The Onion Routing

Research · October 2018

DOI: 10.13140/RG.2.2.10181.09448

2 authors:

Deepanshu Choudhary
Symbiosis Institute of Computer Studies and Research
2 PUBLICATIONS **3** CITATIONS
SEE PROFILE

Rahul Bhagat
Symbiosis Institute of Computer Studies and Research
1 PUBLICATION **0** CITATIONS
SEE PROFILE

The Onion Routing

Deepanshu Sandeep Chaudhary
Symbiosis Institute of Computer Studies & Research
Email : dec1743005@sicsr.ac.in

Rahul Neelratan Bhagat
Symbiosis Institute of Computer Studies & Research
Email : rab1743017@sicsr.ac.in

ABSTRACT

Onion Routing is basically an infrastructure to maintain anonymity over the public network as its functionality provides features against eavesdropping and traffic analysis. Every piece of identifying information is carried within data streams over an anonymous connection. The network provides anonymity to both clients and servers, functioning as a 'black box' that hides the routing information of network participants. This document also provides a historical view of the anonymity systems. An in-depth study of the TOR(The Onion Router) system is also conducted examining its development, its features, its limitations and its weaknesses. The method used to locate any TOR identifying markers is via a packet comparison of TOR and non-TOR identical network packets. Recommendations are given regarding the usage of TOR to mitigate the behavioral actions of users that have inadvertently violated their anonymity.

INTRODUCTION

Onion routing is a technique for anonymous communication over a computer network. In an onion network, messages are encapsulated in layers of encryption, analogous to layers of an onion.

Onion Routing protects its communications against traffic analysis attacks. It makes it very hard for network observers (such as crackers, companies, and governments) to reliably learn who is talking to whom and for what purpose, by examining data packets flowing over the network.

This paper will help the test the extent that TOR provides whenever the traffic is subject to techniques for traffic analysis.It has been noted through traffic analysis and other methods one can determine the origin of the IP packet,but through TOR this can be avoided.Prior moving ahead into the analysis of TOR traffic it is essential to have some understanding about the term anonymity and brief history of the same.We will also be discussing regarding the technology behind the Onion Router and the way it provides anonymity on the internet.

ANONYMITY ON THE INTERNET

Danezis and Diaz define anonymity as “the state of being not identifiable within a set of subjects, the anonymity set”. This definition says that a user on the internet should not be identifiable through their network traffic any more than any user of the internet, that it should not user identity characteristics while analyzing the network traffic. It is to be noted that in network traffic to contain identification features such as Ip headers and port number so that the computers at receiving end should be able to recognize the source and destination Ip addresses. A more suitable term for systems like these could be “unlinkability” which is defined as to ensures that a user may make multiple uses of resources or services without others being able to link these uses together. Unlinkability requires that users and/or subjects are unable to determine whether the same user caused certain specific operations in the system.In simple terms it is defined as a layout where it is impossible to link the network traffic to a particular user which is more precise for describing the level of service that the systems provide over the internet.Few examples of these systems are :- Proxy Server;Anonymous email clients, and MIX or Crowd systems.The idea behind the proxy servers is to route the network traffic through a proxy to hide original Ip address of the internet connection.The motive behind these systems is to provide basic “unlinkability” and allow users

the ability to access Ip location-specific content.These type of systems are called as “one hop” proxy server as the network traffic is only routed through one proxy server at a time.Whereas the MIX networks route traffic through various nodes before reaching to a destination.Another purpose of MIX system is to combine together many messages for an intruder to follow messages through it.One such attack is outlined by Reiter and Rubin who explain that these systems “can be undermined by executable web content that, if downloaded into the user's browser, can open network connections directly from the browser to web servers, thus bypassing Crowds altogether and exposing the user to the end server”.By utilising end-to-end traffic analysis techniques other inadequacies of these systems can also be highlighted:

Another attack tries to correlate events at the endpoints of the system: if a user makes an HTTP request, it is reasonable to assume that this request leaves the last MIX towards a web server shortly later. Similarly, the response sent from the web server to the last MIX will appear on the link between first MIX and user within some seconds.

Onion routing is a technique for anonymous communication over a computer network. In an onion network, messages are encapsulated in layers of encryption, analogous to layers of an onion.

