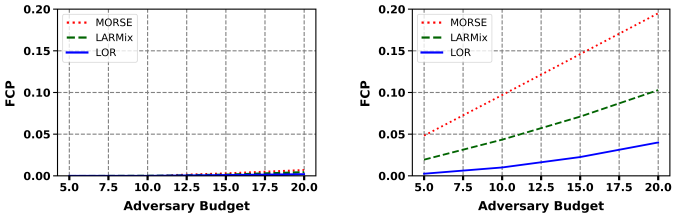
This analysis reveals that although LARMix is asymptotically as fast as LOR, LOR is more efficient as it only requires sorting the latency among mixnodes. In contrast, MORSE is the most computationally expensive in terms of deployment due to its higher complexity.

IV. Mixnode Adversary

In this section, we broaden the scope of adversarial analysis in mixnets, shifting from a GPA to a mixnode adversary. Unlike the GPA, this adversary can also own (corrupt) some mixnodes in the mixnet. The goal of this adversary is to maximize the fraction of fully corrupted paths (FCP), which are paths composed entirely of adversarial mixnodes, potentially leading to the full deanonymization of client communications passing through these paths. The adversary, however, can employ different strategies to corrupt these nodes. We first introduce these strategies and then measure the FCP in our proposed low-latency routing, alongside LARMix.

A. Random Strategy

The Random Strategy is the most basic approach to corrupting mixnodes within the mixnet. Here, the adversary controls C mixnodes selected at random from all mixnet layers.



(a) Random strategy (b) Greedy Corruption strategy

Fig. 6

Depending on the dynamics of selection, this approach can occasionally yield a high probability of intercepting paths that frequently incorporate critical mixnodes. However, at other times, it may result in less optimal node placements, targeting paths with lower selection probabilities.

B. Greedy Corruption Strategy

To comprehensively analyze the capability of a mixnode adversary, we introduce a greedy corruption strategy. This strategy assumes the adversary has complete information about active mixnodes in the network. The strategy begins by randomly corrupting a mixnode within the first layer of the mixnet. The adversary then assesses the distance from this compromised mixnode to all others in the succeeding layer, selecting those closest or with the highest probability of being the next hop. This process continues, choosing subsequent mixnodes that are nearest to the most recently corrupted node or those most likely to be selected next until L mixnodes are corrupted. The adversary then starts from a randomly corrupted mixnode in the first layer and repeats this process until all C mixnodes are corrupted.

C. FCP Evaluations

Considering the adversarial strategies to corrupt the mixnodes, we conducted an experiment where we measured the fraction of fully corrupted paths (FCP) when employing low-latency routing approaches. To do so, we first set $L = 3$ and $W = 60$, and then we modified the budget of the adversary (the percentage of mixnodes in the mixnet they can corrupt) utilizing both random and greedy corruption strategies.

Fig. 6a describes the FCP when deploying the random strategy for corrupting the mixnodes by an adversary whose budget varies from 5% to 20% corruption of the mixnet. As shown, for all the approaches, the FCP slightly increases as the budget increases. A larger budget means a higher chance of corrupting all the nodes in a message route, resulting in a higher FCP. However, when using this strategy, the FCP for all the approaches remains negligible.

On the other hand, Fig. 6b describes the FCP when applying the greedy corruption strategy by the adversary. In this case, the FCP also increases with the budget, but it increases more rapidly due to the greedy corruption algorithm's ability to

corrupt mixnodes in locations receiving the highest proportion of traffic.[13]

Fig. 6b additionally highlights that the LOR approach offers the least advantage to the adversary corrupting mixnodes. Specifically, in the worst case, it results in only four times the advantage compared to the random strategy and has a slow growth rate when increasing the budget. However, the LARMix strategy increases the advantage of the adversary using greedy corruption more rapidly. More accurately, it allows the corruption of 20% of nodes to intercept 10% of the traffic, leading to this amount of traffic being potentially deanonymized. Lastly, the MORSE approach performs the worst in the face of such adversaries, as it shows a linear growth where an adversary with an $\alpha\%$ budget can intercept $\alpha\%$ of the traffic.

V. Conclusion

This work investigated the effect of low-latency routing on anonymity in mixnets. We demonstrated that the anonymity of messages in a mixnet depends on both the mixing process at each intermediary mixnode and the routing policy. Based on these insights, we derived the MORSE and LOR proposals for low-latency routing. Compared to LARMix, these proposals reduce latency more effectively. However, we found that MORSE significantly compromises anonymity compared to LARMix, while LOR does not sacrifice anonymity. We further performed experiments to examine the imbalance in the mixnet after applying these methodologies. We observed that LOR causes the least imbalance among nodes in the network. Additionally, adversarial analysis shows that LOR gives the least advantage to adversaries intelligently corrupting mixnodes.

Overall, our findings indicate that if the goal is to reduce latency without sacrificing anonymity and maintain a balanced network, LOR is a highly suitable option. We hope the analysis and protocols presented in this work can be further developed into practical implementations. This can help expand the use cases of mixnets to latency-sensitive applications such as web browsing and instant messaging.

