POSTER: Traffic Analysis Attacks in Anonymity Networks

Katharina Kohls
Ruhr-University Bochum
katharina.kohls@rub.de

Christina Pöpper
New York University Abu Dhabi
christina.poepper@nyu.edu

ABSTRACT

With more than 1.7 million daily users, Tor is a large-scale anonymity network that helps people to protect their identities in the Internet. Tor provides low-latency transmissions that can serve a wide range of applications including web browsing, which renders it an easily accessible tool for a large user base. Unfortunately, its wide adoption makes Tor a valuable target for de-anonymization attacks. Recent work proved that powerful traffic analysis attacks exist which enable an adversary to relate traffic streams in the network and identify users and accessed contents. One open research question in the field of anonymity networks therefore addresses efficient countermeasures to the class of traffic analysis attacks. Defensive techniques must improve the security features of existing networks while still providing an acceptable performance that can maintain the wide acceptance of a system. The proposed work presents an analysis of mixing strategies as a countermeasure to traffic analysis attacks in Tor. First simulation results indicate the security gains and performance impairments of three main mixing strategies.

Keywords

Anonymity Networks, Mix, Tor

1. INTRODUCTION

While using the Internet, we leave traces of personal information. Such data can reveal sensitive details about personal lives, can harm people that decide to share political statements, or act as whistleblowers. Anonymity networks aim to protect such sensitive user information by separating the identities of individuals from the contents they access in the Internet. With more than 1.7 million daily users, Tor is one prominent example in this context. Tor is a volunteer-operated anonymity network that offers strong security features like onion encryption and provides low-latency transmissions that allow for interactive applications such as web browsing. The latter renders Tor a secure alternative for everyday use cases.

Unfortunately, this performance comes at the expense of known vulnerabilities against traffic analysis attacks. Recent work presented multiple active [12, 8, 3] and passive [10, 15, 18] traffic analysis attacks that aim to de-anonymize users on the basis of transmission traces that get monitored in different nodes of the network. Two main factors influence the success of such attacks: First, the adversary uses a defined set of *attack metrics* to identify relations between monitored traffic streams. In, e.g., a confirmation attack, the ingress traffic that enters the network is monitored along with the egress traffic between the last relay and destination server of a connection. Given the attack metrics the adversary attempts to detect similarities in the ingress and egress traffic to relate the incoming and outgoing streams. If the metric can reliably distinguish the monitored data, the adversary is capable of matching the identity of a user (relates to the ingress connection) to the accessed contents (relates to the egress connection). This allows for the de-anonymization of users.

Second, the number of nodes controlled by the adversary can increase the probability of monitoring related connections. So-called routing attacks [1, 17, 16] help to improve the situation of an adversary by forcing connections to traverse compromised nodes. An empirical study of 2016 [14] revealed that up to 40 % of circuits in the Tor network are vulnerable to traffic analysis attacks, if the adversary acts on the level of autonomous systems. For state-level adversaries or in case of collusion, this can be increased to a coverage of up to 85 %. That is, on average nearly half of Tor circuits are vulnerable to attacks that allow for the de-anonymization of users.

Traffic analysis attacks become possible because the low-latency transmissions in Tor preserve relations of packets in a transmission stream. Metadata information like inter-packet timing or packet counts can be analyzed by the adversary and are used in the attack metrics to identify related streams. In contrast to Tor, classical mix networks [4] and anonymous remailers [5] disrupt metadata relations by adding artificial delays during the transmission process. While this technique protects from traffic analysis attacks, it results in high latencies that prevent interactive applications. Given this trade-off between performance and security, available systems either can serve latency-sensitive use cases or provide superior security features.

Given the current landscape of anonymity networks the need for efficient countermeasures becomes obvious. To over-

ASIA CCS '17 April 02-06, 2017, Abu Dhabi, United Arab Emirates

ACM ISBN 978-1-4503-4944-4/17/04.

DOI: http://dx.doi.org/10.1145/3052973.3055159

come the current shortcomings of Tor in context of traffic analysis attacks, we aim to integrate mixing techniques in the transmission procedures of Tor relays. In contrast to the strategies of classical mix networks, a strict limitation of additional delays should induce short additional latencies for a connection and with that provide acceptable performance rates. At the same time the minimal perturbation of traffic streams should disrupt some relations between ingress and egress traces to an extent that protects against passive traffic analysis attacks.

In the proposed poster we present general mix concepts adapted for an integration in Tor relays. We backup the concept of low-latency mixing in Tor with first simulation results that focus on the performance and security characteristics of different mixing strategies. In short, we make the following main contributions:

- We identify three main mixing strategies, namely batch mixing, continuous-time mixing, and dummy traffic injection, and analyze their security and performance capabilities.
- We provide simulation results that give a first impression of the efficiency of all three mixing strategies for different parameter setups.
- We suggest an experimental setup for realistic measurements with the proposed mixing strategies that allows for a performance and security analysis without harming any real Tor users.

