Multiparty Routing: Secure Routing for Mixnets

Fatemeh Shirazi  
imec - COSIC KU Leuven  
Leuven, Belgium

Elena Andreeva  
imec - COSIC KU Leuven  
Leuven, Belgium

Markulf Kohlweiss  
Microsoft Research  
Cambridge, UK

Claudia Diaz  
imec - COSIC KU Leuven  
Leuven, Belgium

Abstract—Anonymous communication networks are important building blocks for online privacy protection. One approach to achieve anonymity is to relay messages through multiple routers, where each router shuffles messages independently. To achieve anonymity, at least one router needs to be honest. In the presence of an adversary that is controlling a subset of the routers, unbiased routing is important for guaranteeing anonymity. However, the routing strategy also influenced other factors such as the scalability and the performance of the system. One solution is to use a fixed route for relaying all messages with many routers. If the route is not fixed, the routing decision can either be made by the communication initiator or the intermediate routers. However, the existing routing types each have limitations. For example, one faces scalability issues when increasing the throughput of systems with fixed routes. Moreover, when the routing decision is left to the initiator, the initiator needs to maintain an up-to-date view of the system at all times, which also does not scale. If the routing decision is left to intermediate routers, the routing of the communication can be influenced by an adversary. In this work, we propose a novel multiparty routing approach for anonymous communication that addresses these shortcomings. We distribute the routing decision and verify the correctness of routing to achieve routing integrity. More concretely, we provide a mixnet design that uses our routing approach and that in addition, addresses load balancing. We show that our system is secure against a global active adversary.

Keywords—Anonymous Communication; Mixnets; Routing; Privacy.

I. INTRODUCTION

Anonymous communication networks (ACNs) provide secure communication channels that hide not only data contents but also the communications' metadata, thus providing protection against traffic analysis. By routing data through multiple proxies, ACNs prevent network eavesdroppers from linking the source and destination of messages. Furthermore, ACNs enable senders to communicate with potentially malicious destinations without revealing their identity or location.

ACNs are typically overlay networks consisting of a set of dedicated routers, also called relays, that are connected via the Internet. The first ACN design in the early 1980s [1] introduced the concept of mixes for implementing anonymous email. These mixes are relays that collect a batch of N equal-size messages, cryptographically transform them, and output them in a re-shuffled form, so that the output messages cannot be traced to their corresponding inputs based on message content, size, timing or order. The best known ACN is Tor [2]. The Tor network is an ACN that consists of about seven thousands relays run by volunteers, and provides service to an estimated two million daily users who use it primarily to anonymously browse the web [3]. Other examples of ACN networks include for example Mixminion [4], an anonymous re-mailer, Freenet [5], [6] used for anonymous file-sharing, and DC-Nets [7] that can be deployed for broadcast applications such as group messaging.

The goal of ACNs is to anonymize communications by relaying them over multiple routers. There are three main types of anonymous routing in terms of how routers are chosen to form the path.

First, in deterministic routing, the paths are predetermined by the system configuration. Chaum’s original ACN proposal [1] considered a sequence of mixes organized in a cascade. Systems that adopted the cascade network topology in their designs include JAP [8], and voting systems, such as Helios [9]. An advantage of cascades is that routing headers are unnecessary, since the next hop in the path is predetermined. Cascades however, are of limited practicality due to their poor scalability, limited anonymity, and lack of resilience to router failures [10].

Second, in source routing, the full path is chosen by the initiator of the communication. This is the routing used by Tor [2] and by anonymous remailers such as Mixmaster [11] and Mixminion [4]. In comparison to cascades, source routing is more resilient against failures of routers [12]; it is more scalable, meaning that the network can grow to accommodate more users; and consequently, can provide better anonymity to those users.

However, an important requirement in source routing is that the initiator must have a full view of the network, since otherwise partial network knowledge can be exploited by an adversary to fingerprint and identify users based on which relays they choose [13]. As a consequence, all users need to periodically download the full list of currently available routers, their addresses, public keys, and other routing-relevant information. This is a major bottleneck for the scalability of the network, particularly if the routers have a high churn – as is often the case in peer-to-peer networks.

Third, in hop-by-hop routing each router locally decides on the next router of the path. Hop-by-hop routing is advantageous over source routing because it does not require entities to maintain an up-to-date view of the full network, as well as allows for better load balancing. The main disadvantage of hop-by-hop routing is that it is vulnerable to route capture attacks, where a malicious relay can gain full control over the route of a message, so that it is relayed through adversarial relays [14]. The earliest ACN employing hop-by-hop routing is Crowds [15], and another example is Morphmix [16], which uses witnesses to mitigate route capture attacks.

In this paper we propose multiparty routing, a novel type of anonymous routing that broadens the design space with a
fourth kind of routing. Our solution comes with important advantages over the previous routing approaches. Multiparty routing offers a scalability advantage over source routing: in our proposal the communication initiators need to only know one (or a small number) of routers, compared to the full view required in source routing. At the same time, we prevent the route capture attacks of hop-by-hop routing by decentralizing the routing decision using a multiparty computation technique. In addition to anonymity, our protocol guarantees the integrity of both shuffles and routing, which is novel for ACNs.

Our main contribution is a concrete multiparty routing protocol that employs cryptographic primitives, such as commitment schemes, signatures, and cryptographic hash functions. We combine our routing protocol with a mix network that employs provable shuffles, threshold decryption, and re-randomizable encryption to realize secure and verifiable anonymous communication, and load balancing. Our solution achieves security against global active adversaries and offers improved resilience and scalability against relay failures compared to deterministic and source routing. We achieve close to optimal load balancing by integrating relevant information about the throughput of relays into the routing strategy. Our protocol is a mid-latency anonymous communication network and thus can be deployed for applications that are more tolerant to latency such as anonymous wiki, micro-blogging, voting, and auctions.

