DARK WEB

Palash Jain

M.Tech Student, R.N. Modi Engineering College (RMEC)

Kota , Rajasthan, (India)

ABSTRACT

It's just 25 years since the World Wide Web was created.It now touches all of our lives our personal informationand data swirling through the internet on a daily basis. It's now caught in the greatest controversy of its life-surveillance. This paper deals about the world of dark web and its role in preventing anonymity of a user. It also includes a difference between deep search and dark web.

Keywords: TOR,Onion Routing, I2P, PublicKeys, Encrypted, HTTP, Hidden, Network Layer, DeepWeb.

1. INTRODUCTION

The Dark Web is a collection of thousands of websites that use anonymity tools like Tor and I2P to hide their IP address. While it's most famously been used for black market drug sales and even child pornography, the Dark Web also enables anonymous whistle blowing and protects users from surveillance and censorship. The majority of Dark Web sites use the anonymity software Tor, though a smaller number also use a similar tool called I2P. Both of those systems encrypt web traffic in layers and bounce it through randomly-chosen computers around the world, each of which removes a single layer of encryption before passing the data on to its next hop in the network. In theory, that prevents any spy even one who controls one of those computers in the encrypted chain from matching the traffic's origin with its destination.

When web users run Tor, for instance, any sites they visit can't easily see their IP address. But a web site that itself runs Tor what's known as a Tor hidden service can only be visited by Tor users. Traffic from both the user's computer and the web server takes three hops to a randomly chosen meet-up point in the Tor network, Anyone who runs Tor and knows a site's url, which for Tor hidden services ends in ".onion," can easily visit those illegal online marketplaces.

II ONION ROUTING

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 the vegetable onion. The encrypted data is transmitted through a series of network nodes called onion routers, each of which "peels" away a single layer, uncovering the data's next destination. When the final layer is decrypted, the message arrives at its destination. The sender remains anonymous because each intermediary knows only the location of the immediately preceding and following nodes.

Onion routing is implanted by encryption in thethe application layer of a communication protocol stack, nested like the layers of an onion. Tor encrypts the data, including the destination IP address, multiple times and sends it through avirtual circuit comprising successive, randomly selected Tor relays. Each relay decrypts a layer of encryption to reveal only the next relay in the circuit in order to pass the remaining encrypted data on to it. The final relay decrypts the innermost layer of encryption and sends the original data to its destination without revealing, or even knowing, the source IP address. Because the routing of the communication is partly concealed at every hop in the Tor circuit, this method eliminates any single point at which the communicating peers can be determined through network surveillance that relies upon knowing its source and destination.Second. This process can be repeated to build larger and larger chains, but is typically limited to preserve performance.

When the chain is complete, the originator can send data over the Internet anonymously. When the final recipient of the data sends data back, the intermediary nodes maintain the same link back to the originator, with data again layered, but in reverse such that the final node this time removes the first layer of encryption and the first node removes the last layer of encryption before sending the data, for example a web page, to the originator.

2.1 Onion Creation and Transmission

To create and transmit an onion, the originator selects a set of nodes from a list provided by a "directory node". The chosen nodes are arranged into a path, called a "chain" or "circuit", through which the message will be transmitted. To preserve the anonymity of the sender, no node in the circuit is able to tell whether the node before it is the originator or another intermediary like itself. Likewise, no node in the circuit is able to tell how many other nodes are in the circuit and only the final node, the "exit node", is able to determine its own location in the chain.

Using asymmetric key cryptography, the originator obtains a public key from the directory node to send an encrypted message to the first ("entry") node, establishing a connection and a shared secret ("session key"). Using the established encrypted link to the entry node, the originator can then relay a message through the first node to a second node in the chain using encryption that only the second node, and not the first, can decrypt. When the second node receives the message, it establishes a connection with the first node. While this extends the encrypted link from the originator, the second node cannot determine whether the first node is the originator or just another node in the circuit. The originator can then send a message through the first and second nodes to a third node, encrypted such that only the third node is able to decrypt it. The third, as with the second, becomes linked to the originator but connects only with the second. This process can be repeated to build larger and larger chains, but is typically limited to preserve performance.

When the chain is complete, the originator can send data over the Internet anonymously. When the final recipient of the data sends data back, the intermediary nodes maintain the same link back to the originator, with data again layered, but in reverse such that the final node this time removes the first layer of encryption and the first node removes the last layer of encryption before sending the data, for example a web page, to the originator.

III TOR

Tor is free software for enabling anonymous communication. The name is an acronym derived from the original software project name The Onion Router. Tor directs Internet traffic through a free, worldwide, volunteer network consisting of more than six thousand relays to conceal a user's location and usage from anyone conducting network surveillance or traffic analysis. Using Tor makes it more difficult for Internet activity to be traced back to the user: this includes "visits to Web sites, online posts, instant messages, and other communication forms". Tor's use is intended to protect the personal privacy of users, as well as their freedom and ability to conduct confidential communication by keeping their Internet activities from being monitored.

