Detecting TOR Traffic In A Network

**Index**

1. Abstract

2. Introduction

3. Literature Review

4. Features

5. System Design(Methodology)

6. Conclusion

7. References

**Abstract**

To date, there has yet to be a study that characterizes the usage of a real deployed anonymity service. We present observations and analysis obtained by participating in the Tor network. Our primary goals are to better understand Tor as it is deployed and through this understanding, propose improvements. In particular, we are interested in answering the following questions:

(1) How is Tor being used?

(2) How is Tor being misused?

(3) Who is using Tor?

To sample the results, we show that web traffic makes up the majority of the connections and bandwidth, but non-interactive protocols consume a disproportionately large amount of bandwidth when compared to interactive protocols. We provide a survey of how Tor is being misused, both by clients and by Tor router operators. In particular, we develop a method for detecting exit router logging (in certain cases). Finally, we present evidence that Tor is used throughout the world, but router participation is limited to only a few countries been extensively used to design, develop and integrate all system components.

**Introduction**

Background of the Problem

With the rapid increase of the use of the Internet, people consider more about their important and private information that are transmitted on the Internet when people use the Internet for work and daily life activities such as shopping online, managing bank accounts online, visiting private website without being known by others, chatting privately with somebody using instant messaging application, and so on.

My project is about the anonymous communication network. Anonymous communication has been used on the Internet for several years. There are a lot of advantages using it, such as improving Internet users’ privacy and security, keeping websites from recording and tracking visitors’ behavior, and breaking the censor block from the ISP.

With above advantages, there should be a mass of users using this excellent Internet tool. But the fact is that the users of anonymous communication tool are very few compared to the huge quantity of the Internet users because of the following two reasons. First, it is very hard to understand the principle and structure of anonymous communication for existing and potential users because there are very little existing material and information about it. Second, the performance of anonymous communication tool is not so good that users are pleased to use it. The speed of Internet browser through anonymous communication tool is slow compared to the one without it. These two factors impact the wide use of anonymous network.

## Statement of the problem

Through researching the principle and structure of the anonymous communication network, we know that anonymous communication network is a very secure and stable communication tool. But why is the Internet visit speed through it very slow?

In my project, I will test and analyze one of the most popular anonymous network tools, Tor. I will set up an experiment system to test the existing Tor anonymous network and present the reason why Tor suffers low network throughput performance. After finding out the reason of its low performance, I will visualize Tor anonymous communication network in order to promote the use of it. I will set up a database providing comprehensive information of Tor network and publish a website which links to this Tor network database and retrieves the information of Tor network nodes. The website and database will help Tor users get to know the principle, structure, and information about Tor network and improve the use situation of Tor network.

**Literature Review**

**Anonymous communication network review**

With more and more communication on the Internet, there is much higher request for private and secure Internet communication. When we surf the Internet or chat with others using instant messaging application, we consider more private and secure problems and want to protect our privacy better.

Anonymity is righteous and necessary in many scenarios, such as protecting Internet user privacy, improving system security, bypassing Internet censorship, satisfying some antivirus requirement, and protecting Internet users’ computer from hackers’ attacks. With the wider use of the Internet and more improving computer hacker technologies, the Internet and network users are anxious to gain more powerful ability to keep their privacy and security. They want to request a better network tool or technology to protect their private information from being monitored by the Internet sniffers and hackers. With this requirement, anonymous communication has become more and more popular on the Internet.

As one classic anonymous communication tool, Tor gets more implementation on the world. There are many companies paying much attention to network communication security and using Tor network now. Indymedia Group recommends that their employees use Tor when communicating with others on the Internet. Recently some departments of US Navy used Tor communication network for the activities in Middle East area. 1

At the same time, besides the positive side of Tor communication network, there are some reports about the weakness of Tor too. For example, in September 2007, Dan Egerstad, a [Swedish](http://en.wikipedia.org/wiki/Sweden) security specialist, found that he can capture a lot of sensitive information such as usernames and passwords for emails account though monitoring the data packets transmitted from Tor exit nodes.2 With the design principle of Tor network, there is no encryption like SSL technology between exit nodes and the final destinations. Although this attack does not violate the anonymity for the source and destination of Tor network nodes, it is a big weakness of Tor network.