II. RELATED WORK

Golle and Juels introduced Parallel Mixes [17], an ACN that uses both deterministic and hop-by-hop routing in two different phases that they call rotation and distribution phases, respectively. Typical configurations start with a rotation phase, then go through a distribution phase, and finish with a rotation phase. During a rotation phase, mixes exchange their entire output ciphertexts with another mix in a rotating fashion (that is, the first mix passes its ciphertexts to the second, and so on, and the last mix passes them to the first). In the distribution phase, each mix splits output ciphertexts among other mixes. Our routing protocol has similarities with the distribution phase of Parallel Mixes, but takes away one primary requirement of Parallel Mixes: Parallel Mixes require that in the distribution phase mixes must not be aware of the sender of any of the ciphertexts. Consequently, users have to submit input ciphertexts uniformly to the system and ciphertexts have to be mixed by at least one honest mix during the rotation phase. This is necessary to prevent route-capture attacks, because mixes in the distribution phase are free to choose how to distribute their output ciphertexts to other mixes (contrary to our protocol, where this decision is made jointly by multiple entities). However, these countermeasures are still insufficient when the adversary controls a fraction of the messages going through the network, as shown by an attack by Borisov [18].

Movahedi et al. [19] proposed a multiparty shuffling protocol that allows multiple parties to jointly compute a private random permutation of input messages. The parties that perform the multiparty shuffling protocol are split up into quorums that use a sorting scheme on inputs that have been shared among the quorum to achieve random permutation. While the approach is resilient to parties aborting the protocol or dropping out because of technical difficulties, there is no mechanism for preventing traffic tampering before the shuffling starts. Therefore, parties who hold user inputs at the beginning of the shuffling can replace them with arbitrary adversarial inputs. The approach is therefore only suitable for passive adversary. Furthermore, this scheme requires that the links between honest parties must be private and cannot be observed by the adversary, and the adversary needs to be static, meaning it cannot corrupt alternative parties once the protocol is running.

III. GOALS, ASSUMPTIONS, AND SYSTEM ARCHITECTURE

In this section, we describe the security model, the goals of the system, the entities and their roles, and the threat model.

A. Adversary Model

We assume a global adversary that can monitor all the communication links between entities in the system, including channels from users to the system and vice versa. The adversary is active, meaning that it can corrupt a subset of entities in the system, as well as submit messages herself. We assume that all cryptographic primitives are secure and cannot be broken by the adversary. Our protocol does not enforce restrictions on how many users need to be honest. However, a user is anonymous only among honest users that submitted messages to the network during the same time frame, and whose messages have traversed at least one honest relay.

B. Goals of the System

The goal of our system are as follows:

- Anonymity: the adversary must not be able to link a message leaving the system to the user who sent it, provided that the message has traversed at least one honest relay. This goal is equivalent to sender anonymity that has been defined by Pfitzmann et al. [20].
- Routing Integrity: the adversary must not be able to influence the route followed by a message, and consequently cannot bias the routing by forcing the data to only traverse adversarial relays.
- Availability: the system must be robust to the adversarial removal of relays, and output all messages even if some relays are taken down once the routing has started.
- Load-balancing: we aim at providing load balancing such that relays route an amount of traffic that is proportional to their throughput.

C. Entities and Roles

We consider a broadcast channel that we call bulletin board. Although the bulletin board is readable by all parties in the system: users and servers, users can not write into it.

Our system comprises a set of users and dedicated servers as depicted in Fig. 1. The users in the system want to send messages anonymously, whereas the servers cooperate with each other to achieve secure anonymous routing. We distinguish between three sets of servers: \( R \) - relays, \( RE \) - routing entities, and \( AS \) - auditing servers. The relays \( R \) are mixes that relay messages via publishing ciphertext messages on the bulletin board and fetching them. More precisely, the mixes first collect
from the bulletin board their set of input ciphertexts, then re-encrypt and shuffle those, a process referred to as mixing, and then relay their outputs to the bulletin board, so that each of them can be retrieved by the next mix in the path. The relays are organized in \( l \) sequential layers formed by disjoint subsets of \( R \).

The second set of servers are called routing entities \( RE \). These are responsible for computing the joint randomness that is used in the routing process.

The third set of servers are the auditors (auditing servers) \( AS \). Auditors are responsible for: 1. generating and broadcasting the public encryption key; 2. verifying the correctness of message shuffling and routing and 3. final message decryption.

For clarity, we keep the roles of auditors and routing entities separate from mixes. However, from a practical point of view all three roles can be carried out by the same set of servers.

Next, we briefly review the required cryptographic building blocks and proceed by describing our protocol.

**D. Cryptographic Building Blocks**

Next, we review relevant cryptographic building blocks.

**Public Key Encryption and Signing** We assume the reader is familiar with the principles of public key encryption and signature schemes. In our protocol, we employ a variant of public key encryption, referred to as **threshold decryption** [21], [22], which allows for the decryption to be carried out by a quorum of participants (the auditors in our scheme). More precisely, our scheme uses a distributed key generation protocol **KeyGen** such as the ones of [23], [24]. It generates a public key which is made available to all users of the system, while the corresponding secret key is distributed to all the auditors \( AS \) in a secret shared form. During the decryption protocol **Dec** a threshold of the secret key share holders (auditors) jointly decrypt the input ciphertext. This process can be can be realized either with a trusted dealer, such as in the work of [25], or without [26], [27].

To conceal how ciphertexts traverse a mix, we use re-randomizable encryption, such as ElGamal [28]. This allows a mix to update the randomness of the traversing ciphertexts. We propose standard cryptographic signatures, such as [29], for the routing verification.

**Shuffle and Correctness of Shuffling** A shuffle algorithm **Shuffle** implements a permutation on its inputs. In our protocol we use shuffles to permute the ciphertexts to conceal their ordering. To guarantee the correct execution of the shuffling algorithm we apply a **proof of correct shuffling** which is realized by a zero-knowledge (ZK) proof systems.

In our scheme, given a set of pre-shuffling and post-shuffling ciphertexts, a prover (the mix) needs to prove to a verifier (the auditors or, in case the auditor role is carried out by mixes, other mixes) that the set of post-shuffling ciphertexts is a re-randomization of the set of the pre-shuffling ciphertexts (without revealing the permutation and the re-randomization randomness used by **Shuffle**). Our protocol utilizes non-interactive variants of ZK proofs, such as the ones proposed in [30] [31].

