Demo: Traffic Splitting for Tor — A Defense against Fingerprinting Attacks

Sebastian Reuter[1], Jens Hiller[1], Jan Pennekamp[1], Andriy Panchenko[2], and Klaus Wehrle[1]

[1] *Communication and Distributed Systems*, RWTH Aachen University, Germany
[2] *IT Security*, Brandenburg University of Technology, Germany
{lastname}@comsys.rwth-aachen.de, {firstname.lastname}@b-tu.de

Abstract: Website fingerprinting (WFP) attacks on the anonymity network Tor have become ever more effective. Furthermore, research discovered that proposed defenses are insufficient or cause high overhead. In previous work, we presented a new WFP defense for Tor that incorporates multipath transmissions to repel malicious Tor nodes from conducting WFP attacks. In this demo, we showcase the operation of our traffic splitting defense by visually illustrating the underlying Tor multipath transmission using LED-equipped Raspberry Pis.

Keywords: Onion Routing; Website Fingerprinting; Multipath Traffic; Privacy

1 Introduction

Tor is one of the most popular anonymization networks [CMH+20]. It is commonly used to bypass censorship, to maintain the freedom of speech, or to repel trackers on the Internet and in the IoT domain [HPD+19, PHR+19]. To provide anonymity to its users, Tor leverages the onion routing principle [CMH+20]: All traffic is dismembered into fixed-size *cells* and routed from the client (onion proxy, OP) to the server through a virtual *circuit* consisting of multiple relaying onion routers (ORs), known as *entry*, *middle*, and *exit* ORs (depending on their position in the circuit). During transmission, Tor applies a multi-layered encryption scheme to all cells, which obfuscates their content and the communication endpoints. More specifically, Tor achieves anonymity as each OR knows only its predecessor and successor in the circuit and, consequently, no OR knows the identity of both the client and the server at the same time. Hence, Tor hides the client's identity (IP address) from all involved entities except nodes on the path between the OP and the entry node, which, however, are unaware of the accessed server [CMH+20].

Website Fingerprinting (WFP) WFP attacks enable attackers to subvert Tor's anonymity guarantees [CMH+20, PLZ+16]. To this end, they exploit the fact that Tor's multi-layered encryption does *not* properly obfuscate revealing timing and volume *patterns* within the transmitted data stream. A *local passive adversary* located at the first hop of a Tor circuit (i.e., a malicious entry OR or its ISP) can exploit this knowledge to unveil the website retrieved over a Tor circuit. Specifically, the attacker employs machine learning to identify traffic patterns for the loading of websites over Tor circuits. As its privileged position also provides the attacker with the IP address of the circuit's Tor user, it can thus link the Tor user to the visited website. To counter such

Authors' version of a manuscript that was accepted for publication in *Electronic Communications of the EASST* (Volume 080) as part of the Conference on Networked Systems (NetSys 2021). Changes may have been made to this work since it was submitted for publication. Please cite the published version: http://doi.org/10.14279/tuj.eceasst.80.1151 – ISSN: 1863-2122

WFP attacks, several defenses have been proposed [CNJ14, WG17, JIP+16], which, however, usually introduce undesirable bandwidth or latency overheads [CMH+20].

Our Contributions In previous work [CMH+20, CMP+19, PHR+19], we presented a novel defense against WFP attacks based on a multipath approach that distributes Tor cells over multiple entry ORs, and showed that a well-tailored distribution strategy achieves a superior defense against malicious entry nodes while incurring only limited overhead. In this demo, we visualize the traffic distribution of our defense to provide Tor users with an intuition on the operation of the defense and thus foster its acceptance.

2 Related Work

The discovery of WFP attacks started an arms race between corresponding defenses and improved versions of the attack. We briefly discuss previous defenses and refer to [CMH+20] for a comprehensive overview. Walkie-Talkie [WG17] combines half-duplex data transmission with traffic shaping to distort distinguishable traffic patterns. CS-BuFLO [CNJ14] and WTF-PAD [JIP+16] rely on transmitting padding data to warp the traffic patterns. Yet, research has found that existing defenses either introduce huge overhead or are not sufficiently effective, even though WTF-PAD is currently considered for inclusion into Tor [CMH+20].