Our proposed concept of mixing is fully software-based and therefore can be integrated into existing Tor relays. It furthermore is backwards compatible in a sense that the mixing algorithm is self-contained and does not disrupt the transmissions of other relays that do not implement this new functionality. The mixing procedure is parametrized and can be adapted to trade-off between performance and security.

2. COUNTERMEASURES TO TRAFFIC ANALYSIS ATTACKS

We focus on two main design aspects for anonymity systems. First, their security features define the expected protection against traffic analysis attacks. Second, the performance of a network limits the range of applications that can be served. Mix-based countermeasures for Tor can be considered satisfactory if they reduce the success of attacks on Tor, e.g., passive traffic analysis attacks like confirmation, and preserve acceptable performance rates at the same time.

There are three abstract mix concepts that are candidate countermeasures in the described context. *Batch* mixes [6, 5] store all incoming packets at a node and flush a defined portion of packets after an event was triggered, e.g., the delay duration expired or a fixed number of packets was received. *Continuous-time* mixes [9, 7] assign individual (random) delays to packets in a node. Other than in batch mixes this allows for a constant emission of packets while at the same time relations between incoming and outgoing packets should be disrupted. *Dummy traffic injection* [2, 15] uses additional packets for traffic stream perturbation. Such injections do not necessarily carry any reasonable payload data and disrupt patterns without depending on additional delays. All three mix concepts can be adapted through individual parameters, e.g., the injection rate for dummy packets, flush rates, or delay durations.

In a first step, we define an abstract simulation model that analyzes different mixing strategies with respect to their effect on the traffic stream. These preliminary measurements help to understand the dynamics of mixing and give first insights regarding the capabilities of individual mixing strategies. Outgoing from these theoretical results we can continue to design an explicit mixing system that attaches to the current transmission procedures of Tor.

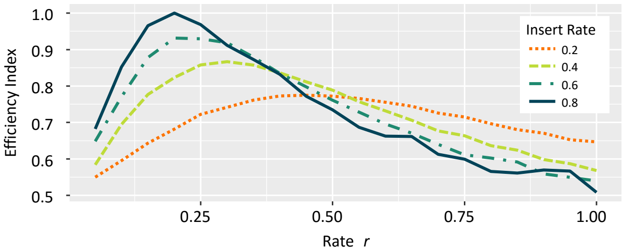

Figure 1: Results for 1000 random repetitions with a comparison of increasing insert rates for dummy traffic. We tested increasing insert rates (number of packets injected in a window) and increasing chaffing rates (number of windows affected by an injection).

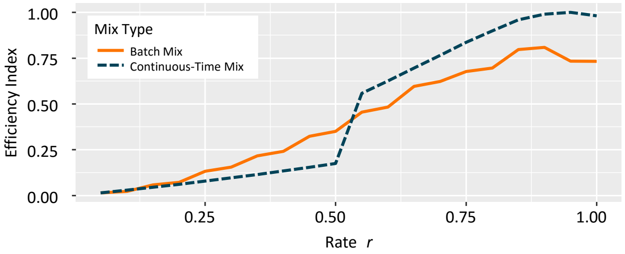

Figure 2: Results for 1000 random repetitions with a comparison of two mixing concepts. Both mixes use a rate parameter r that denotes increasing delays.

The results of both, the abstract comparison of different mixing strategies along with an explicit implementation of a mix for the Tor anonymity network reveal whether mixing in general can be used as a countermeasure to confirmation attacks.

3. RELATED WORK

Mix networks were originally introduced by Chaum [4] and provide the anonymous transmission of information at the expense of high latencies. In this original concept, a mix node gathers messages from multiple transmitters and stores them to disrupt the relation between incoming and outgoing traffic. In contrast to this, modern anonymity networks such as Tor, e.g., establish relay circuits that forward onion-encrypted packets with low latencies and with that can serve interactive applications such as web browsing.

Transmissions with low latency come at the expense of vulnerabilities to traffic analysis attacks [12, 3, 11, 13, 15]. In such attacks, an adversary passively monitors ingress and egress traffic to the network and de-anonymizes users through correlating similarities in the metadata of transmissions. Active attacks take the traffic analysis one step further and interact with transmissions, e.g., inject easily identifiable fingerprints.

While much offensive work was presented throughout the last years, there is a lack of realistic countermeasures to traffic analysis attacks. This applies especially for open issues in the Tor network which is still vulnerable against confirmation attacks. With its large user base, Tor represents a valuable target for the conduction of such attacks, which makes the analysis of potential protection mechanisms even more relevant.

