

Abstract

The transmission of anonymous information on the Internet is a research hotspot in the field of data privacy protection. On the basis of studying traditional information security (the confidentiality, integrity and authenticity of transmitted data), scholars are paying more and more attention to how to protect the identity information of communication users. That is, the anonymity of transmission. In response to the new needs of network communication for information security, this article conducts in-depth research and analysis of the most successful anonymous communication system Tor, and analyzes the Tor network protocol, the status quo and the challenges it faces; and based on this Improved, designed an anonymous information transmission method based on the dual-cloud structure in response to the problems of Tor, and conducted performance analysis and testing of the system.

SYSTEM STRUCTURE

1. Tor, an anonymous communication system

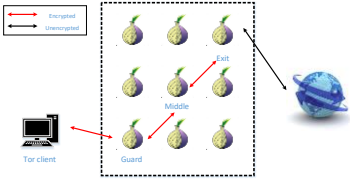


Fig. 1 Tor network structure

Tor can be used to prevent widespread traffic filtering and sniffing analysis on the Internet, and can realize anonymous external connections and anonymous hidden services. Referring to Fig.1, the Tor system consists of four parts[10]: Onion Router OR (Onion Router), Client Onion Proxy OP (Onion Proxy), Directory Server (Directory Server), and Application Server (App Server).

- (1) Onion router: OR is Tor relay node, without any special authority, each OR maintains TLS connection with each other OR. Tor generally defaults to three ORs to form a link, which are the entry node (Entry), the intermediate node (Middle) and the exit node (Exit) in Fig.1.
- (2) Client onion proxy: OP is a program that runs on the user's PC. OP is used by the user to obtain the node directory to facilitate the establishment of the link, and the OP encapsulates the user data into cells and multi-layer encryption, which is based on TCP. The application provides anonymous proxy services.
- (3) Directory server: It is a very critical server in Tor. Its main function is to provide OP with OR lists and related parameters that can be used to ensure that anonymous links can be effectively established.
- (4) Application server: it is the destination server visited by OP.

2. Anonymity system assumption

The analysis of the entire anonymous system is based on the following settings:

This system does not consider the inherent shortcomings of the anonymous system, that is, it is assumed that the association cannot be obtained only through anonymous data packets;

The internal communication data package of the cloud platform cannot be obtained. Under this setting, if the adversary does not obtain control of the internal VM of the cloud platform, theoretically complete anonymity can be achieved;

Data traffic in and out of the cloud platform can be hidden by other traffic on the cloud platform that does not participate in anonymous network communications. The more virtual nodes provided by the cloud platform, the better the anonymity.

3. Anonymous network based on dual clouds



Fig. 2 Anonymous network structure based on dual clouds

The Fig.2 shows two public clouds, and a Tor network is built on them. When the system client accesses the first cloud, it uses SSL-encrypted remote login, which is encrypted video streaming media data. The communication between the cloud platform and the exit is encrypted using SSL/TLS, and ordinary users of the cloud platform will also have such requests. The attacker cannot detect the sending and receiving of messages, which is unobservable.

When we borrow the cloud platform Tor network resources, there are also ordinary normal users who are borrowing the cloud platform. Therefore, it is more difficult for an attacker to judge me because the network data is mixed with normal users. Assuming that the attacker has distinguished the ordinary nodes on the cloud platform, he must consider that if the judgment is wrong, it may affect the regular services of ordinary users on the cloud platform.

Moreover, as an ordinary adversary, it is more difficult to obtain internal data on the cloud platform, because as a cloud storage service provider, it will give users basic protection, that is, internal data is characterized as invisible to the outside world, although the Tor built on the cloud platform The information exchange between the internal nodes of the network is also invisible to the outside world.

Consider a specific scenario Fig.3. Alice uses an anonymous network based on a dual cloud structure to access the Twitter website through HTTPS:

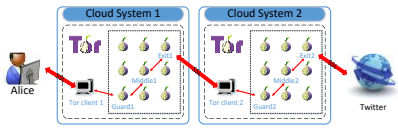


Fig. 3 Alice uses the anonymous network to access the Twitter website through the HTTPS protocol

- (1) Through remote login, Alice accesses Tor client 1;
- (2) Specify the transit node Tor client 2 on Tor client 1, and send network requests from Exit 1 to Tor client 2 through the Tor network of Cloud system 1;
- (3) Tor client 2 will simply perform network request forwarding, and send network requests to Twitter through the Tor network on Cloud system 2;
- (4) After the Twitter website responds to the network request, it will reply to the message, and the reply network information will be transmitted to the Tor client 2 through the Tor network on Cloud system 2;
- (5) Information forwarding will be sent by Tor client 2 from Exit 1 to Tor client 1;
- (6) Finally, the access result obtained by Tor client 1 is fed back to Alice in the encrypted video stream.

The above communication between cloud platforms appears to the attacker to be the normal service of ordinary cloud storage service providers, which guarantees the unobservability of communication messages.

PRACTICE TEST

The specific practical operating environment is: the server to build the cloud platform is CentOS Linux release 7.8.2003 (Core), the cloud platform is built using CloudStack 4.14.0.0, the cloud platform node operating system is Ubuntu 18.04.5 LTS, and the Tor network version is tor-0.4.3.6, data tool acquisition uses python to call pycurl library for data feedback.

In order to ensure the accuracy of the experimental data, under the condition that the network environment remains unchanged, the access operations were performed 100 times and the average value was taken. And in the case of four specific tool networks, all are HTTPS access.



Fig. 4 Histogram of network response data

The data in Fig.4 is divided into four aspects: the time to establish a connection, the time to prepare for transmission, the time to transmit the first byte, the time to complete, and to show the difference in website access. The private Tor network built on the cloud platform is faster than the nodes on the physical machine. It is obvious that the network response speed of the anonymous system with dual cloud structure is almost close to that of the private Tor network built on the physical machine, and the network response speed of the public Tor network is so slow that it is not indicative.

SUMMARY AND OUTLOOK

This article mainly proposes and designs an anonymous information transmission system based on dual clouds, and analyzes its anonymity and actual network access speed. Next, we will combine the existing De-Anonymization attack methods to test the actual security of the system.

Reference

- [1] Karunanayake, Ishan, et al. "Anonymity with Tor: A Survey on Tor Attacks." arXiv preprint arXiv:2009.13018 (2020).
- [2] Basyoni, Lamiaa, et al. "Traffic Analysis Attacks on Tor: A Survey." 2020 IEEE International Conference on Informatics, IoT, and Enabling Technologies (ICIoT). IEEE, 2020.
- [3] L Wang, H Mei and V S Sheng. Multilevel Identification and Classification Analysis of Tor on Mobile and PC Platforms[J]. IEEE Transactions on Industrial Informatics, 2020, doi : 10.1109/TII.2020.2988870.
- [4] Leberknight, Christopher S., et al. "A taxonomy of Internet censorship and anti-censorship." Fifth International Conference on Fun with Algorithms. 2010.

Contact information

Contact person: Fan Wen
 Phone number: 18852866567
 Mailbox: wenfan@ujs.edu.cn