Onion Routing protects its communications against traffic analysis attacks. It makes it very hard for network observers (such as crackers, companies, and governments) to reliably learn who is talking to whom and for what purpose, by examining data packets flowing over the network.

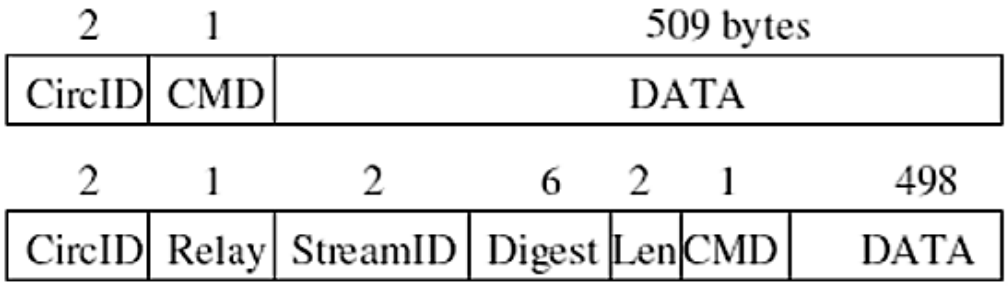
HISTORY AND INTENDED USE OF THE ONION ROUTER

One of the most widely deployed onion-routing anonymizing systems is TOR (The Onion Router).Tor's intention is to allow untied and anonymous communication over the Internet. Tor allows anyone to connect to websites that may be blocked by oppressive governments, allows whistleblowers to communicate with officials anonymously, and gives a means for legitimate communication between businesses and persons who desire to keep their private conversations private. Despite multiple helpful and positive purposes, the Tor browser can be used to either facilitate crimes or commit crimes. Although Tor was initially developed by the US government in 2002, it is not presently controlled by the US government. In fact, Tor is practically not controlled by any one entity but rather open for improvements by virtually anyone with the technical ability to test and improve it. For that reason alone, Tor receives worldwide input from privacy motivated experts to ensure it remains relevant and effective. As a point of irony, the US government not only created Tor but is also researching methods to deanonymize users of it.

DESIGN

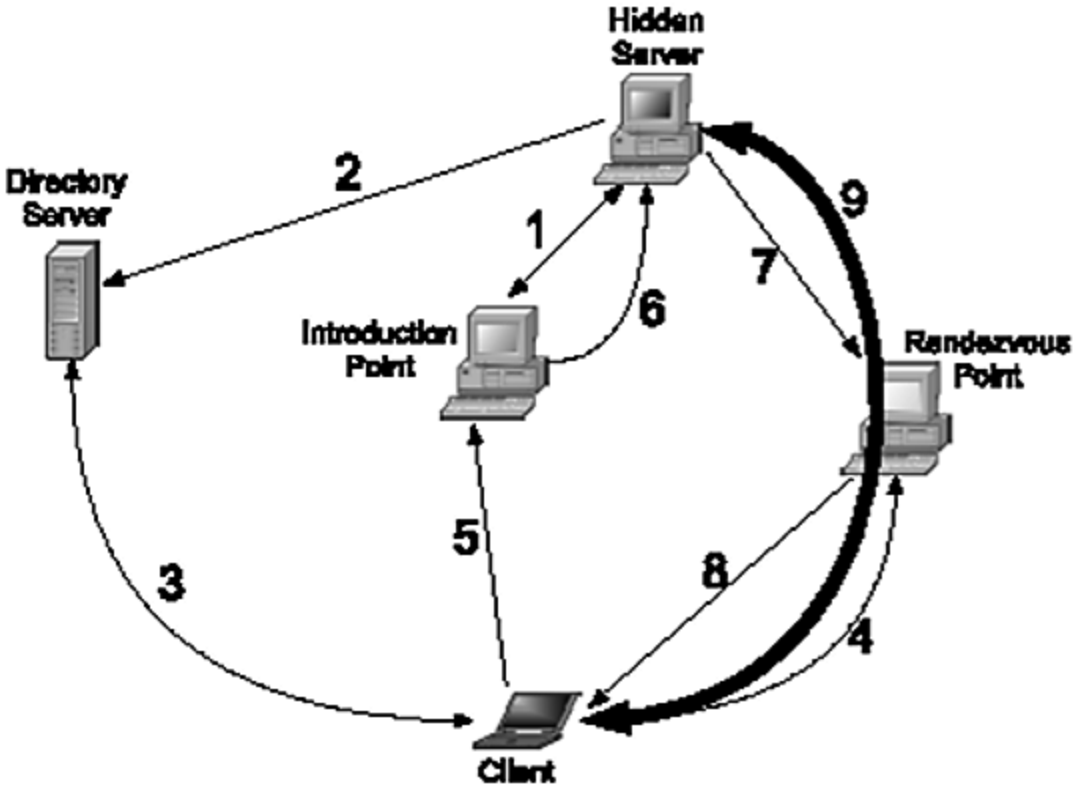
The TOR system follows on from traditional onion routing services, that is it utilizes proxy servers in order to spoof an IP address so that the originating IP address remains unknown. TOR can be seen as a mix between onion routing and crowd systems. The TOR system "tunnels everything over TCP Port 80 over a network of relays, and is particularly well tuned to work for web traffic, with the help of the 'Privoxy' content sanitizer. Privoxy is a web proxy service which modifies "web page data and HTTP headers and is commonly used for "removing ads and other obnoxious Internet junk. In the case of TOR this web cache assists in removing web traffic that could reveal the true IP address of the user, such as Javascript or Flash content. Unlike traditional onion routing services TOR does not send the traffic through in its original packet format, instead TOR uses fixed-length "Cells" to transfer data. Each Cell consists of a header and a payload . As stated by Fraser et. al, "TOR operates using fixed 512 byte cells (or packets) for stronger anonymity and the Transport Layer Security (TLS) protocol for authentication and privacy. Coupled with this Cell-based