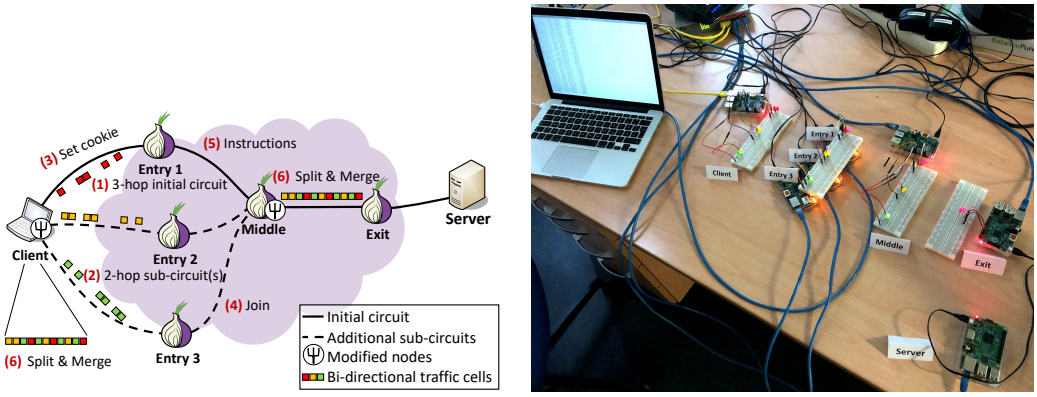
Related to traffic splitting, research considered implementing multipath approaches in Tor to improve its performance, e.g., Conflux [ABEG13], but not to counter WFP attacks [PHR+19].

3 Design of the Traffic Splitting Defense

Our WFP defense splits traffic between the Tor client and middle OR to distribute it over multiple entry ORs (cf. Figure 1a). Thereby, a traffic splitting *strategy* distorts traffic patterns with respect to malicious entry nodes [CMH+20]. With a suitable strategy, this traffic splitting is highly effective in repelling WFP attacks ($< 14\%$ accuracy compared to more than 98% without defense) while introducing hardly noticeable overhead [CMH+20].

Multipath Circuit Setup: To set up an m-fold multipath circuit, our design re-uses Tor's existing circuit mechanisms [CMH+20]: First, the client builds a common three-hop *initial* circuit and then establishes $m-1$ further two-hop *sub-circuits* via different entry ORs to the middle OR. To enable the middle OR to join the created sub-circuits, the client sends a cookie value (20-byte nonce) to the middle OR using the initial circuit and waits for an acknowledgement (`SET_COOKIE`, `COOKIE_SET`). Following, the client sends the same cookie via each remaining sub-circuit to the middle OR (`JOIN`). For each sub-circuit, the middle OR uses the received cookie to join the sub-circuit with the initial circuit and sends an acknowledgement (`JOINED`).

Multipath Traffic Transmission: A distribution strategy determines the order in which cells are sent over the sub-circuits such that the middle OR can send cells to and merge cells received from the client in the correct order [CMH+20]. To set up this strategy, the client regularly sends *splitting instructions* to the middle OR (`INSTRUCTION`, `INFO`). The proper selection of this strategy is essential to prevent any fingerprintable traffic patterns and thus influences the effectiveness of the defense [CMH+20].



(a) Schematic illustration.

(b) Demo setup.

Figure 1: Overview of our Tor traffic splitting WFP defense.

Implementation: We implemented our defense and added demo functionality in Tor 0.3.5.11, the latest version with long-term support. Our source code is publicly available [cod21].