**IV. OUR MULTIPARTY ROUTING PROTOCOL**

**A. System Overview**

As mentioned before, mixes are divided into a number of subsets (layers). Messages are relayed by one mix from each layer of the system. Therefore, each message is relayed by a set of mixes. Users send their messages to an arbitrary mix from the first layer of mixes \( R_1 \). Only messages that are submitted within the same time frame are mixed with each other. Hence, a user is only anonymous among honest users that submitted their messages at the same time frame.

Each relay mixes all message it receives. The next mix is determined in a hop-by-hop fashion. However, the mixes do not have freedom to choose to which mixes they send any given output message. The routing entities, \( RE \), determine how each message is routed by jointly producing the randomness used for the routing; this process, called **Multiparty Routing**, can be summarized as follows.

1) All routing entities choose privately a random number and use a commitment scheme to commit to their random numbers on the bulletin board.
2) After all routing entities have committed to their random numbers, they all open their commitments.
3) The random numbers are combined to produce a joint random number.
4) This random number is used to compute the routing assignment of the ciphertext messages to the mixes in the next layer in proportion to their capacity/throughput.

This procedure is carried out for the output of each mix. In parallel to this, the auditing servers need to verify the correctness of the shuffling and routing. After a message has been mixed by a mix from each layer and the auditing servers successfully verify the correctness of all routing and shuffling processes the message is delivered to the intended recipient of the message. If at least one of the mixes that has relayed the message is honest, the adversary cannot trace back a message leaving the network to any particular sender.

Figure 2 shows the topology for a network that has four layers (disjoint subset) of mixes. After processing input messages, the MPR protocol assigns output messages to mixes in the next layer; meaning that a portion of the output messages of each mix is assigned to mixes in the next layer in proportion to the throughput of those mixes within their own layer. Hence, the
routing pattern between the mixes follows a stratified topology as investigated by Diaz et al. \cite{32}.

B. Main Protocol

We denote the set of users by $U$. The set $\mathcal{R}$ of relays is divided into $l$ layers formed by mutually exclusive subsets $\mathcal{R}_i$ with the union $\bigcup_{i=1}^l \mathcal{R}_i = \mathcal{R}$. An individual mix is denoted by $r$.

The bulletin board contains lists of routing and verification information (input and output ciphertexts, corresponding signatures, commitments and their corresponding openings), and NIZK proofs of shuffles, grouped and indexed by $ctr$ going from 1 to $l$).

Encryption: The auditors (auditing servers) $\mathcal{AS}$ are responsible for the generation of the encryption key $pke$. Each auditing server has only a share of the private key $\{skd_{AS_1}, \ldots, skd_{AS_d}\}$ where $|\mathcal{AS}| = d$. Only qualified subsets of the auditing servers can jointly decrypt ciphertexts using their private key shares.

The protocol starts with each user encrypting her message $m$ together with its intended receiver under the public key $pke$ (known to all users) and sending the corresponding ciphertext to an arbitrary mix of $\mathcal{R}_1$. The set of ciphertexts that enter the system in the same time (by all senders) is denoted by $C_0$. Our system mixes ciphertexts in multiple layers with layer $\mathcal{R}_i$ producing the ciphertext set $C_i$. After completing $l$ layers the protocol terminates with a verification step and the decryption of $C_l$.

Mixing: Once a subset of the ciphertexts are assigned to a mix $r$ the mixing and routing take place. Mixes in the first layer receive ciphertexts from users, while mixes for other layers take ciphertexts from the bulletin board. The mixing consists of the re-encryption of the ciphertext inputs and their shuffling by a random permutation $\phi$.

After the output ciphertexts have been submitted to the bulletin board, the next mix to shuffle a ciphertext is determined by the following routing assignment procedure.

Routing Assignment: We use a coin-tossing protocol such as \cite{33} to obtain joint randomness. In the coin-tossing protocol the routing entities commit to their randomness using a cryptographic commitment schemes. Once the commitments are opened and the random strings are revealed on the bulletin board, a simple function, e.g. bitwise XOR or modular addition, is applied to obtain the joint randomness $Rand$.

Then we apply a function Assign on $Rand$. Assign then outputs the next layer mix index assignment list $\mathcal{L} = (z_1, \ldots, z_m)$ with $k$ elements in correspondence with the capacity/throughput of these mixes.

We give an example instantiation of Assign via the two procedures $H$ and Map in Appendix A in Algorithms 2 and 5.

The example in Figure 3 illustrates the routing assignment. Let us have a relay which outputs ciphertexts $c_1, c_2, c_3$ and the next layer of relays consist of two relays $r_1, r_2$ with corresponding capacities $b_1 = 1, b_2 = 2$. By applying Assign we want to determine which relays in the layer receive which ciphertexts. Namely, for an output vector of indices $(z_1 = 2, z_2 = 1, z_3 = 2)$ output by Assign the outgoing ciphertext is going to be fetched in the next step as follows. In our case, $r_1$ will then fetch the first $\frac{1}{1+2} = \frac{1}{3}$ of the outgoing ciphertexts, namely $c_{z_1} = c_2$, and $r_2$ fetches the rest $\frac{2}{1+2} = \frac{2}{3}$ ciphertexts, $c_{z_2} = c_1$ and $c_{z_3} = c_3$ respectively.

Proof of Mixing and Routing: The correctness of the mixing process is accommodated by the bulletin board. After mixing the set of input and corresponding set of output ciphertexts the mix needs to sign input and output ciphertexts. For signing, both the routing and mixing, $\mathcal{R}$ use their signing private key $sks$. The corresponding verifying public key $pke$ is available to all parties for later verification purposes. These keys are used as long-term keys, unlike $(pke, (skd_1, \ldots, skd_d))$ that are issued fresh for each time-frame.

Then, to prove correctness of protocol execution each mix needs to post on the bulletin board: a) inputs and outputs and their corresponding signatures, and b) a non-interactive zero-knowledge (NIZK) proof of correct mixing. Also, the bulletin board maintains the commitments for the randomness used for the MPR routing algorithm and their subsequent openings posted by the routing entities.