design, TOR utilises “Circuits” to choose the path that the data will take as well as which protocol layer to anonymise: “they may intercept IP packets directly, and relay them whole (stripping the source address) along the circuit.



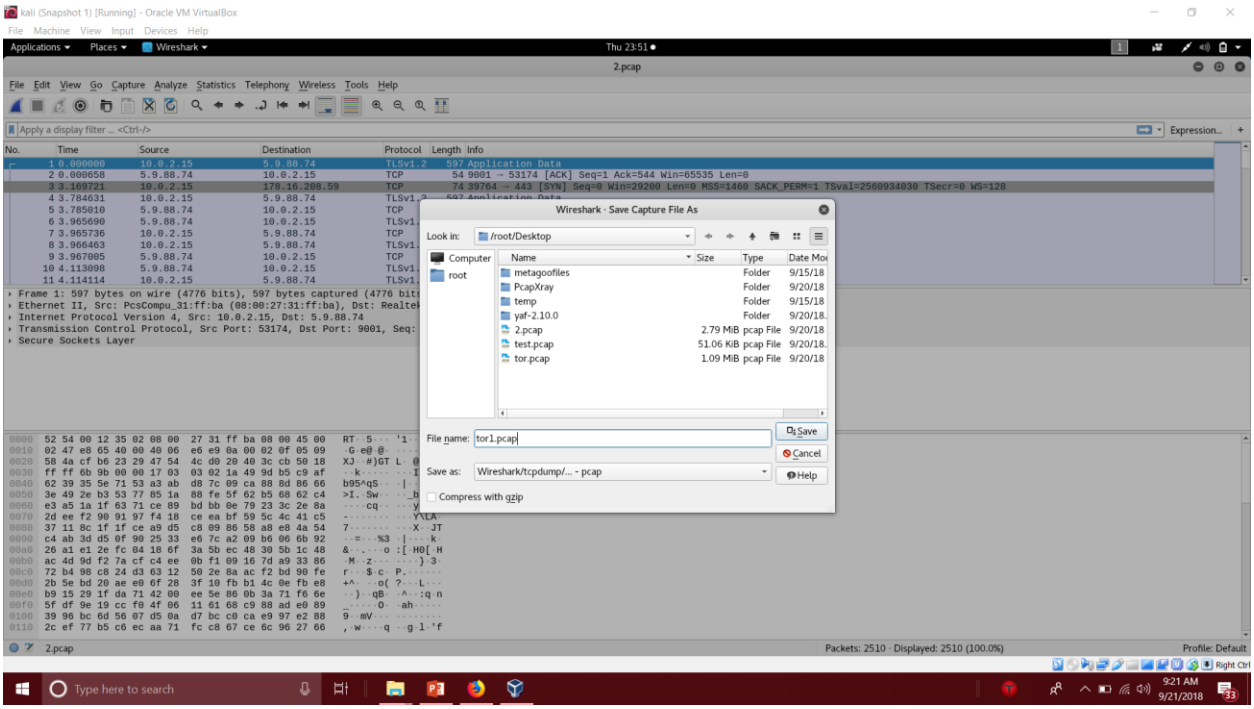
HIDDEN SERVICES

Another benefit of the TOR system over traditional onion routers is that it allows users to host content on the internet that can only be accessed via the use of the TOR system, these are known as "Hidden Services". These hidden services are denoted by the use of the virtual Top Level Domain (TLD) ".onion" which is the address entered by the user to connect to this type of service. When connecting to a hidden service a user creates a new