Mixim: A General Purpose Simulator for mixnet

Iness Ben Guirat
KU Leuven
ibenguir@esat.kuleuven.be

Devashish Gosain
IIIT-Delhi
devashishg@iiitd.ac.in

Claudia Diaz
KU Leuven
cdiaz@esat.kuleuven.be

1. INTRODUCTION

Mixes are routers that transform the appearance and order of messages so that their inputs and outputs are unlinkable. Mixes are organized in networks called *mixnets* that route messages via multiple mixes in order to provide anonymity. A variety of mixnets have been proposed in prior work, including Mixminion [2], Vuvuzela [10], and Loopix [9]. These systems have different architectures, threat models, design principles, features and parameters. Moreover, the proposals have been evaluated using different methods and anonymity metrics, making it difficult to benchmark and compare them.

Anonymity systems, including mixnets, are known to require a tradeoff between anonymity, latency and bandwidth, in what is called the *Anonymity Trilemma* [4]. Modifying design features and parameters has an impact on these tradeoffs that is difficult to model analytically. Instead, the evaluation of mixnets is normally done empirically via simulations. Existing mixnet evaluations are done with *ad hoc* simulators that do not enable comparisons between mixnet designs.

In this work we present Mixim and its supported features. With this general-purpose simulator we can evaluate entropy-based and indistinguishability-based anonymity for a wide variety of design options and parameters across mixnet designs. In section 3 we present two experiments studying the impact of different parameters on the entropy-based anonymity, we leave the indistinguishability out of scope for this paper.

2. SUPPORTED FEATURES OF THE SIMULATOR

In this section we present the features that are supported by our simulator. In Table 1, we highlight the three broad categories of features currently supported by our simulator including different types of network topology, mixes and routing.

Table 1: Features of the simulator

Topologies	Mix-type	Routing
Stratified	Timed	Source Routing
Free route	Pool	Hop-by-Hop
Cascade	Stop-and-Go	

2.1 Topologies

The topology is the structure of the mixes in the mixnet. In our simulator we consider three main topologies: cascades, free routing and stratified topologies. In a **cascade topology** every message goes through a number of mixes in a pre-determined order, with each mix sending all messages to the next mix in the cascade. Some variations on this design, like cMix [1] allow multiple cascades to run in parallel which is also feasible on our simulator. In a **free routing topology**, a message may go through any path and the path length may be variable [7]. Note that it becomes very difficult to analyze anonymity in these topologies as the network scales [3] so our simulator does not currently support the computation of anonymity in such topologies. Mixes in a **stratified topology** are arranged in a fixed number of layers where each mix, at any given time, is assigned to one specific layer. The layers are interconnected such that each mix in layer i is connected with every mix in layers $i-1$ and $i+1$, while the first layer only receives incoming messages from senders and the last layer sends outgoing messages to the final recipients.

2.2 Mix-type

Each mix takes input messages and outputs (flushes) them with a different appearance and in a different order. Thereby, mixes hide the correspondence between inputs and outputs, such that an adversary is not able to establish a correlation between input and output messages neither based on message content nor on timing. Our simulator supports the following types of mixes:

- **Timed Mix**: At each time $T = t$, the mix flushes all the messages it contains.
- **Pool Mix**: The mix will flush N messages when $n+N$ messages have been received, keeping n messages in an internal pool to be mixed with messages received in the next round.
- **Stop-and-Go Mix**: The delays messages independently with the delay following an exponential distribution. The memoryless property of exponential distributions implies that when a mix that contains N messages sends one message out, all N input messages are equally likely to be the output [8, 9].

2.3 Routing

Routing determines how messages are sent through the network, including the way mixes are selected for relaying a message. Note that in anonymous communication networks, routing determines to a large extent the security and performance of the entire system [5]. Although there exists different types of routing such that rendezvous and multi-party routing, we have only implemented hop-by-hop and source routing as they are the two most popular types.

- **Hop-by-hop**: The sender only selects the first mix, which in turn picks the second, and so on, until the message reaches its final destination.
- **Source routing**: When the sender of the communication selects the set of mixes that will form the entire route. This is the design that is currently the most common in Tor [6] and modern mixnet designs [9].

3. EXPERIMENTS

In this section we present two sets of experiments that demonstrate the efficacy of Mixim. In both experiments we used a stratified topology (shown in Figure 1) where all the mixes are Poisson mixes and each layer has 3 mixes.

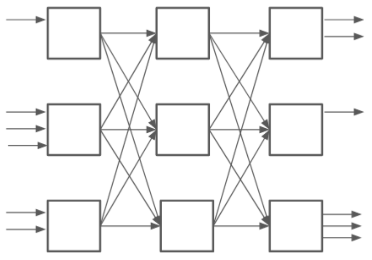

Figure 1: Stratified topology

3.1 Mix corruption

In the first experiment we study the impact of the position of corrupted mixes on anonymity. We first test a baseline of our experiment: We generate 100k messages per time unit and each mix has an average delay of 1/100 time units[1]. Then we study the impact of having 3 corrupted mixes in the first layer of the network and compare it to 3 corrupted mixes each mix in a different layer. We can see from Figure 2 that the position of the corrupted mixes has a significant impact on anonymity. Even though the capability of the adversary is similar, the entropy has a much higher variance in the last experiment due to the fact that certain messages will go through more than just one corrupted mix. In the worst case of the third scenario, a message may go through a fully corrupted path and in that case the adversary can fully deanonymize the message.