1 <http://www.torproject.org/overview.html.en>

2 <http://en.wikipedia.org/wiki/Tor_(anonymity_network)>

**Features**

* Design a system that can identify the tor traffic in network.
* Sends alert who is trying to access.
* With list of tor exit nodes,we can tell our system to set up a policy that will trigger if any of these ips are identified as a source/destination of network.
* Detecting tor traffic and send giving updater related to that

# **System Design (Research methodology)**

**Tor Network**

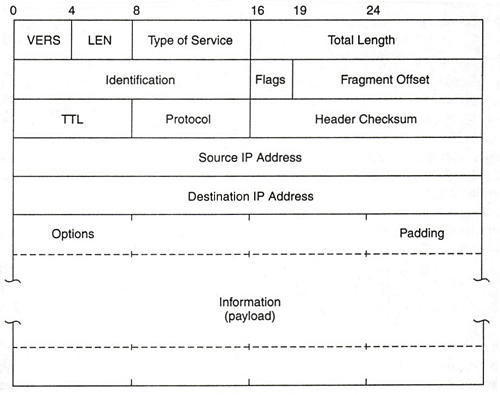
**Tor network introduction**

Tor is the second generation onion router system with many advantages compared to the previous generation.

Tor is one of anonymous communication tools. It is a virtual tunnel for people on the Internet to communicate privately and securely. Over the public networks such as the Internet, Tor can keep the communication between people safe. Tor can be used with many computer applications including several web browsers, instant messaging, and so on. It can improve the privacy and security when people use these applications. There are several applications of Tor described as follows:

Tor users can surf websites without someone tracking or spying the content and information they visit. Tor user can use instant messaging application without any leak for the conversation contents and objects. Tor users also can penetrate the blocked website by Internet service provider and do not leave any identity information to that Internet service provider. When using Tor to publish a website, Tor users can get a hidden service that publishes the website but does not let visitors know the location information of this website. Tor users can avoid some harassment when they chat in some chat rooms or in some chat forums. Journalists or staff of some organizations can use Tor to finish their work abroad without letting the foreign spies know their work contents and their employers.

“Traffic analysis” is a common surveillance on the Internet. It can disclose the address and location information of communication parties. There are two parts in one Internet data packet. One is the data payload part and the other is packet header for Internet routing. Data payload is the content that we want to transmit to the recipient. E.g. a web page content, one transmitting file, one email message, and so on. Data payload part is encrypted commonly. But the routing header of the Internet packet is not encrypted. It records the source and destination information of the packet such as the IP addresses of sender and recipient. Traffic analysis use the weakness of the header part of Internet packet to get the source and destination information of the packet. Figure 1 is the structure of IP packet.

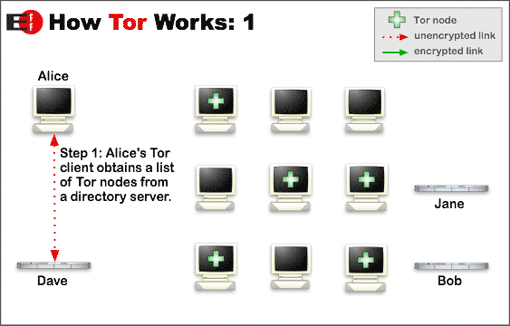


**Figure 1 IP packet structure1**

The header part reveals the packet’s source, destination, size, timing, and so on. The recipient of the packet can know much information through checking the header. If some intermediaries sit between the source and the destination, it also can get above information easily through looking at the header.