Verification and Decryption: Ciphertext which have passed through all layers are decrypted by the auditing servers upon successful verification of the correctness of: 1) Mixing and 2) Routing.

After all mixes in a layer published their results on the bulletin board, the auditing servers verify their mixing process by verifying all NIZK proofs on the bulletin board. Moreover, the auditing servers carry out routing verification on the input/output ciphertext signed on the bulletin board for each layer as soon as input ciphertexts for the next layer start being published to the bulletin board. The routing verification takes
the randomness (commitments openings) and computes the assignment for the next mix layer. Then, correct routing is confirmed by verifying that the computed assignment corresponds to the correct input and output ciphertexts in the bulletin board. With Algorithm 4 we give an example instantiation of the routing verification algorithm in Appendix B.

If routing or mixing verification fails, meaning that a mix has taken wrong ciphertext as input or produced an invalid NIZK proof, the auditing servers record the verification failure of the corresponding mix on the bulletin board. If routing and mixing of a mix does not verify the routing entities generate a fresh \( \text{Rand} \) to assign the input ciphertexts from the mix that has performed wrong routing to the remaining mixes in his layer. All remaining output ciphertexts in the layer need to wait until this stage is carried out to be routed to next layers.

The integrity and authenticity of bulletin board submissions is verified by the use of the signatures schemes. In our protocol we use randomized algorithms and with \( \xi \) we denote the randomized output generation, so \( \xi \leftarrow S_w \) samples randomly from all permutations of size \( w \) and \( s \) refers to the randomness used by the ciphertext re-randomization algorithms.

Our protocol is summarized in Algorithm 1.

### Algorithm 1 Multiparty Routing

**Setup**
- Auditing servers \( \mathcal{AS} \) generate encryption keys:
  - \( (pke, skd_{\mathcal{AS}1}, \ldots, skd_{\mathcal{AS}d}) \) \( \xleftarrow{\$} \) KeyGen
  - publish \( pke \) to bulletin board

**Encrypt**
- Each sender:
  - \( c \xleftarrow{\$} \text{Enc}_{pke}(\text{receiver's address}, m) \)
  - sends \( c \) to \( r \in \mathcal{R} \)
  - \( ctr \xleftarrow{\$} 1 \)

**while** \( ctr \leq l \)**

**Mixing and Prove Mixing**
- Each \( r \in \mathcal{R}_{ctr} \), with inputs \( C = \{c_1, \ldots, c_w\} \) and \( \phi \xleftarrow{\$} S_w \):
  - shuffles \( C^\phi \xleftarrow{\$} \text{Shuffle}(C, \phi) \)
  - re-encrypts \( (C', s) \xleftarrow{\$} \text{ReEnc}(C^\phi) \)
- Each \( r \in \mathcal{R}_{ctr} \):
  - computes \( \pi \xleftarrow{\$} \text{ProveNIZK}(C, C'; \phi, s) \)
  - publish \( \pi \) to bulletin board
  - publish \( (\text{Sign}_{skv_r}(C), \text{Sign}_{skv_r}(C'; C, C')) \) to the bulletin board

**Next hop and Prove Routing**
- For each \( r \in \mathcal{R}_{ctr} \), all \( v \) routing entities \( \mathcal{RE} \):
  - publish \( \text{Com}_{\text{rand}_1}, \ldots, \text{Com}_{\text{rand}_v} \) to the bulletin board
  - publish \( \text{rand}_1, \ldots, \text{rand}_v \) to the bulletin board
- For each \( r_{out} \in \mathcal{R}_{ctr} \), each \( r \in \mathcal{R}_{ctr+1} \) computes:
  - \( \text{Rand} \xleftarrow{\$} \text{SUM}(\text{rand}_1, \ldots, \text{rand}_v) \)
  - \( (z_1, \ldots, z_w) \xleftarrow{\$} \text{Assign}(\text{Rand}) \)
  - For \( i = 1 \) to \( w \):
    - \( r_{zi} \in \mathcal{R}_{ctr+1} \) fetch \( c_i \)

**Verify Mixing and Routing**
- For each \( r \in \mathcal{R}_{ctr} \), a threshold \( z \) of auditing servers \( \mathcal{AS} \) runs
  - \( \text{VerifyNIZK}(C_r, C'_r, \pi) \)
  - \( \text{Verify}_{\text{pkv}_r}(\text{Sign}_{skv_r}(C_r), C_r) \)
  - \( \text{Verify}_{\text{pkv}_r}(\text{Sign}_{skv_r}(C'_r), C'_r) \)
  - and computes
    - \( \text{VerifyRouting}(\text{Rand}_r, C'_r, \mathcal{R}_{ctr+1}) \)

**Decrypt**
- Each \( r \in \mathcal{R}_d \):
  - sends all \( c' \) to \( \mathcal{AS} \)
- For a threshold \( z \) of auditing servers \( \mathcal{AS} \):
  - if \( \text{Verify Mixing and Routing} \) is successful for all layers
    - decrypts (receiver’s address, \( m \)) \( \xleftarrow{\$} \text{Dec}_{skd_1 \ldots skd_d}(c') \)
    - otherwise returns an error
    - send \( m \) to \( \text{receiver from } U \)

C. Parameter Choices

Our protocol has similarities to parallel mixing in terms of distributing ciphertexts among a set of mixes, hence, we revisit suggested parameters by the literature and adopt these parameters for our system.

The number of mixes that relay the ciphertexts, or so-called path length, is an important aspect of our system because it determines how many times a batch of ciphertexts \( C \) is re-arranged. The path length has a direct effect on whether a ciphertext batch \( C \) entering the system is fully mixed with the other batches entering at the same time. Our result confirms the intuition that a larger number of layers leads to a better mixing.

For a passive adversary, Czumaj shows that the path length in any switching networks should be \( \log n \), where \( n \) is the total number of mixes in the system \([34] \ [35]\). Similarly, Klonowski et al. \([36]\) have also shown that in parallel mixing the number of distribution steps should be \( \log n \).