circuit to the hidden service's rendezvous point which adds an extra layer of protection. As claimed by Dingledine, "this type of anonymity protects against Distributed-Denial-of-Service attacks: attackers are forced to attack the onion routing network because they do not know the host's IP address


Hidden
Server
Directory
Server
Introduction
Point
Rendezvous
Point
Client
1
2
3
4
5
6
7
8
9

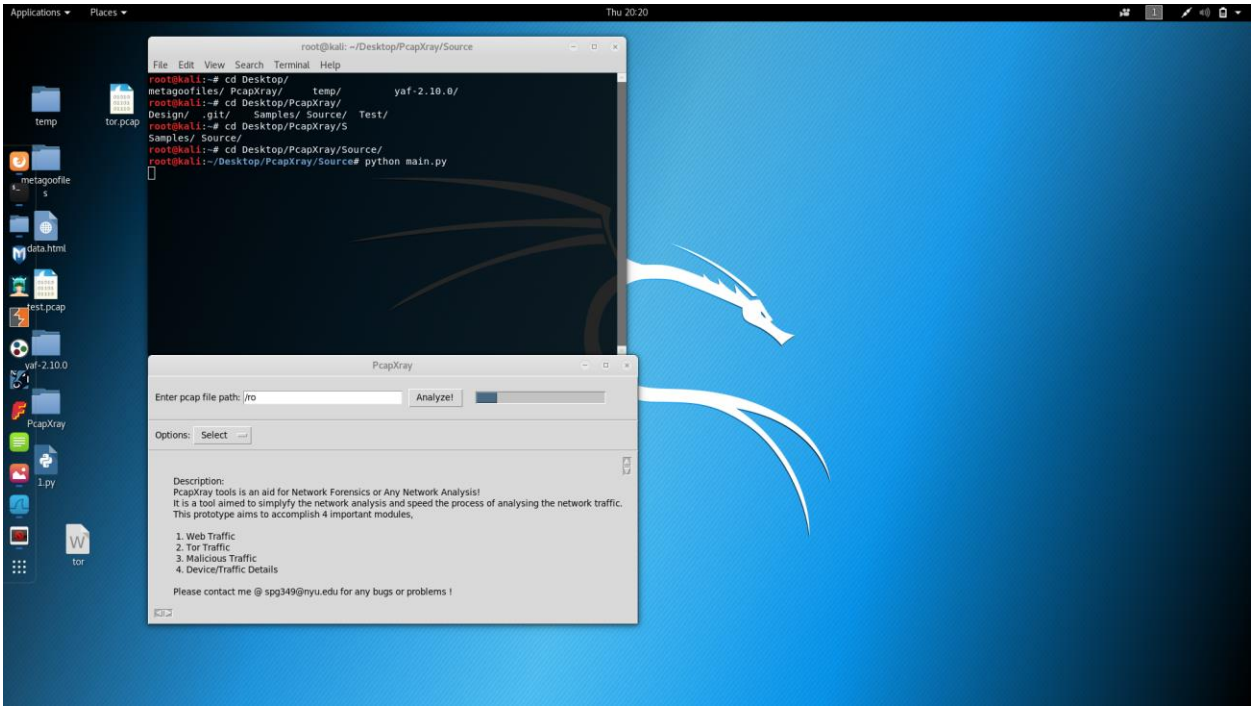
METHODOLOGY

We carried out a TOR packet analysis experiment to find out whether it is easily possible to detect whether a packet is a TOR packet or not. Here we used a Kali VM to send TOR packets over the network and a packet sniffing and capturing tool known as Wireshark to capture the packets.

