Compromising Tor Anonymity Exploiting P2P Information Leakage

Pere Manils, Abdelberi Chaabane, Stevens Le Blond,
Mohamed Ali Kaafar, Claude Castelluccia, Arnaud Legout, Walid Dabbous

ABSTRACT

Privacy of users in P2P networks goes far beyond their current usage and is a fundamental requirement to the adoption of P2P protocols for legal usage. In a climate of cold war between these users and anti-piracy groups, more and more users are moving to anonymizing networks in an attempt to hide their identity. However, when not designed to protect users information, a P2P protocol would leak information that may compromise the identity of its users. In this paper, we first present three attacks targeting BitTorrent users on top of Tor that reveal their real IP addresses. In a second step, we analyze the Tor usage by BitTorrent users and compare it to its usage outside of Tor. Finally, we depict the risks induced by this de-anonymization and show that users' privacy violation goes beyond BitTorrent traffic and contaminates other protocols such as HTTP.

Keywords

Anonymizing Networks, Privacy, Tor, BitTorrent

1. INTRODUCTION

The Tor network was designed to provide freedom of speech by guaranteeing anonymous communications. Whereas the cryptographic foundations of Tor, based on onion-routing [3, 9, 22, 24], are known to be robust, identity leaks at the application level can be exploited by adversaries to reveal Tor users identity. Indeed, Tor does not cipher data streams end-to-end, but from the source to a Tor exit node. Then, streams from the Tor exit node to the destinations are in plain text (if the application layer does not encrypt the data). Therefore, it is possible to analyze the data stream seeking for identity leaks at the application level. Tor does not consider protocol normalization, that is, the removal of any identity leak at the application level, as one of its design goals. Whereas this assumption is fair, Tor focuses on anonymizing the network layer, it makes the task of users that want to anonymize their communications much harder. As an illustration, the Web communications on Tor are the subject of many documented attacks. For instance, attacks can leverage from misbehaving browsers to third party plugins or web components (JavaScript, Flash, CCS, cookies, etc.) present in the victim's browser to reveal browser's history, location information, and other sensitive data [7, 2, 4, 17].

In order to prevent or at least reduce these attacks, the Tor project recommends the use of web proxy solutions like Privoxy or Polipo [19, 5, 21]. The Tor project is even maintaining a Firefox plugin (Torbutton [20]) that, by disabling potentially vulnerable browser components, aims to countermeasure most of the well-known techniques an adversary can exploit to compromise identity information of Tor users. Thus a big effort has been invested and is heading on improvement and protection of the HTTP protocol on top of Tor, but surprisingly, such an effort is limited to this protocol.

P2P applications and more specifically BitTorrent, an application that is being daily used by millions of users [12], have been so far neglected and excluded from anonymizing studies. The crux of the problem is that BitTorrent easily allows any adversary to retrieve users' IP addresses from the tracker for torrents they are participating to. Indeed, by design BitTorrent exposes the IP address of peers connected to torrents. This implies important anonymity and privacy issues, as it is possible to know who is downloading what. To go around this issue, many BitTorrent users that care about their anonymity use Tor, although the Tor project explicitly not recommend the use of BitTorrent on top of the Tor network, because of the major risk of overloading the network.

BitTorrent is a complex protocol with many potential identity leaks, as user privacy is not among its design goals. However, this serious issue is overlooked by BitTorrent users who believe that they can hide their identify when using Tor.

Today's reality is that BitTorrent is one of the most used protocols on top of Tor (with HTTP/HTTPS) in terms of traffic size and number of connections as reported by [16] and observed during our own experiments. Surprisingly, no studies have been conducted on the way BitTorrent may leak the identity of users when the application is running over an anonymizing network. Although, it might be argued that BitTorrent is mostly used for piracy (distribution of illegal content), we believe that privacy issue is a major impediment for the commercial and legal use of BitTorrent. Moreover, identity leaks at the level of a stream might also contaminate