For an active adversary, Klonowski \([36]\) shows that the number of mixing layers must depend on the number of messages that are relayed through the system. Similarly, Goodrich and Mitzenmacher \([37]\) have shown that if some mixes are controlled by an active adversary then the number of layers needs to be \( \log N \) where \( N \) is the number of input messages in a given time frame.

Since, we are considering an active adversary we set the number of layers to \( \log N \). For example, if we assume the routing of 1000 messages our system needs at least 3 layers where each layer has the throughput to relay 1000 messages. For relaying 100,000 messages the system needs to have at least 5 layers where each layer has the throughput to relay 100,000 messages. To improve availability each layer should have at least two mixes.
Furthermore, to increase the cost of an adversary we assume that mixes are set up across distinct IP subnets by multiple organizations or institutions. For example, an organization can offer more than one mix for the network, however, all mixes of an organization need to be in the same layer. This is important because otherwise a message can be effectively mixed by a single organization.

Another important parameter is the capacity or throughput of the mixes. From a load balancing point of view the total throughput of mixes in each layer needs to be equal in order to prevent any congestion in our system.

V. SECURITY ANALYSIS

In this section we start by revisiting the security properties of interest and go on to analyze the security of our protocol based on those security goals against a global, active adversary. We show that an active adversary in our solution has no advantage over a passive one.

Security Properties:

- **Routing Integrity**: guarantees that an active adversary is not able to influence the routing decision to de anonymize messages by forcing them to route through adversarial mixes.
- **Routing Correctness**: guarantees the detection of malicious routing and subsequently only correctly routed messages leaving the system.
- **Anonymity (routing confidentiality)**: guarantees that the adversary cannot find the relationship between messages leaving the system and a user sending a message to the system, provided that the message has traversed at least some honest mixes.
- **Availability**: guarantees that the protocol is robust against removal of a subset of entities by an adversary and that the messages entering the system will be output by the system in the face of such attacks.

A. Routing Integrity

To address the integrity property we answer the following question: Can a malicious mix (or a subset of mixes) influence the routing of a ciphertext $c$ to a mix of his choice $r_A$? We investigate two ways of manipulating the routing decision. The first way deals with adversaries who try to influence the shuffling procedure. This attack path is however not possible in our protocol because all mixes publish their shuffled ciphertexts before the routing decision ($\text{Rand}$ outcome) is revealed. The second way to manipulate the routing decision is through adversarial bias in the computation of the joint randomness $\text{Rand}$.

More concretely, we investigate the advantage of an active adversary $A$ who gains control over a subset of routing entities denoted as $\mathcal{RE}_A$. In order to influence the routing decision $A$ has to be able to bias the computation of $\text{Rand}$ where $\text{Rand} = \text{Rand}_{\text{honest}} + rand_A$ is the mod sum addition of the random numbers from all honest entities $\text{Rand}_{\text{honest}}$ and the adversarially controlled randomness $rand_A$. The goal of $A$ is to produce a new $rand'_A$, such that the updated sum $\text{Rand}' = \text{Rand}_{\text{honest}} + rand'_A$ is valid and is used by the auditing server to verify that the ciphertext is from a valid mix. Since $A$ does not know in advance the openings of the honest routing entities, $A$ needs to ‘predict’ in advance a valid $rand'_A$, such that $T = \text{Com}(rand'_A)$. Our assumption here is that $\text{Com}$ is a secure scheme against collision, preimage and second preimage attacks. Under this assumption $A$’s best strategy is to perform precomputation to find the largest set of preimages $\mathcal{N}$ with $|\mathcal{N}| = n$ mapping to a commitment value $T$. Then, once $A$ commits to $T$, he ‘hopes’ that the needed randomness for his routing purposes $rand'_A$ belongs to the precomputed set $\mathcal{N}$. Then, what is left to find is the size of the set of valid randomness values $V$ for the routing purposes of $A$. Since the adversary controls $b/B$ of the throughput $w$ ciphertexts, then the set of valid values $V$ amounts to $\frac{b}{w} \times \frac{1}{2^{2|\text{Rand}|}}$. Our assumption here is that $\mathcal{Com}$ is collision secure, finding $n = 2$ values takes already precomputation time of birthday bound complexity $2^{\frac{|\text{Rand}|}{2}}$. This means that unless $A$ has a precomputational power of order $2^{\frac{|\text{Rand}|}{2}}$ its success probability in biasing the joint randomness $\text{Rand}$ is $\frac{1}{2^{2|\text{Rand}|} \times b/B}$.

Furthermore, if the adversary removes the mix that $c$ is assigned to to force a new routing assignment his chances to be assigned to $c$ are increasing only if he removes a large proportion of throughput of the layer of $r_A$.

B. Routing Correctness

In this scenario we investigate if a mix can unnoticeably deviate from the routing assignment. The adversary goal (acting as one of the mixes $R$) is to get a ciphertext of his choice regardless of the ciphertext’s next hop assignment. To achieve her goal the adversary here needs to manipulate the routing verification. To do so the adversary needs to control a subset of the auditing entities that read the routing information from the bulletin board and carry out the verification. But this is impossible in our protocol because this adversarial behavior is going to be discovered in the routing verification phase if at least one of the auditing servers is honest. Remember that one of the tasks of the auditing servers is the verification of the correct input/output ciphertext routing and their signatures.

C. Anonymity

The adversary’s aim is to trace the path of a ciphertext $c$ until it leaves the system. To link an input ciphertext to an output ciphertext the adversary needs to control all mixes that have relayed the input ciphertext. Otherwise, if at least one mix on the ciphertext’s route is honest, the adversary will not
be able to trace this link. Hence, our adversary needs to control at least one mix in each layer of the system. Our decision to follow a stratified topology disallows the adversary to gain any advantage from placing his efforts (i.e. throughput) unevenly as compared to balancing out its efforts for each layer. For example, if an adversary controls 20% of the total throughput of the system, we assume that he is in fact controlling a 20% of the throughput of the mixes in each layer.