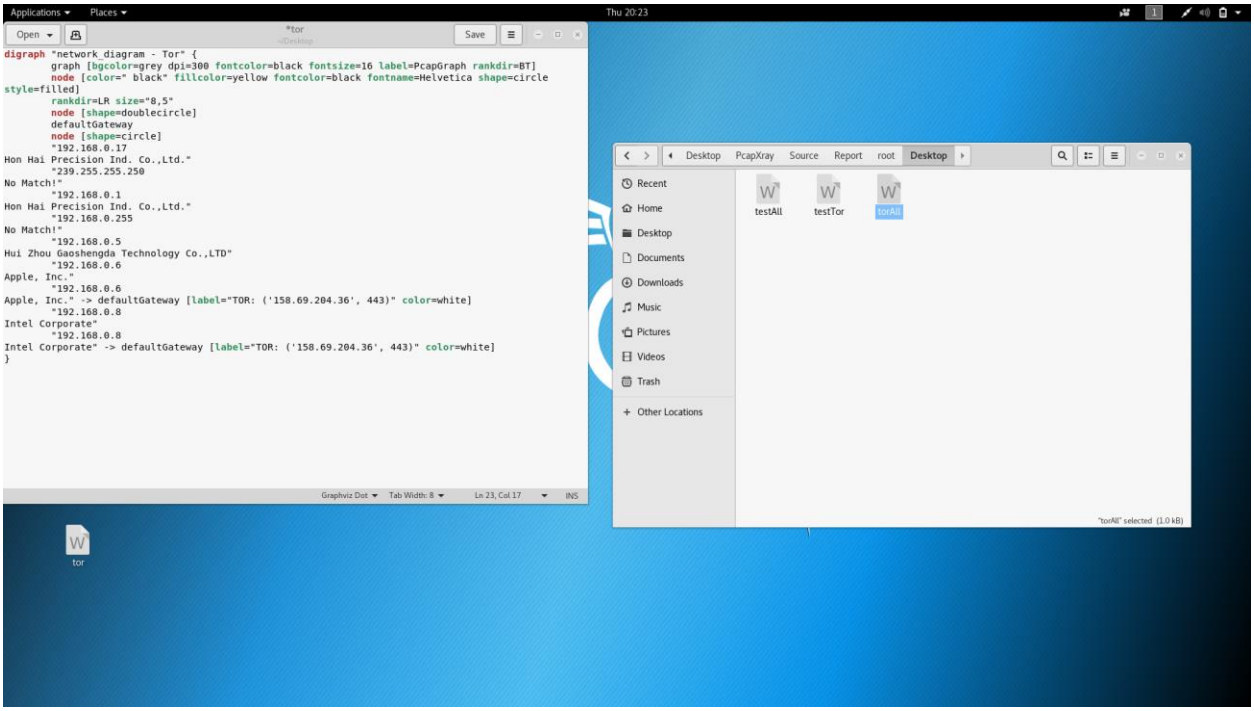


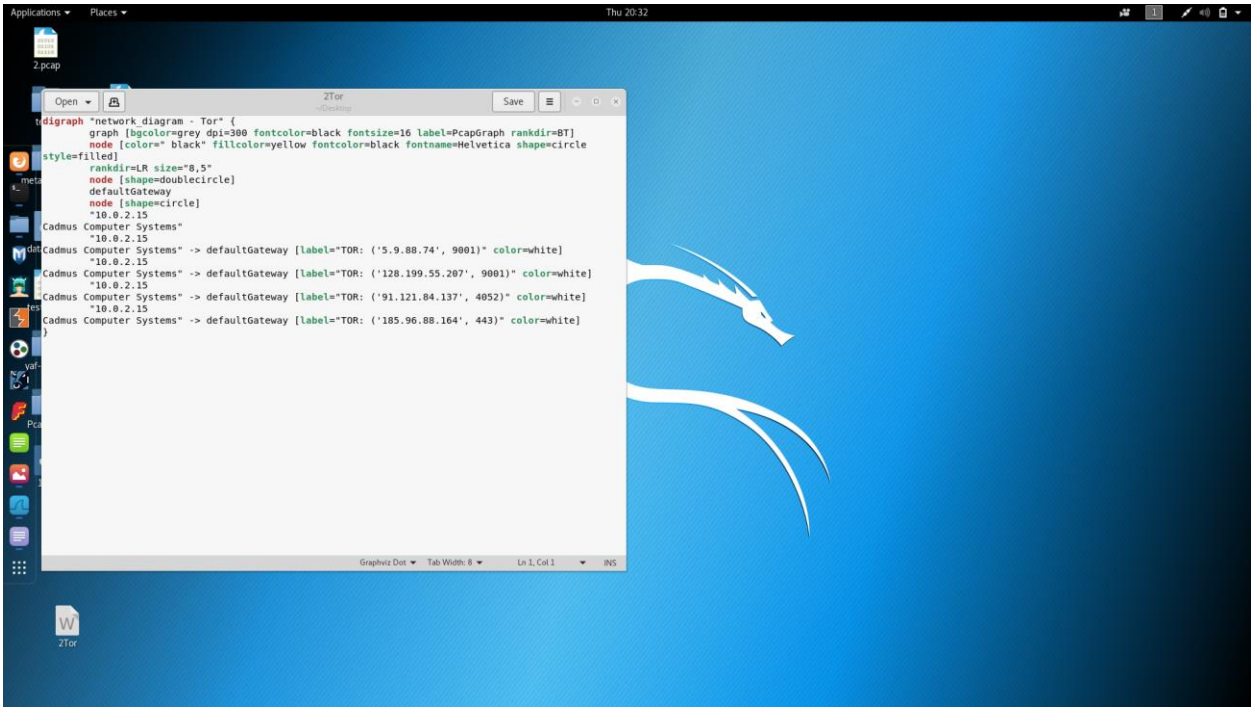
Next, we saved the captured packet list in a ".pcap" file. This pcap file can be used with a pcap Tool. Here, we used PcapXray which is a network forensic tool

used to visualize a packet capture offline as a network diagram including