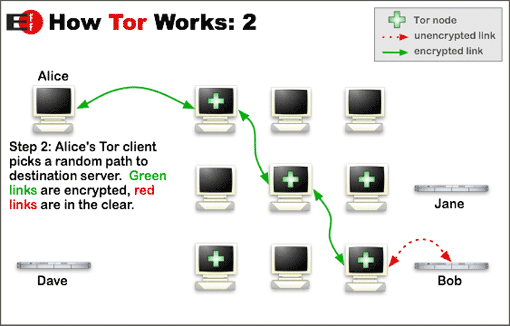
With the disclosure of source and destination information, the Internet users may suffer some loss, for example some e-commerce website may give different prices according to buyers’ IP addresses. Another instance is that eavesdropper can know your nationality through traffic analysis for IP packet header and apply some terrorism attacks aiming to some countries’ users.

How does Tor anonymous communication network overcome the traffic analysis to hide the address information? Tor network distribute the Internet data packets through a sequence of nodes from the packets’ source to destination. These Tor network nodes are independent and can not know the whole path of the data packets. So no one point can link the source and destination. It is just like using a twisty route to get rid of any eavesdropper who is tailing your data packets. At the same time, to confirm the privacy and security, Tor users select random nodes from Tor network system to firm the pathway from source to destination. After some time, Tor users change selected nodes to set up a new pathway to keep the pathway from monitoring. We will discuss the setup of Tor pathway as follows.



**Figure 2 Tor router setup step 1 1**

There are three kinds of Tor nodes on Tor anonymous communication network. One is server node which not only can visit their own destination websites but also can relay other Tor nodes’ data packets. Server can initiate a Tor router and can be the mediate relay point for other Tor nodes’ router. Another kind of Tor node is client node which only begins a Tor router and can not relay other nodes’ data packets. The third one is directory server of Tor network which stores the information of all Tor server nodes. As the presentation of above figure, Tor client, Alice wants to visit the website Bob. Alice is a Tor client node only. Between Alice and Bob, there are many Tor server nodes. First, Tor client node, Alice needs to get enough quantity of Tor server nodes’ information from a directory server node, Dave here. When the client node gets enough server nodes’ information, it can begin to set up Tor routers or circuits from itself to destinations.

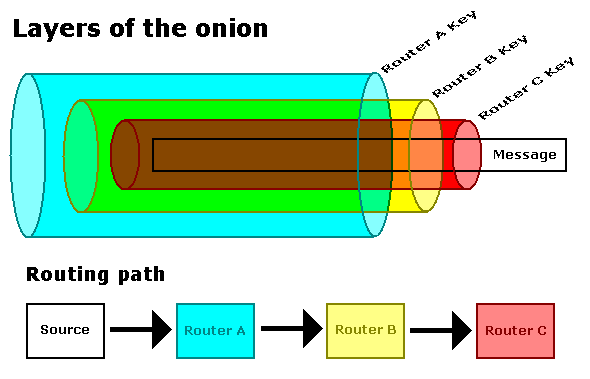


**Figure 3 Tor router setup step 21**

After the client node gets enough quantity of server nodes, the client node builds a Tor circuit with 3 other server nodes incrementally. This circuit is set up step by step with TLS encryption and private key. The client node extends the circuit one hop at a time. Each server node along the circuit only knows the information of nodes which give date to it, and to which it gives data. No any server node knows all the nodes’ information through the circuit.

