DE GRUYTER OPEN

Proceedings on Privacy Enhancing Technologies ; 2017 (1):100–117

Chen Chen* and Adrian Perrig

PHI: Path-Hidden Lightweight Anonymity Protocol at Network Layer

Abstract: We identify two vulnerabilities for existing high-speed network-layer anonymity protocols, such as LAP and Dovetail. First, the header formats of LAP and Dovetail leak path information, reducing the anonymity-set size when an adversary launches topological attacks. Second, ASes can launch session hijacking attacks to deanonymize destinations. HORNET addresses these problems but incurs additional bandwidth overhead and latency.

In this paper, we propose PHI, a Path-HIdden lightweight anonymity protocol that solves both challenges while maintaining the same level of efficiency as LAP and Dovetail. We present an efficient packet header format that hides path information and a new back-off setup method that is compatible with current and future network architectures. Our experiments demonstrate that PHI expands anonymity sets of LAP and Dovetail by over 30x and reaches 120 Gbps forwarding speed on a commodity software router.

Keywords: Anonymity, path-hidden protocols

DOI 10.1515/popets-2017-0007

Received 2016-05-31; revised 2016-09-01; accepted 2016-09-02.

1 Introduction

Revelations about governments' mass-surveillance programs have demonstrated their capability of conducting pervasive surveillance on huge volumes of domestic and international traffic [4, 9]. Meanwhile, an increasing number of users have begun using anonymous communication software to protect their privacy. For instance, Tor [27] has on average 2 million active users per day [13]. However, most anonymity software today is built as an overlay network composed of end hosts' voluntarily-contributed nodes [6, 8, 14]. As a consequence, users experience poor performance due to long propagation delays and limited bandwidth, along with intrinsic queuing and retransmission delays of the protocols [28].

Recent research has proposed anonymity as a principal network function to benefit from short paths and high throughput of network devices [21, 34, 46]. These proposals demonstrate that it is viable to build lightweight cryptography into network routers to help anonymize the huge volumes of traffic accessible to mass surveillance programs today. We compare these protocols in Table 1.

Protocol	Throughput	Bandwidth Overhead	Path Info. Leakage	Session Hijacking	Trust in 1st hop	Hosts Control Paths
LAP	Low	Low	Yes	Yes	Yes	No
Dovetail	Low	Low	Yes	Yes	No	Yes
HORNET	Medium	Medium	No	No	No	Yes
PHI	Low	Low	No	No	No	No

Table 1. Comparison of high-speed network-layer anonymity protocols.

LAP [34], for example, adds an encrypted path into each packet, and LAP routers forward packets using only symmetric cryptography. However, a compromised first-hop Autonomous System (AS) can deanonymize both the source and the destination and thus immediately compromise anonymity because LAP's setup process leaks the destination address. Dovetail [46] overcomes the limitation by using an indirection node, called "match maker", to conceal the destination node from the first-hop AS. However, Dovetail requires that the source have full control over the traversed path, which harms its compatibility with current network architectures.

Furthermore, we focus on two vulnerabilities of LAP and Dovetail. First, their headers, even with the proposed defense to hide a path's length and routers' positions on a path [34], still leak such path information. Thus, an on-path router can reduce the size of the source's anonymity set based on publicly available network topology. Second, payload encryption is detached from the path in Dovetail, enabling a session hijacking attack to deanonymize destinations.

In contrast to LAP and Dovetail, HORNET [21] hides the path information by using an onion-encrypted data structure to embed path information and prevents the session hijacking attack. However, HORNET's solution incurs additional costs: first, HORNET's connection setup requires Elliptic-Curve Diffie-Hellman (ECDH) computation between a sender and each intermediate on-path nodes, adding computational la-

***Corresponding Author: Chen Chen:** Carnegie Mellon University / ETH Zurich, E-mail: chen.chen@inf.ethz.ch

Adrian Perrig: ETH Zurich, E-mail: adrian.perrig@inf.ethz.ch

tency; second, HORNET requires the sender to anonymously retrieve and verify the public keys of on-path nodes, introducing further latency and potential identity leakage vectors.