device identification, highlight important communication and file extraction.



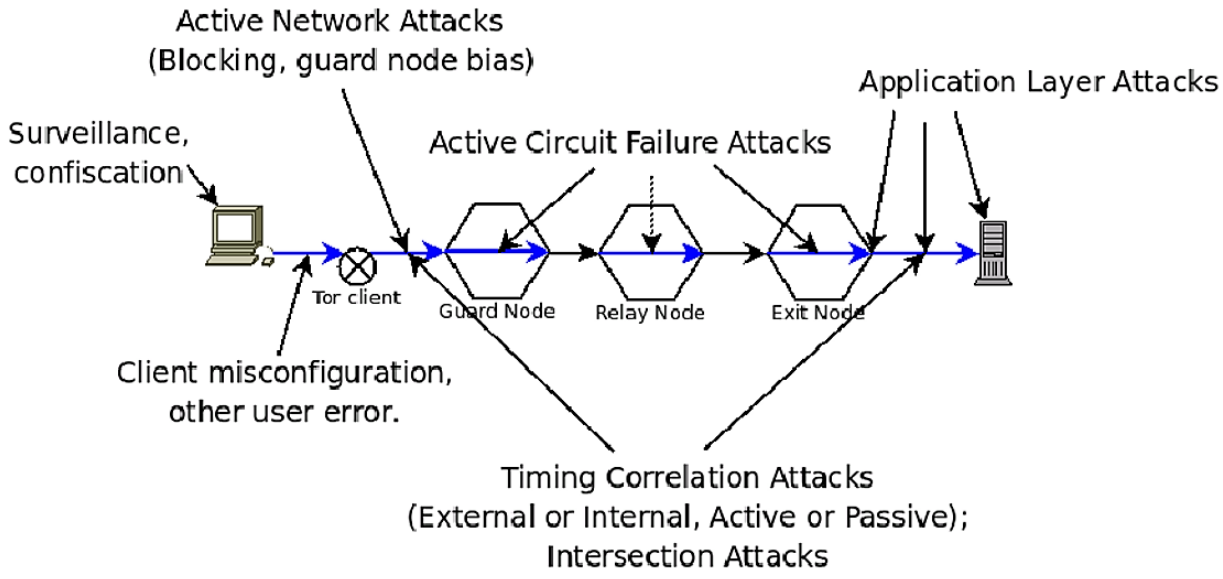
Using this tool, we are able to identify whether a packet is a TOR packet or not. This could help an attacker further investigate the TOR packets and reach the origin or the source client.