To estimate the probability that a ciphertext follows a fixed path, denoted by $c \rightarrow p$, then we just need to multiply the probability of that ciphertext being routed by the individual mixes in each layer of the path. Let us assume that a path $p$ consists of an order list of mixes $p = \{r_1, r_2, \ldots, r_l\}$, where $l$ is the number of layers of the system.

$$Pr[c \rightarrow p] = \prod_{i=1}^{l} \frac{b_i}{B}.$$  

Moreover, the probability of choosing each of the mixes in a path is computed as follows. where $b_i$ is the relative throughput of the mix $r_i$ compared to the total throughput $B$ of all mixes (recall that it is the same for all layers). If all mixes in the subset have the same capacity $b$, then $Pr[c \rightarrow p] = \left(\frac{b}{B}\right)^l$.

We are interested in the probability that a ciphertext is routed only by adversarial mixes because this leads to deanonymize the ciphertext leaving the system by the adversary. We compare the probability that the ciphertext relayed by our MPR protocol is only routed through adversarial mixes to the probability that the ciphertext relayed by regular parallel mixing system is only relayed by adversarial mixes. Figure 4 shows the probability that a ciphertext routed through only adversarial mixes, for our MPR protocol and for parallel mixing system, when the adversary is controlling 0.1%, 0.15%, 0.20%, 0.25%, 0.30%, 0.35%, and 0.40% of the throughput in each layer of the system. For this figure we assume the system has 4 layers. When regular parallel mixing is used the probability that a ciphertext is only routed along adversarial mixes is in proportion to the resources of the adversary. While if our MPR protocol is used, this probability is increasing much slower due to the fact that adversarial mixes cannot steer ciphertexts toward further adversarial mixes.

Figure 5 shows the probability that a ciphertext is routed through only adversarial mixes, for our MPR protocol and for parallel mixing system, when the system has 3, 4, 5, 6, or 7 layers. We assume for this figure that the adversary is controlling 0.25% of the throughput of each layer of the system. If regular parallel mixing is used, the probability that a ciphertext is only routed along adversarial mixes is constant and equal to the proportion of adversarial mixes irrespective of the number of layers. While when our MPR protocol is used, this probability is significantly lower compared to parallel mixes even when the system consists of only 3 layers. This probability decreases further as the number of layers in the system increases.

Moreover, recall that in Parallel Mixes it is required to distribute the input messages uniformly before submitting them to the system to decrease the chances of carrying out attacks such as blending attacks. However, permuting incoming messages before entering the system to obtain a uniform distribution would need another randomly distributing mechanism and complicates the trust model. Furthermore, Parallel Mixes require two rotating phases before and after a single redistribution phase where the input ciphertexts are rotated $m + 1$ times, where $m$ is the number of malicious mixes. Using MPR there is no need to enter only uniformly distributed messages to the system and the rotating phases of Parallel Mixes can be eliminated; this allows to use multiple distribution phases instead of only one, which contributes to obtaining a complete random permutation of all ciphertexts.

### D. Availability

The goal of the adversary is to take out stakeholders of the protocol in order to obstruct the protocol from running.

There is a restriction on how many mixes, routing entities, and auditors the adversary can take out. Taking out auditors, $AS$ is least beneficial, because the attack is not resulting in any advantage for the adversary. While taking out routing entities, $RE$ increases the adversary’s success in biasing routing to a
very small extent. Taking out mixes will not abrupt the protocol execution but leads to traffic congestion. Below we investigate the effect of removing each type of entity from the system.

Routing entities RE: If the routing entity fails before committing to a random number the protocol continues to operate as long as it is assumed that at least one honest routing entity has remained. If the routing entity fails after committing to a random number it is considered an abortion, although it might have been a system failure. If this happens more than once, the remaining routing entities vote the corresponding routing out.

Auditors AS: If a limited number of auditors fail to carry out the distributed decryption, then the protocol is guaranteed to continue execution because we are using threshold cryptography. To carry out the routing and mixing verification the only restriction is that at least one remaining auditor is not controlled by the adversary.

Mixes R: Our protocol is resilient to the failure of a number of mixes and preserves load balanced among the remaining mixes. We describe how the protocol recovers after the failure of a mix or a number of mixes in different positions as follows.

1) If a mix fails at the first layer after the users have sent their ciphertexts the users machine needs to notice in the bulletin board that her ciphertext is not in the input of any of the entry mixes and send their ciphertext to another entry mix.
2) If the adversary takes out all mixes in the first layer the user’s machine will notice that the ciphertext sent by her is not in the bulletin board and the auditors notice the absence of any output from the first layer and introduce new entry points for the users.
3) If a mix fails at one of the layers except of the first layer before generating the output ciphertext and the corresponding routing information, the auditors notice this and inform the routing entities to generate a new Rand that assigns the input ciphertext of the failed mix to the remaining mixes in its layer.
4) If the adversary takes out all mixes in one layers except of the first layer, the auditors inform the next layer of mixes to take over with fresh random numbers.

VI. DISCUSSION

Decentralization Advantages: MPR decentralizes the routing decision which leads to several advantages over a routing decision that is not decentralized. Troncoso et al. [38] present numerous advantages of decentralization in protocol design such as increasing the adversaries cost and facilitating public verifiability. Our MPR protocol also increases the cost of the adversary for attacking the routing integrity and facilitates public verifiability.

Latency We estimate for a system with 9 mixes having equal throughput, arranged in 3 layers each layer having the throughput to relay 900 messages, and where 1000 messages were sent to the system, messages takes about 20 seconds to traverse the system. This is, however, only a preliminary result tested on a Intel Xeon E5-2640 CPU, 2.50 GHz machine with 24 cores, where all computation are performed online. If some of the computations are made offline the latency can improve further. Increasing the input traffic and consequently the number of layers of the system strengthens anonymity, but also increases the latency. There is a trade-off between latency and anonymity. Our system offers almost optimal load balancing, when congestion is minimized. As long as the adversary does not take out mixes the latency of the system thus increases linearly with the amount of traffic. Our protocol is a medium-latency anonymous communication network and thus can be deployed for applications that are more tolerant to latency such as anonymous wiki, micro-blogging, voting, and auctions. Note that for application such as voting with homomorphic tallying the decryption phase of our protocol may not be necessary for all messages.