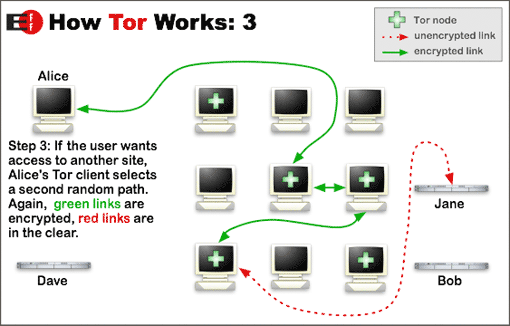
After these Tor circuits are set up, the client node can transmit its data packets through them. Every circuit can transmit several TCP data streams which are Internet requests such as http stream request and so on. This method that several streams share one circuit is very high-efficient for Tor anonymous communication network. Through these 3 mediate server nodes, the client node can exchange data with the final destination privately and securely without

disclosing any sensitive address information such as IP address, packet time, and packet size. The data packets transmitted by the client node are like a three-layer onion with encryption. One hop along the circuit only can peel one layer with its private key. Every hop through the Tor circuit only knows its previous and next point. The Tor network data packet structure is as follows:



**Figure 4 The structure of Tor data packet1**

The last mediate server node of Tor circuit is called exit node. From the initial client node to the exit server node, the data packets are encrypted by TLS encryption. The data packet is not encrypted from exit node to the final destination. As we discussed before, there are some weakness for this unencrypted segment of Tor circuit. Some sniffers can collect some sensitive information of the data packet although they can not know the detailed address and location information.



**Figure 5 Tor router setup step 3 1**

In order to be high efficient, every Tor circuit is used about ten minutes for many connections. But to keep network sniffers from monitoring Tor users according to their earlier circuit information, every ten minutes Tor client user will set up a new Tor circuit for new TCP streams. On the above figure, when the client node Alice used one Tor circuit about ten minutes, and she wants to visit another website Jane, the client node Alice will select three different server nodes to set up a new Tor circuit to transmit data packets to web server Jane.

Tor anonymous communication network use above methodology to transmit network data among Tor nodes. Tor network is an excellent supplement tool for all kinds of encryption ways that encrypt data packet payload part. It improves the privacy of packets’ location, address, time, and size information dramatically. Tor network achieves anonymity through the distributed and twisted route. Next we will discuss Tor network node types.

|  |
| --- |
|  |
|  |
|  |

**Hardware/ software requirements**:

Hardware Requirements

* Windows 7 or latest version
* 2 GB ram
* Pentium 4 + Processor
* 1GB free Hard Disk
* Wifi Router

Software Requirements

* Eclipse
* Server on Java or php

**Conclusion**

We have presented a practical traffic analysis attack against Tor, the most popular low latency anonymity preserving system, that relies on using data from existing monitoring framework, already installed in network devices. We have demonstrated the feasibility of launching such attack to determine source of anonymous traffic.

**Reference**

1. Roger Dingledine, Nick Mathewson, and Paul Syverson. Tor: The Second-Generation Onion Router. Retrieved June 25, 2008, from http://www.torproject.org/svn/trunk/doc/design-paper/tor-design.html

## 2. Tor project organization. Tor knowledge and information. Retrieved June 26, 2008, from https://www.torproject.org/

## 3. Wikepedia. Tor (anonymity network). Retrieved June 27, 2008, from http://en.wikipedia.org/wiki/Tor\_(anonymity\_network)

## 4. Nadir Soualem. Crontab: Schedule Tasks. Retrieved June 27, 2008, from http://www.math-linux.com/spip.php?article45

## 5. Wikepedia. Traffic analysis. Retrieved June 27, 2008, from http://en.wikipedia.org/wiki/Traffic\_analysis

## 6. Wikepedia. Symmetric-key algorithm, Retrieved June 27, 2008, from http://en.wikipedia.org/wiki/Symmetric\_key

## 7. Wikepedia. Onion routing, Retrieved June 28, 2008, from http://en.wikipedia.org/wiki/Onion\_routing

## 8. Wikepedia. Vidalia project, Retrieved June 29, 2008, from http://en.wikipedia.org/wiki/Vidalia\_project

## 9. Wikepedia. Privoxy, Retrieved June 30, 2008, from http://en.wikipedia.org/wiki/Privoxy

## 10. Tor project organization. Tor Protocol Specification, Retrieved June 27, 2008, from https://www.torproject.org/svn/trunk/doc/spec/tor-spec.txt