Towards an Evaluation Framework for Tor Load Balancing

André Greuibel
University of Wuerzburg
Wuerzburg, Germany
andre.greuibel@uni-wuerzburg.de

Samuel Kounev
University of Wuerzburg
Wuerzburg, Germany
samuel.kounev@uni-wuerzburg.de

ABSTRACT
Tor is a widely used anonymization network. Traffic is routed over different relay nodes to conceal the communication partner. The Tor Load Balancing Mechanism (TLBM) is responsible for assigning traffic to relays. However, up to this point, there are no agreed-upon quality and security metrics for evaluating the quality of the TLBM. In practice, multiple paper provide their own evaluation using custom metrics incomparable to related work. This paper aims to start bridging this gap by arguing for the necessity of a unified framework to evaluate TLBMs and by giving an overview of aspects that should be included in such a framework.

1 INTRODUCTION
In Tor, clients can build privacy-preserving paths (circuits) consisting of relay nodes (relays) that forward encrypted TCP packets to other relays in a circuit or destination server on the internet. This way, only the communication from the client to Tor and from the last relay to the destination can be observed. As the available bandwidth of each relay is limited, the Tor Load Balancing Mechanism (TLBM) has to ensure that no relay attracts more traffic than it can handle. Additionally, circuits in Tor are anonymous, only if the attacker is not able to simultaneously surveil the first and last node of the circuit. Otherwise, a powerful traffic correlation attack potentially enables complete de-anonymization in that circuit [4, 6].

To decrease feasibility of this and similar (cf. [4]) attacks, traffic should be distributed to as many entities as possible. However, recent research has shown that the current measurement method determining the load distribution in Tor lacks in both security and quality [1, 3, 4, 12]. Furthermore, the current load distribution highly favors high-bandwidth relays [2, 4, 7, 8]. This leads to a reduction of resources necessary to perform several large-scale de-anonymization attacks by more than 80% [4]. Based on these shortcomings, several projects proposed replacements or improvements of the current TLBM, both in security and quality [3, 5, 9, 11].

2 GOALS OF A FRAMEWORK
Yet, despite the importance of the TLBM on the security of Tor and its known shortcomings, there currently is no consensus on how to evaluate a TLBM. In practice, all evaluations of the existing TLBM and all proposals for replacing or improving it provide their own evaluation method and evaluation metrics.

3 BLUEPRINT OF A FRAMEWORK
Such a framework should provide (1) a classification of TLBM approaches, (2) unified terminology, (3) metrics and criteria for evaluating a given TLBM.

The classification should include the general ideas used to design the TLBM. Most notably, it should be able to clearly communicate the security assumptions used to design the TLBM and its design approach (decentralized measurements, trusted measurement entities, semi-trusted measurement entities, ...). Furthermore, there should be exact definitions of the terms used. Most importantly, the term “bandwidth” and its various interpretations during measurements, aggregation, and load distribution in the TLBM should be standardized. An example of such an attempt (including nine different definitions of the term bandwidth) is presented in [4]. However, these definitions are limited to aspects relevant to a centralized measurement process (like the current TLBM) and are not directly applicable to decentralized approaches.

Lastly, such a framework should also include metrics and criteria that a secure TLBM of high quality should fulfill. In the remainder of this section, we will describe such criteria and metrics of the quality and security of the TLBM that we consider to be relevant for an evaluation framework. For better readability, we loosely categorized them into three categories (“security”, “load distribution”, and “other”).

3.1 Security Aspects
In this section, we describe security aspects that a TLBM should fulfill to reduce its attack surface.

Trust-Less The TLBM should not require trust in self-reported data from a non-trusted entity. This is desirable as it prevents relays from influencing the load distribution by lying about self-reported information.

Fair The TLBM is fair if no untrusted entity can be assigned more traffic by other means than actually providing more bandwidth in real-world communication. This is desirable as it prevents relays from influencing the TLBM by other
means than lying about its bandwidth. For example, if it is possible to detect bandwidth measurements in a TLBM, relays might artificially inflate their traffic by providing additional bandwidth during measurement only.

Anonymous The TLBM should not rely on collection, verification, or publication of real-world communication data – or information that could be used to reconstruct this data. This is desirable as Tor is designed to conceal communication between relays in a circuit and the latency between the relays. It is also dependent on client parameters like geolocation and relays already chosen for the circuit. As many of the current deficiencies are caused by deliberate choices aimed at increasing the user performance [4], this aspect should be included in a framework.

4 FUTURE WORK AND CONCLUSION

The most important step for building such a framework includes proposals on how to exactly define, quantify, and measure these aspects. Note that this is not trivial: For example, there are currently at least three different approaches to measure decentralization [2, 4, 8]. However, none of the approaches are able to measure a “ground truth”, and all three metrics proposed allow for various interpretations themselves. But even with a consensus on how to define, measure, and interpret these metrics, one has to discuss trade-offs: Some aspects, even in our preliminary list, directly contradict each other: For example, if Tor has differently sized entities, it cannot simultaneously be perfectly balanced and decentralized. Up to this date, there is no consensus on how these trade-offs should be designed and which aspects are most important.

As such, a complete framework most likely cannot be built by a single research group but rather has to be a project of the TLBM research community at large. Overall, we hope that, in this work-in-progress paper, we showed the need for such a project and offered fellow researchers a quick overview of aspects of interest for it.

REFERENCES