Acknowledgment

We would like to thank the anonymous reviewers for their valuable feedback. This research is partially supported by CyberSecurity Research Flanders with reference number VR20192203.

References

[1] D. Chaum, "The dining cryptographers problem: Unconditional sender and recipient untraceability," *Journal of cryptology*, vol. 1, pp. 65–75, 1988.

[2] D. L. Chaum, "Untraceable electronic mail, return addresses, and digital pseudonyms," *Communications of the ACM*, vol. 24, no. 2, pp. 84–90, 1981.

[3] R. Dingledine, N. Mathewson, P. F. Syverson *et al.*, "Tor: The second-generation onion router." in *USENIX security symposium*, vol. 4, 2004, pp. 303–320.

[4] A. M. Piotrowska, J. Hayes, T. Elahi, S. Meiser, and G. Danezis, "The loopix anonymity system," in *26th USENIX Security Symposium (USENIX Security 17)*, 2017, pp. 1199–1216.

[5] F. Shirazi, M. Simeonovski, M. R. Asghar, M. Backes, and C. Diaz, "A survey on routing in anonymous communication protocols," *ACM Computing Surveys (CSUR)*, vol. 51, no. 3, pp. 1–39, 2018.

[6] C. Diaz, H. Halpin, and A. Kiayias, "The nym network," 2021.

[7] C. Diaz, S. Seys, J. Claessens, and B. Preneel, "Towards measuring anonymity," in *International Workshop on Privacy Enhancing Technologies*. Springer, 2002, pp. 54–68.

[8] C. Diaz, "Anonymity and privacy in electronic services," *Heverlee: Katholieke Universiteit Leuven. Faculteit Ingenieurswetenschappen*, 2005.

[9] C. Diaz and B. Preneel, "Taxonomy of mixes and dummy traffic," in *Information Security Management, Education and Privacy: IFIP 18th World Computer Congress TC11 19th International Information Security Workshops 22–27 August 2004 Toulouse, France*. Springer, 2004, pp. 217–232.

[10] D. Kesdogan, J. Egner, and R. Büschkes, "Stop-and-go-mixes providing probabilistic anonymity in an open system," in *International Workshop on Information Hiding*. Springer, 1998, pp. 83–98.

[11] C. Diaz, S. J. Murdoch, and C. Troncoso, "Impact of network topology on anonymity and overhead in low-latency anonymity networks," in *Privacy Enhancing Technologies: 10th International Symposium, PETS 2010, Berlin, Germany, July 21-23, 2010. Proceedings 10*. Springer, 2010, pp. 184–201.

[12] D. Chaum, D. Das, F. Javani, A. Kate, A. Krasnova, J. De Ruiter, and A. T. Sherman, "cmix: Mixing with minimal real-time asymmetric cryptographic operations," in *Applied Cryptography and Network Security: 15th International Conference, ACNS 2017, Kanazawa, Japan, July 10-12, 2017, Proceedings 15*. Springer, 2017, pp. 557–578.

[13] A. Kwon, D. Lu, and S. Devadas, "{XRD}: Scalable messaging system with cryptographic privacy," in *17th USENIX Symposium on Networked Systems Design and Implementation (NSDI 20)*, 2020, pp. 759–776.

[14] U. Möller, L. Cottrell, P. Palfrader, and L. Sassaman, "Mixmaster protocol—version 2," 2003.

[15] M. Rahimi, P. K. Sharma, and C. Diaz, "Larmix: Latency-aware routing in mix networks," in *The Network and Distributed System Security Symposium*. Internet Society, 2024.

[16] M. Rahimi, "CLAM: client-aware routing in mix networks," in *Proceedings of the ACM Workshop on Information Hiding and Multimedia Security, IH&MMSec 2024, Baiona, Spain, June 24-26, 2024*. ACM, 2024, pp. 199–209. [Online]. Available: https://doi.org/10.1145/3658664.3659631

[17] ——, "Larmix++: Latency-aware routing in mix networks with free routes topology," in *International Conference on Cryptology and Network Security*. Springer, 2024, pp. 187–211.

[18] C. E. Shannon, "Communication theory of secrecy systems," *The Bell system technical journal*, vol. 28, no. 4, pp. 656–715, 1949.

[19] Python, "Event discrete, process based simulation for python." https://pypi.org/project/simpy/, 2013.

[20] R. N. Staff, "Ripe atlas: A global internet measurement network," *Internet Protocol Journal*, vol. 18, no. 3, pp. 2–26, 2015.

[21] I. Ben Guirat, D. Gosain, and C. Diaz, "Mixim: Mixnet design decisions and empirical evaluation," in *Proceedings of the 20th Workshop on Workshop on Privacy in the Electronic Society*, 2021, pp. 33–37.

[13] For FCP to be more prohibitive, we only consider the actual traffic to be intercepted, not cover traffic.