The Effect of DNS on Tor's Anonymity† (abstract)

Benjamin Greschbach* *KTH Royal Institute of Technology*
Tobias Pulls* *Karlstad University*
Laura M. Roberts* *Princeton University*
Philipp Winter* *Princeton University*
Nick Feamster *Princeton University*

I. Introduction

We have yet to learn how to build anonymity networks that resist global adversaries, provide low latency, and scale well. Remailer systems such as Mixmaster [6] and Mixminion [1] eschew low latency in favor of strong anonymity. In contrast, Tor [2] trades off strong anonymity to achieve low latency; Tor therefore enables latency-sensitive applications such as web browsing but is vulnerable to adversaries that can observe traffic both entering and exiting its network, thus enabling deanonymization. Although Tor does not consider global adversaries in its threat model, adversaries that can observe traffic for extended periods of time in multiple network locations (*i.e.*, "semi-global" adversaries) are a real concern [3, 5]; we need to better understand the nature to which these adversaries exist in operational networks and their ability to deanonymize users.

Past work has quantified the extent to which an adversary that observes TCP flows between clients and servers (*e.g.*, HTTP requests, BitTorrent connections, and IRC sessions) can correlate traffic flows between the client and the entry to the anonymity network and between the exit of the anonymity network and its ultimate destination [5, 7]. The ability to correlate these two flows—a so-called *correlation attack*—can link the sender and receiver of a traffic flow, thus compromising the anonymity of both endpoints. Although TCP connections are an important part of communications, the Domain Name System (DNS) traffic is also quite revealing: for example, even loading a single webpage can generate hundreds of DNS requests to many different domains. No previous analysis of correlation attacks has studied how DNS traffic can exacerbate these attacks.

DNS traffic is highly relevant for correlation attacks because it often traverses completely different paths and autonomous systems (ASes) than the subsequent corresponding TCP connections. An attacker that can observe occasional DNS requests may still be able to link both ends of the communication, even if the attacker cannot observe TCP traffic between the exit of the anonymity network and the server. Figure 1 illustrates how an adversary may monitor the connection between a user and the guard relay, and between the exit relay and its DNS resolvers or servers. This territory—to-date, completely unexplored—is the focus of this work.

We first explore how Tor exit relays resolve DNS names. By developing a new method to identify all exit relays' DNS resolvers, we learn that Google currently sees almost 40% of all DNS requests exiting the Tor network. Second, we investigate which organizations can observe DNS requests that originate from Tor exit relays. To answer this question, we emulate DNS resolution for the Alexa Top 1,000 domains from an autonomous system that is popular among exit relays. We find that DNS resolution for half of these domains traverses numerous ASes that are not traversed for the subsequent HTTP connection to the web site. Next, we show how the ability to observe DNS traffic from Tor exit relays can augment existing website fingerprinting attacks, yielding perfectly precise DefecTor[1] attacks for unpopular websites. We further introduce a new method to perform traceroutes from the networks where exit relays are located, making our results significantly more accurate and comprehensive than previous work. Finally, we use the Tor Path Simulator (TorPS) [4] to investigate the effects of Internet-scale DefecTor attacks.

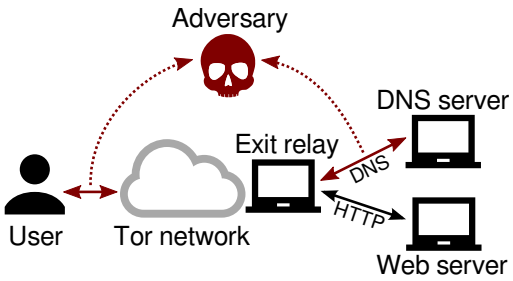

Fig. 1. Past traffic correlation studies have focused on linking the TCP stream entering the Tor network to the one(s) exiting the network. We show that an adversary can also link the associated DNS traffic, which can be exposed to many more ASes than the TCP stream.

We demonstrate that DNS requests significantly increase the opportunity for adversaries to perform correlation attacks. This finding should encourage future work on correlation attacks to consider both TCP traffic and the corresponding DNS traffic; future design decisions should also be cognizant of this threat. The measurement methods we use to evaluate the effects of traffic correlation attacks are also more accurate than past work. Our work (*i*) serves as guidance to Tor exit relay operators and Tor network developers, (*ii*) improves state-of-the-art measurement techniques for analysis of correlation attacks, and (*iii*) provides even stronger justification for introducing website fingerprinting defenses in Tor. To foster future work and facilitate the replication of our results, we publish both our code and datasets.[2]

References

[1] G. Danezis, R. Dingledine, and N. Mathewson, "Mixminion: Design of a type III anonymous remailer protocol," in *Security & Privacy*. IEEE, 2003. URL: https://nymity.ch/tor-dns/pdf/Danezis2003a.pdf

[2] R. Dingledine, N. Mathewson, and P. Syverson, "Tor: The second-generation onion router," in *USENIX Security*. USENIX, 2004. URL: https://nymity.ch/tor-dns/pdf/Dingledine2004a.pdf

[3] S. Farrell and H. Tschofenig, "RFC 7258 – pervasive monitoring is an attack," May 2014. URL: https://tools.ietf.org/html/rfc7258

[4] A. Johnson, "The Tor path simulator." URL: https://github.com/torps/torps

[5] A. Johnson, C. Wacek, R. Jansen, M. Sherr, and P. Syverson, "Users get routed: Traffic correlation on Tor by realistic adversaries," in *CCS*. ACM, 2013. URL: https://nymity.ch/tor-dns/pdf/Johnson2013a.pdf

[6] "Mixmaster." URL: http://mixmaster.sourceforge.net

[7] S. J. Murdoch and P. Zieliński, "Sampled traffic analysis by Internet-exchange-level adversaries," in *PET*. Springer, 2007. URL: https://nymity.ch/tor-dns/pdf/Murdoch2007a.pdf

†This paper will appear in NDSS 2017.

*All four authors contributed substantially and share first authorship.

[1] The acronym is short for DNS-enhanced fingerprinting and egress correlation on Tor.

[2] Our project page is available at https://nymity.ch/tor-dns/.