Applications
Places
Thu 20:23
Open
*tor
~/Desktop
Save
digraph "network_diagram - Tor" {
graph [bgcolor=grey dpi=300 fontcolor=black fontsize=16 label=PcapGraph rankdir=BT]
node [color=" black" fillcolor=yellow fontcolor=black fontname=Helvetica shape=circle
style=filled]
rankdir=LR size="8,5"
node [shape=doublecircle]
defaultGateway
node [shape=circle]
"192.168.0.17
Hon Hai Precision Ind. Co.,Ltd."
"239.255.255.250
No Match!"
"192.168.0.1
Hon Hai Precision Ind. Co.,Ltd."
"192.168.0.255
No Match!"
"192.168.0.5
Hui Zhou Gaoshengda Technology Co.,LTD"
"192.168.0.6
Apple, Inc."
"192.168.0.6
Apple, Inc." -> defaultGateway [label="TOR: ('158.69.204.36', 443)" color=white]
"192.168.0.8
Intel Corporate"
"192.168.0.8
Intel Corporate" -> defaultGateway [label="TOR: ('158.69.204.36', 443)" color=white]
}
Graphviz Dot
Tab Width: 8
Ln 23, Col 17
INS
Desktop
PcapXray
Source
Report
root
Desktop
Recent
Home
Desktop
Documents
Downloads
Music
Pictures
Videos
Trash
Other Locations
testAll
testTor
torAll
"torAll" selected (1.0 kB)
tor


Applications
Places
Thu 20:32
2.pcap
Open
2Tor
~/Desktop
Save
digraph "network_diagram - Tor" {
graph [bgcolor=grey dpi=300 fontcolor=black fontsize=16 label=PcapGraph rankdir=BT]
node [color=" black" fillcolor=yellow fontcolor=black fontname=Helvetica shape=circle
style=filled]
rankdir=LR size="8,5"
node [shape=doublecircle]
defaultGateway
node [shape=circle]
"10.0.2.15
Cadmus Computer Systems"
"10.0.2.15
Cadmus Computer Systems" -> defaultGateway [label="TOR: ('5.9.88.74', 9001)" color=white]
"10.0.2.15
Cadmus Computer Systems" -> defaultGateway [label="TOR: ('128.199.55.207', 9001)" color=white]
"10.0.2.15
Cadmus Computer Systems" -> defaultGateway [label="TOR: ('91.121.84.137', 4052)" color=white]
"10.0.2.15
Cadmus Computer Systems" -> defaultGateway [label="TOR: ('185.96.88.164', 443)" color=white]
}
Graphviz Dot
Tab Width: 8
Ln 1, Col 1
INS
2Tor

LIMITATIONS & WEAKNESSES

The TOR system is not without its share of limitations; Danezis and Diaz raise the point that "one notable difference between TOR and previous attempts at anonymizing streams of traffic, is that it does not claim to offer security against even passive global observers. In fact Lemos states that "the problem is known to both the Tor Project, which advises everyone to use end-to-end encryption, and to security researchers. This limitation accumulates to the following point, "an adversary, who can observe a stream at two different points, can trivially realize it is the same traffic.This limitation leads to weaknesses that can be exploited to undermine the anonymity of the TOR system



One of the weaknesses, if not the biggest weakness, of the Tor browser is the user. As the browser is preconfigured with security in mind, customization is not recommended. In fact, the best thing a Tor user can do is not to change any settings of the browser because anyone setting can leak information out of Tor. The man-in-the-middle attack is yet another method to bypass the security of Tor users by interjecting a capture service between the Tor user and destination. Nation-states have the resources for these types of attacks on Tor, but even then, compromising Tor is very difficult. Each of these methods requires more resources and time than will ever be given to the common criminal unless special situations exist, such as a terror connection. Even then, the number of agencies with access to such resources is very few. Given that, the few remaining methods rely on the suspect and the suspect's errors.