Load Balancing: Since, our MPR protocol offers load balancing the messages leaving the system have a more uniform anonymity set size compared to mixnets using source routing because in source routing the routes are selected without considering load balancing. This increases the minimum anonymity that a message can achieve.

Limitation As mentioned before, MPR addresses the scalability issue of source routing in terms of not requiring to maintain a complete system view for users. Growing our system is, however, not as scalable as growing a system with a free topology because the system’s throughput can be only expanded by adding mixes to all layers, while if there is no restrictions for selecting a relay and the routing can choose any relay in the system adding a single relay increases the system’s overall throughput.

VII. CONCLUSION

We proposed a novel routing approach, called multiparty routing, to carry out secure routing for anonymous communication networks. More concretely, our MPR protocol achieves routing integrity, routing correctness, load balancing, in addition to anonymity, and is relatively scalable.

We leave as future work a full implementation of our protocol and the addition of resilience mechanism against failing mixes explained in the security analysis. Moreover, we leave as an open question measurements of the performance and resilience of our system under DoS attacks.

APPENDIX

In this section, we give instantions of routing assignment and routing verification.

A. Routing Assignment

To route a ciphertext, at each layer, a relay is assigned to the ciphertext using a mapping function. The mapping function is composed of a permutation function keyed by the jointly generated random number Rand and a load-balancing function that distributes ciphertext based on the throughput of relays in the next layer.

The permutation algorithm is the function \( H \) out of hash function \( h \). \( H \) is applied iteratively on the inputs: joint random number \( Rand \) and the ciphertext indexes \( 1, \ldots, w \) and returns the permuted indexes \( z_1, \ldots, z_w \). To realize the permutation functionality we have to avoid hash function collisions by
For this purpose we use the function $fetch$. $C_r$ is assigned to them as their input from the bulletin board. These relays fetch the ciphertext that performs a check of membership on a set $Z$ of already computed index values.

We apply the mapping function $Map$, e.g., Algorithm 3 in appendices, to make the assignment of the permuted ciphertext (indexed by $H$) to the relays in the next layer in proportion to their throughput. We refer to the throughput of relays by $b_i$, where throughput refers to the number of messages it can route in proportion to all the number of message that can be routed by a layer and the throughput of a single layer is denoted $B$. For this purpose we use the function $fetch$, which ciphertexts to mixes in the next layer. These relays fetch the ciphertext that is assigned to them as their input from the bulletin board.

**B. Routing Verification**

Let $C_r$ refer to the set of input ciphertext that relay $r$ has submitted and signed and $C'_r$ to the set of output ciphertexts that relay $r$ has submitted and signed. Moreover, let us denote the $ctr$-th layer of routers in our system by $R_{ctr}$.

Then all the auditors must carry out the routing verification as described below. Our algorithm applies a Boolean $VerifyRand$ function which simply verifies both the correctness of the opening of the commitments $Com(rand)$ with regard to the actual $rand$ on the bulletin board and checks that their sum equals $Rand$. Moreover, the algorithm verifies whether the input of all relays in the $ctr + 1$ layer is correct according to the output ciphertexts of relays in layer $ctr$ and their corresponding $Rand$.

**REFERENCES**

[1] D. Chaum, “Untraceable electronic mail, return addresses, and digital pseudonyms,” Commun. ACM, vol. 24, no. 2, pp. 84–88, 1981.

[2] R. Dingledine, N. Mathewson, and P. Syverson, “Tor: The second-generation onion router,” in Proceedings of the 13th Conference on USENIX Security Symposium - Volume 13, SSYM '04, pp. 303–320, USENIX Association, 2004.

[3] “Tor Project: Anonymity Online.” [https://www.torproject.org/][17]. Last accessed: July 30, 2017.

[4] G. Danezis, R. Dingledine, and N. Mathewson, “Mixminion: Design of a type III anonymous remailer protocol,” in 2003 IEEE Symposium on Security and Privacy (SP 2003), 11–14 May 2003, Berkeley, CA, USA, pp. 2–15, 2003.

[5] I. Clarke, O. Sandberg, B. Wiley, and T. W. Hong, “Freenet: A distributed anonymous information storage and retrieval system,” in International Workshop on Designing Privacy Enhancing Technologies: Design Issues in Anonymity and Unobservability, pp. 46–66, Springer-Verlag New York, Inc., 2001.

[6] I. Clarke, O. Sandberg, M. Toseland, and V. Verendel, “Private communication through a network of trusted connections: The dark freenet,” Network, 2010.

[7] H. Corrigan-Gibbs and B. Ford, “Dissent: Accountable anonymous group messaging,” in Proceedings of the 17th ACM Conference on Computer and Communications Security, CCS ’10, pp. 340–350, 2010.

[8] O. Berthold, H. Federrath, and S. Köpsell, “Web MIXes: A system for anonymous and unobservable internet access,” in Designing Privacy Enhancing Technologies, International Workshop on Design Issues in Anonymity and Unobservability, Berkeley, CA, USA, July 25-26, 2000, Proceedings, pp. 115–129, 2000.

[9] B. Adida, “Helios: Web-based open-audit voting,” in Proceedings of the 17th Conference on Security Symposium, SS ’08, (Berkeley, CA, USA), pp. 335–346, USENIX Association, 2008.

[10] A. Kwon, H. Corrigan-Gibbs, S. Devadas, and B. Ford, “Atom: Scalable anonymity resistant to traffic analysis,” arXiv preprint arXiv:1612.07841, 2016.

[11] U. Möller, L. Cottrell, P. Falstrader, and L. Sassaman, “Mixmaster protocol - version 2,” 2003.

[12] F. Shirazi, C. Diaz, and J. Wright, “Towards measuring resilience in anonymous communication networks,” in Proceedings of the 14th ACM Workshop on Privacy in the Electronic Society, WPES ’15, (New York, NY, USA), pp. 95–99, ACM, 2015.

[13] G. Danezis and P. Syverson, “Bridging and fingerprinting: Epistemic attacks on route selection,” in Proceedings of the 8th International Symposium on Privacy Enhancing Technologies, PETS ‘08, (Berlin, Heidelberg), pp. 151–166, Springer-Verlag, 2008.