3.2 Impact of the Mixnet layers on anonymity

In Figure 3, we show the impact of adding layers to a stratified topology. As we can see, having just one layer partitions the anonymity set into as many subsets as mixes existing in the layer. Adding a second layer substantially strengthens anonymity: an increase of 2.5 bits implies that the anonymity set becomes more than five times larger. Additional layers bring diminishing returns while increasing the cost in terms of latency and bandwidth.

[1] Each mix will delay every message according to a Poisson process with parameter μ, therefore the average delay is $1/\mu$

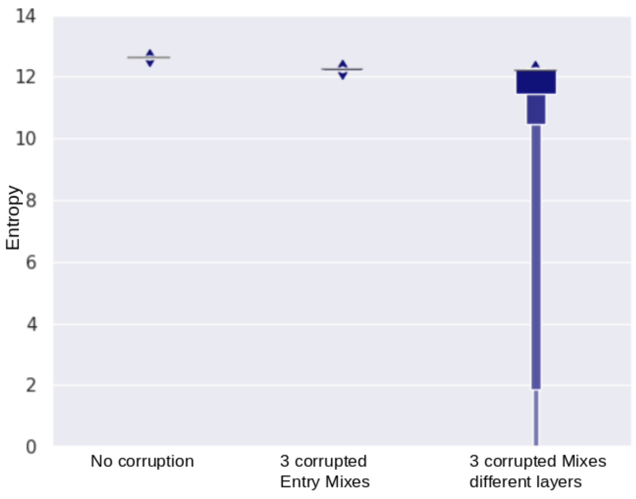

Figure 2: Position of corrupted mixes

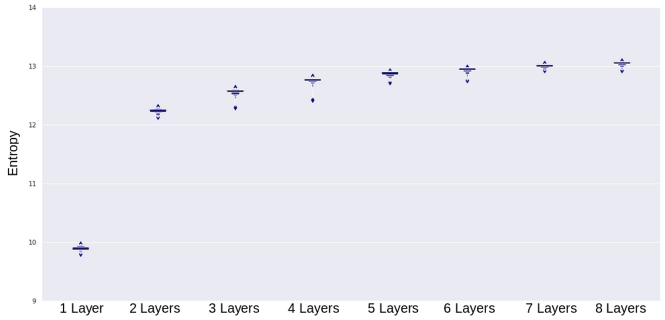

Figure 3: Impact of different layers on anonymity

4. CONCLUSION

In this paper, we presented the design and implementation of Mixim, a mixnet simulator that is able to design and evaluate mixnet-based systems. We performed experiments that show the power of simulations to analyze and compare mixnet designs. As future work, we will further develop Mixim to be suitable for studying a wide range of problems including network congestion, dummy traffic strategies, and the impact of different traffic distributions on privacy properties. Mixim may also be used to study the impact of different attacks on different designs and ultimately the different privacy properties these designs provide.

5. ACKNOWLEDGEMENTS

We would like to thank Harry Halpin and Karim Eldafrawi and Tariq Elahi for discussions and comments on this work.

6. REFERENCES

[1] D. Chaum, F. Javani, A. Kate, A. Krasnova, J. de Ruiter, A. T. Sherman, and D. Das. cmix: Anonymization by high-performance scalable mixing. In *USENIX Security*, 2016.

[2] G. Danezis, R. Dingledine, and N. Mathewson. Mixminion: Design of a type iii anonymous remailer protocol. In *2003 Symposium on Security and Privacy, 2003.*, pages 2–15. IEEE, 2003.

[3] G. Danezis and C. Troncoso. Vida: How to use bayesian inference to de-anonymize persistent communications. In *International Symposium on Privacy Enhancing Technologies Symposium*, pages 56–72. Springer, 2009.

[4] D. Das, S. Meiser, E. Mohammadi, and A. Kate. Anonymity trilemma: Strong anonymity, low bandwidth overhead, low latency-choose two. In *2018 IEEE Symposium on Security and Privacy (SP)*, pages 108–126. IEEE, 2018.

[5] R. Dingledine and N. Mathewson. Anonymity loves company: Usability and the network effect. In *WEIS*, 2006.

[6] R. Dingledine, N. Mathewson, and P. Syverson. Tor: The second-generation onion router. Technical report, Naval Research Lab Washington DC, 2004.

[7] R. Dingledine, V. Shmatikov, and P. Syverson. Synchronous batching: From cascades to free routes. In *International Workshop on Privacy Enhancing Technologies*, pages 186–206. Springer, 2004.

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

[9] A. M. Piotrowska, J. Hayes, T. Elahi, S. Meiser, and G. Danezis. The loopix anonymity system. In *26th {USENIX} Security Symposium ({USENIX} Security 17)*, pages 1199–1216, 2017.

[10] J. Van Den Hooff, D. Lazar, M. Zaharia, and N. Zeldovich. Vuvuzela: Scalable private messaging resistant to traffic analysis. In *Proceedings of the 25th Symposium on Operating Systems Principles*, pages 137–152, 2015.