At a first glance, one is bound to an unfortunate choice between weaker anonymity and additional latency. In this paper, we demonstrate that it is possible to achieve the best of both worlds. We propose a Path-HIdden lightweight anonymity protocol, named PHI, that improves anonymity over LAP and Dovetail and is equally efficient.

PHI improves on LAP and Dovetail by introducing three new techniques. First, PHI places nodes' state in a pseudo-random order in a packet header to conceal information about node positions. Second, PHI leverages a back-off path construction method to eliminate the need for a source to fully control the path traversed. Third, PHI prevents session hijacking attacks by binding payload encryption to paths.

Our paper makes the following contributions:

1. We identify two attacks that reduce sizes of anonymity sets in LAP and Dovetail. In particular, we model and analyze the path information leakage when LAP and Dovetail intentionally obscure path information by using *variable-size segments* (see Section 2). In comparison, existing work, HORNET, only shows that LAP and Dovetail leak path information without such protection mechanism.
2. We propose a path-hidden header format that is more efficient than the ones used by onion routing protocols.
3. We present a new approach to establish an end-to-end path for a source node with no control over the path traversed.
4. We design the Path-HIdden lightweight anonymity protocol (PHI), an efficient network layer protocol that provides stronger anonymity properties than LAP and Dovetail with the same level of efficiency.
5. We evaluate PHI's security and performance. Evaluation results confirm that PHI's performance is comparable to, or more efficient than LAP and Dovetail, while expanding the sizes of anonymity sets.

2 Background

2.1 Network-layer Anonymity Protocols

Network-layer anonymity protocols assume that network infrastructure (e.g., switches and routers) perform anonymization operations when forwarding packets. They function at the network layer as a complementary or as an alternative option to the Internet Protocol (IP) to anonymize packets' sources and destinations. Compared to existing anonymity systems built on overlay networks such as Tor [27], network-layer anonymity protocols aim to offer low-latency and high-throughput packet forwarding and scale to handle high volumes of traffic seen in the Internet [21, 34, 46].

To achieve fast forwarding and high scalability on network devices whose per-packet computation and per-flow storage are usually extremely limited, network-layer anonymity protocols share two design choices: 1) that packet forwarding only needs symmetric crytography, and 2) that forwarding state should be carried by packets instead of stored on network devices. While using symmetric cryptography is also common in latest overlay-based anonymity systems, the second choice mainly distinguishes network-layer anonymity protocols. At the beginning of each connection, a source node and a destination node exchange setup packets that traverse each Autonomous System (AS) on the path. Within a setup packet, each on-path AS (or a *node*) creates a *path segment* containing its forwarding state. These path segments are carried by data packet headers so that a node can retrieve its state and know how to forward the packets.

Lightweight vs. onion-routing protocols. We divide existing network-layer anonymity protocols into onion-routing protocols and lightweight protocols based on their different requirements for computation and header overhead. An onion-routing protocol, like HORNET [21], requires an on-path network device to compute an expensive asymmetric crytographic operation for setting up a flow and applies per-hop authenticated encryption over every data packet. Each on-path node also needs to store necessary keys within its path segment, resulting in large packet headers. In comparison, lightweight anonymity protocols, such as LAP [34] and Dovetail [46], only require a network device to decrypt and verify a path segment that contains minimal information for forwarding packets, and use end-to-end packet encryption to offer confidentiality.

Topology-based attacks. Unlike a node in an overlay-based anonymity system that can forward packets to any other node, an AS can only forward packets according to its physical connections and business contracts with its neighbors. Therefore, network-layer anonymity protocols are inherently subject to so called "topology-based attacks". With publicly-available network topology information, a compromised node receiving a single packet from a victim can narrow down the victim's anonymity set. For example, in Figure 1, if we assume that AS2 is AS3's customer, by knowing the topology and receiving a packet from AS2, AS3 can conclude that the source node must be within {AS0, AS1, AS2, AS4, AS5}, which also forms the anonymity set. If the adversarial AS further discovers its AS-level distance from the source's AS, it can reduce the size of the victim's anonymity set. In Figure 1, if AS2 is AS3's customer and AS3 uncovers that the source of a packet from AS2 is 3 hops away, AS3 can infer that the source's