CONCLUSION / RECOMMENDATIONS

Through this paper, we have shown that TOR packets can be captured and analyzed at the exit or the entry nodes. However, tracking back to the origin is not easy. In a LAN environment it is possible, but when it comes to internet, it would be quite hectic to get back to the source node.

USED IN COMBINATION OF OTHER TOOLS AND METHODS

Tor, when used by itself works well. When used in combination with other security methods works perfectly as an anonymous communication tool. As you have seen, the last hop of data in the Tor circuit is unencrypted. Theoretically, this data could be compromised. However, if the data was encrypted, not only is the transmission anonymous, but the information is encrypted. At that point, even if both ends of the communication are identified, the encrypted contents are secure if end-to-end encryption has been employed. If a Tor user adds an additional layer of protection by using a nonowned computer on a nonowned network, the odds of being identified are even smaller. This would be the case of using Tor on a public computer at a library or hotel lobby as both the sender and receiver of electronic communications.

TAILS

As Tor is a browser that can be used anonymously, Tails (The Amnesic Incognito Live System) is a complete operating system that can be used anonymously. Tails is based on Debian/Linux and runs from a DVD, USB flash

drive, or SD card and does not need to be installed. In the context of the Tor browser, the Tor browser is preinstalled and preconfigured in Tails. The operation is the same as described previously with all the benefits of anonymity. However, Tails adds even more security and anonymity to covert communications and web surfing. To run Tails, a computer must be able to boot to an external device that contains Tails. Once booted to Tails, the user needs to only connect to an active Internet connection and the Tor browser can be used immediately. The most substantial difference in using Tails compared to using the Tor browser in Windows or other installed operating system is that Tails does not leave any trace on the host computer system. A Tails user can prevent any forensic artifacts being created since Tails does not touch the host computer hard drive. Additionally, you have seen that Tor forensic artifacts and even URLs can be found in memory and the hiberfil.sys file. Tails does not create, use, or write to a pagefile.sys or hiberfil.sys. Moreover, the live memory is wiped on shutdown. Even when run from a writable USB flash drive, data is not saved to the flash drive from the Tails use.

RELATED TOR TOOLS AND APPLICATIONS

The Tor network works well for its intended purpose and because of its effectiveness, Third-party tools capitalize on it. One commercial example is the Anonabox (http://www.anonabox.com). The Anonabox is a hardware router that routes all Internet traffic through Tor, rather than only the Tor browser being routed through the Tor network. Devices such as the Anonabox should be considered when seizing computer systems for analysis. Devices such as the Anonabox reduce security errors made by Tor browser users by eliminating the risk of using a non-Tor browser connection as the entire Internet connection runs through Tor. By running the entire Internet connection through a device

such as this, suspects can use any web browser, not just Tor, without their true IP address being disclosed.

SUMMARY

Tor is the most commonly used anonymous Internet tool in the world and is used for both legitimate and illicit communication. Although identifying the illicit users of Tor is nearly impossible, any forensics investigation should not discount the possibility of Tor use by a suspect as a means of covert communication. Identifying communications between persons requires more than just identifying the Internet traffic or identifying the persons involved. A complete picture is identifying the suspects as well as the contents of their communications.

REFERENCES

Wikipedia.org, *"Onion Routing"* - https://en.wikipedia.org/wiki/Onion_routing

Make Use of, *"What is onion routing, exactly"* – https://www.makeuseof.com/tag/what-is-onion-routing-exactly-makeuseof-explains/

E-security planet, *"The problem with Tor"* - https://www.esecurityplanet.com/open-source-security/the-trouble-with-tor.html

Geeks for Geeks, *"How onion routing works"* - https://www.geeksforgeeks.org/onion-routing/

YouTube, *"Computerphile"* - https://www.youtube.com/watch?v=QRYzre4bf7Ihttps://www.y

https://www.torproject.org/docs/documentation.html.en

https://metrics.torproject.org/exonerator.html