4. REFERENCES

[1] K. Bauer, D. McCoy, D. Grunwald, T. Kohno, and D. Sicker. Low-resource routing attacks against tor. In *Proceedings of the 2007 ACM workshop on Privacy in electronic society*, pages 11–20. ACM, 2007.

[2] O. Berthold and H. Langos. Dummy traffic against long term intersection attacks. In *International Workshop on Privacy Enhancing Technologies*, pages 110–128. Springer, 2002.

[3] S. Chakravarty, A. Stavrou, and A. D. Keromytis. Traffc Analysis Against Low-Latency Anonymity Networks Using Available Bandwidth Estimation. In *European Conference on Research in Computer Security*, ESORICS '04, pages 249–267, Athens, Greece, Sept. 2010. Springer.

[4] D. L. Chaum. Untraceable Electronic Mail, Return Addresses, and Digital Pseudonyms. *Communications of the ACM*, 24(2):84–90, Feb. 1981.

[5] L. Cottrell. Mixmaster 2.0 remailer release! Usenet post, May 1995.

[6] G. Danezis, R. Dingledine, and N. Mathewson. Mixminion: Design of a Type III Anonymous Remailer Protocol. In *IEEE Symposium on Security and Privacy*, SP '03, pages 2–15, Oakland, CA, USA, May 2003. IEEE.

[7] R. Dingledine, A. Serjantov, and P. Syverson. Blending Different Latency Traffic with Alpha-mixing. In *Workshop on Privacy Enhancing Technologies*, PET '06, pages 245–257, Cambridge, UK, June 2006. Springer.

[8] X. Fu, B. Graham, R. Bettati, and W. Zhao. Active Traffic Analysis Attacks and Countermeasures. In *International Conference on Computer Networks and Mobile Computing*, ICCNMC '03, pages 31–39, Shanghai, China, Oct. 2003. IEEE.

[9] D. Kesdogan, J. Egner, and R. Büschkes. Stop-And-Go-MIXes Providing Probabilistic Anonymity in an Open System. In *International Workshop on Information Hiding*, IH '98, pages 83–98, Portland, OR, USA, Apr. 1998. Springer.

[10] B. N. Levine, M. K. Reiter, C. Wang, and M. Wright. Timing attacks in low-latency mix systems. In *International Conference on Financial Cryptography*, pages 251–265. Springer, 2004.

[11] P. Mittal, A. Khurshid, J. Juen, M. Caesar, and N. Borisov. Stealthy Traffic Analysis of Low-Latency Anonymous Communication Using Throughput Fingerprinting. In *ACM Conference on Computer and Communications Security*, CCS '11, pages 215–226, Chicago, IL, USA, Oct. 2011. ACM.

[12] S. J. Murdoch and G. Danezis. Low-Cost Traffic Analysis of Tor. In *IEEE Symposium on Security and Privacy*, SP '05, pages 183–195, Oakland, CA, USA, May 2005. IEEE.

[13] S. J. Murdoch and P. Zieliński. Sampled Traffic Analysis by Internet-Exchange-Level Adversaries. In *Workshop on Privacy Enhancing Technologies*, PET '07, pages 167–183, Ottawa, ON, Canada, June 2007. Springer.

[14] R. Nithyanand, O. Starov, A. Zair, P. Gill, and M. Schapira. Measuring and Mitigating AS-level Adversaries Against Tor. In *Symposium on Network and Distributed System Security*, NDSS '16, San Diego, CA, USA, Feb. 2016. Internet Society.

[15] V. Shmatikov and M.-H. Wang. Timing Analysis in Low-Latency Mix Networks: Attacks and Defenses. In *European Symposium on Research in Computer Security*, ESORICS '06, pages 18–33, Hamburg, Germany, Sept. 2006. Springer.

[16] Y. Sun, A. Edmundson, L. Vanbever, O. Li, J. Rexford, M. Chiang, and P. Mittal. RAPTOR: Routing Attacks on Privacy in Tor. In *USENIX Security Symposium*, USENIX '16, pages 271–286, Washington, DC, USA, Aug. 2015. USENIX.

[17] L. Vanbever, O. Li, J. Rexford, and P. Mittal. Anonymity on QuickSand: Using BGP to Compromise Tor. In *ACM Workshop on Hot Topics in Networks*, HotNets-XIII, Los Angeles, CA, USA, Oct. 2014. ACM.

[18] Y. Zhu and R. Bettati. Unmixing mix traffic. In *International Workshop on Privacy Enhancing Technologies*, pages 110–127. Springer, 2005.