4 Illustrating the Traffic Splitting Operation (Demo Setup)

To illustrate the operation of our WFP defense, we use seven Raspberry Pis (RPi) (cf. Figure 1b). These RPis are interconnected among each other using Ethernet (**local setup**) as well as the public Internet (use of **real Tor network**). In the **local setup**, we configure all RPis to form a separate local Tor network and build multipath circuits using one RPi as client, three RPis as entry ORs, and the remaining three RPis as middle OR, exit OR, and server, respectively. During the demo, the client repeatedly builds new multipath circuits to the exit OR and requests a website from the server or transfers large amounts of data using the bandwidth measurement tool *iPerf*. Thereby, the circuit setup (building of sub-circuits as well as the joining operation) is visualized with live log output from the respective RPis. More importantly, to visualize the following multipath traffic, each RPi is equipped with LEDs which represent transmissions on each sub-circuit with a different color. Furthermore, to also differentiate the transmission direction, we use pairs of LEDs (same color), one LED for each direction. Thus, the client and middle RPis each have three pairs of LEDs (one for each sub-circuit), whereas the other RPis are equipped with one pair of LEDs. During the data transmission, each RPi uses the LED corresponding to the transmission direction (and circuit, if applicable) to signal the transmission of Tor traffic cells, thereby evidently illustrating the transmission behavior of our multipath defense. For improved visibility in the face of high networking speeds, the RPis are configured to signal only every 100^{th} traffic cell per sub-circuit and transmission direction in case of the iPerf (bulk) data transmission. In the case of the website request, due to the comparably small amount of data, we maintain the observability by artificially restricting the bandwidth of the contacted server RPi. The demo can alter the number of used paths and show different splitting strategies, i.e., from easily conceivable strategies like round-robin up to the best defense strategy determined in [CMH+20].

Beyond this local setup, we also showcase the application of our traffic splitting approach within the **real Tor network**. To this end, we replace the local ORs (RPis) with nodes de-

ployed in the live Tor consensus and only use an RPi to act as the traffic-splitting enabled client. Specifically, we use one of our own splitting-enabled public Tor nodes as remote middle OR for multipath circuits, which allows us to observe the splitting and merging activities via the live log output. This setup allows us to showcase the use of our defense while accessing virtually any Internet resource reachable via the real Tor network.

5 Conclusion

In this demo, we visualize and observe the traffic distribution of our WFP defense using multiple LED-equipped RPis within both a local autonomous setup and the live public Tor network. Thereby, we provide Tor users with an intuition on the operation of the defense, thus fostering its acceptance and raise further awareness towards this active area of research.

Bibliography

[ABEG13] M. AlSabah, K. Bauer, T. Elahi, I. Goldberg. The Path Less Travelled: Overcoming Tor's Bottlenecks with Traffic Splitting. In *PETS*. 2013.

[CMH+20] W. De la Cadena, A. Mitseva, J. Hiller, J. Pennekamp, S. Reuter, J. Filter, T. Engel, K. Wehrle, A. Panchenko. Traffic Sliver: Fighting Website Fingerprinting Attacks with Traffic Splitting. In *ACM CCS*. 2020.

[CMP+19] W. De la Cadena, A. Mitseva, J. Pennekamp, J. Hiller, F. Lanze, T. Engel, K. Wehrle, A. Panchenko. POSTER: Traffic Splitting to Counter Website Fingerprinting. In *ACM CCS*. 2019.

[CNJ14] X. Cai, R. Nithyanand, R. Johnson. CS-BuFLO: A Congestion Sensitive Website Fingerprinting Defense. In *ACM WPES*. 2014.

[cod21] https://github.com/TrafficSliver/trafficsliver-net-demo, 2021.

[HPD+19] J. Hiller, J. Pennekamp, M. Dahlmanns, M. Henze, A. Panchenko, K. Wehrle. Tailoring Onion Routing to the Internet of Things: Security and Privacy in Untrusted Environments. In *IEEE ICNP*. 2019.

[JIP+16] M. Juarez, M. Imani, M. Perry, C. Diaz, M. Wright. Toward an Efficient Website Fingerprinting Defense. In *ESORICS*. 2016.

[PHR+19] J. Pennekamp, J. Hiller, S. Reuter, W. De la Cadena, A. Mitseva, M. Henze, T. Engel, K. Wehrle, A. Panchenko. Multipathing Traffic to Reduce Entry Node Exposure in Onion Routing. In *IEEE ICNP*. 2019.

[PLZ+16] A. Panchenko, F. Lanze, A. Zinnen, M. Henze, J. Pennekamp, K. Wehrle, T. Engel. Website Fingerprinting at Internet Scale. In *NDSS*. 2016.

[WG17] T. Wang, I. Goldberg. Walkie-Talkie: An Efficient Defense Against Passive Website Fingerprinting Attacks. In *USENIX Security*. 2017.