[14] G. Danezis and R. Clayton, “Route fingerprinting in anonymous communications,” in Peer-to-Peer Computing, 2006. P2P 2006. Sixth IEEE International Conference on, pp. 69–72, IEEE, 2006.

[15] M. K. Reiter and A. D. Rubin, “Crowds: Anonymity for web transactions,” ACM Trans. Inf. Syst. Secur., vol. 1, pp. 66–92, November 1998.

[16] M. Rennhard and B. Plattner, “Introducing MorphMix: Peer-to-peer based anonymous internet usage with collusion detection,” in Proceed-
ings of the 2002 ACM Workshop on Privacy in the Electronic Society, WPES ’02, (New York, NY, USA), pp. 91–102, ACM, 2002.

[17] P. Golle and A. Juels, “Parallel mixing,” in Proceedings of the 11th ACM Conference on Computer and Communications Security, CCS ’04, (New York, NY, USA), pp. 220–226, ACM, 2004.

[18] N. Borisov, “An analysis of parallel mixing with attacker-controlled inputs,” in Proceedings of the 5th International Conference on Privacy Enhancing Technologies, PET’05, (Berlin, Heidelberg), pp. 12–25, Springer-Verlag, 2006.

[19] M. Movahedi, J. Saia, and M. Zamani, Secure Multi-party Shuffling, pp. 459–473. Cham: Springer International Publishing, 2015.

[20] A. Pfitzmann and M. Köhntopp, “Anonymity, unobservability, and pseudonymity - A proposal for terminology,” in Designing Privacy Enhancing Technologies, International Workshop on Design Issues in Anonymity and Unobservability, Berkeley, CA, USA, July 25-26, 2000, Proceedings, pp. 1–9, 2000.

[21] Y. G. Desmedt and Y. Frankel, “Threshold cryptosystems,” in Proceedings on Advances in Cryptology, CRYPTO ’89, (New York, NY, USA), pp. 307–315, Springer-Verlag New York, Inc., 1989.

[22] A. Shamir, “How to share a secret,” Commun. ACM, vol. 22, pp. 612–613, Nov. 1979.

[23] T. P. Pedersen, “Non-interactive and information-theoretic secure verifiable secret sharing,” in Proceedings of the 11th Annual International Conference on Advances in Cryptology, CRYPTO ’91, (London, UK, UK), pp. 129–140, Springer-Verlag, 1992.

[24] R. Gennaro, S. Jarecki, H. Krawczyk, and T. Rabin, Secure Distributed Key Generation for Discrete-Log Based Cryptosystems, pp. 295–310. Berlin, Heidelberg: Springer Berlin Heidelberg, 1999.

[25] V. Shoup and R. Gennaro, Securing threshold cryptosystems against chosen ciphertext attack, pp. 1–16. Berlin, Heidelberg: Springer Berlin Heidelberg, 1998.

[26] T. P. Pedersen, “A threshold cryptosystem without a trusted party,” in Proceedings of the 10th Annual International Conference on Theory and Application of Cryptographic Techniques, EUROCRYPT’91, (Berlin, Heidelberg), pp. 522–526, Springer-Verlag, 1991.

[27] R. Cramer, R. Gennaro, and B. Schoenmakers, “A secure and optimally efficient multi-authority election scheme,” in Proceedings of the 16th Annual International Conference on Theory and Application of Cryptographic Techniques, EUROCRYPT’97, (Berlin, Heidelberg), pp. 103–118, Springer-Verlag, 1997.

[28] T. ElGamal, “A public key cryptosystem and a signature scheme based on discrete logarithms,” in Advances in Cryptology (G. Blakley and D. Chaum, eds.), vol. 196 of Lecture Notes in Computer Science, pp. 10–18, Springer Berlin Heidelberg, 1985.

[29] C.-P. Schnorr, “Efficient identification and signatures for smart cards,” in Proceedings of the 9th Annual International Conference on Advances in Cryptology, CRYPTO ’89, (London, UK, UK), pp. 239–252, Springer-Verlag, 1990.

[30] D. Wikström, “A commitment-consistent proof of a shuffle,” in Information Security and Privacy, 14th Australasian Conference, ACISP 2009, Brisbane, Australia, July 1-3, 2009, Proceedings, pp. 407–421, 2009.

[31] S. Bayer and J. Groth, Efficient Zero-Knowledge Argument for Correctness of a Shuffle, pp. 263–280. Berlin, Heidelberg: Springer Berlin Heidelberg, 2012.

[32] C. Diaz, S. J. Murdoch, and C. Troncoso, “Impact of network topology on anonymity and overhead in low-latency anonymity networks,” in Proceedings of the 10th International Conference on Privacy Enhancing Technologies, PETS’10, (Berlin, Heidelberg), pp. 184–201, Springer-Verlag, 2010.

[33] M. Blum, “Coin flipping by telephone a protocol for solving impossible problems,” SIGACT News, vol. 15, pp. 23–27, Jan. 1983.

[34] A. Czumaj, P. Kanarek, K. Loryś, and M. Kutylowski, Switching Networks for Generating Random Permutations, pp. 25–61. Boston, MA: Springer US, 2001.

[35] A. Czumaj, “Random permutations using switching networks,” in Proceedings of the Forty-seventh Annual ACM Symposium on Theory of Computing, STOC ’15, (New York, NY, USA), pp. 703–712, ACM, 2015.

[36] M. Klionowski and M. Kutylowski, Provable Anonymity for Networks of Mixes, pp. 26–38. Berlin, Heidelberg: Springer Berlin Heidelberg, 2005.

[37] M. T. Goodrich and M. Mitzenmacher, Anonymous Card Shuffling and Its Applications to Parallel Mixnets, pp. 549–560. Berlin, Heidelberg: Springer Berlin Heidelberg, 2012.

[38] C. Troncoso, G. Danezis, M. Isaakidis, and H. Halpin, “Systematizing decentralization and privacy: Lessons from 15 years of research and deployments,” arXiv preprint arXiv:1704.08065, 2017.