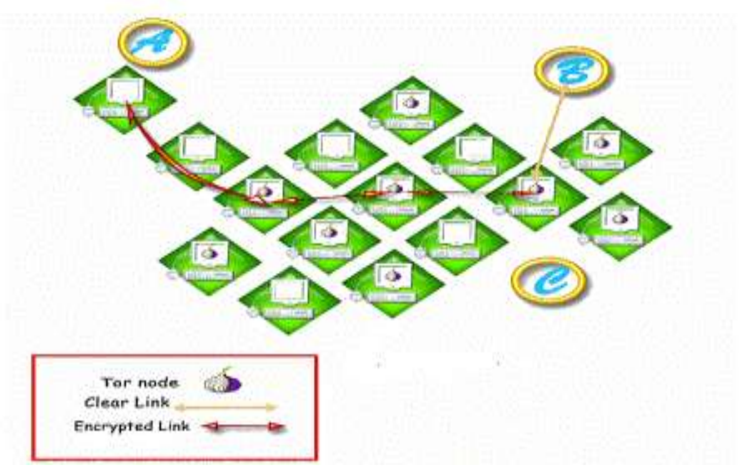

Figure 1. Tor Circuit

A. Operation

Tor aims to conceal its users' identities and their online activity from surveillance and traffic analysis by separating identification and routing. It is an implementation of onion routing, which encrypts and then randomly bounces communications through a network of relays run by volunteers around the globe. These onion routers employ encryption in a multi-layered manner to ensureperfect forward secrecy between relays, thereby providing users with anonymity in network location. That anonymity extends to the hosting of censorship-resistant content by Tor's anonymous hidden service feature Furthermore, by keeping some of the entry relays (bridge relays) secret, users can evade Internet censorship that relies upon blocking public Tor relay Because the IP address of the sender and the recipient are not both in cleartext at any hop along the way, anyone eavesdropping at any point along the communication channel cannot directly identify both ends. Furthermore, to the recipient it appears that the last Tor node (called the exit node), rather than the sender, is the originator of the communication.

B. Originating

A Tor user's SOCKS-aware applications can be configured to direct their network traffic through a Tor instance's SOCKS interface. Tor periodically creates virtual circuits through the Tor network through which it can multiplex and onion-route that traffic to its destination. Once inside a Tor network, the traffic is sent from router to router along the circuit, ultimately reaching an exit node at which point thecleartext packet is available and is forwarded on to its original destination. Viewed from the destination, the traffic appears to originate at the Tor exit node.

C. Hidden services

Tor can also provide anonymity to websites and other servers. Servers configured to receive inbound connections only through Tor are called hidden services. Rather than revealing a server's IP address (and thus its network location), a hidden service is accessed through its onion address. The Tor network understands these addresses and can route data to and from hidden services, even those hosted behind firewalls or network address translators (NAT), while preserving the anonymity of both parties. Tor is necessary to access hidden services.

Hidden services have been deployed on the Tor network since 2004Other than the database that stores the hidden-service descriptors Tor is decentralized by design; there is no direct readable list of all hidden services, although a number of hidden services catalog publicly known onion addressesBecause hidden services do not use exit nodes, connection to a hidden service is encrypted end-to-end and not subject to eavesdropping. There are, however, security issues involving Tor hidden services. For example, services that are reachable through Tor hidden servicesand the public Internet are susceptible to correlation attacks and thus not perfectly hidden. Other pitfalls include misconfigured services (e.g. identifying information included by default in web server error responses), uptime and downtime statistics, intersection attacks, and user error.

Hidden services could be also accessed from a standard web browser without client-side connection to the Tor network;using services like Tor2web Popular sources of dark web .onion links include Pastebin, Twitter, Reddit and other Internet forums.

D. Attack

1. Traffic-analysis attack

Steven J. Murdoch and George Danezis from University of Cambridge presented an article at the 2005 IEEE Symposium on security and privacy on traffic-analysis techniques that allow adversaries with only a partial view of the network to infer which nodes are being used to relay the anonymous streams. These techniques greatly reduce the anonymity provided by Tor. Murdoch and Danezis have also shown that otherwise unrelated streams can be linked back to the same initiator. This attack, however, fails to reveal the identity of the original user. Murdoch has been working with and has been funded by Tor since 2006.

2 Tor exit node block

Operators of Internet sites have the ability to prevent traffic from Tor exit nodes or to offer reduced functionality to Tor users. For example, it is not generally possible to editWikipedia when using Tor or when using an IP address that also is used by a Tor exit node, due to the use of the Tor Block MediaWiki extension, unless an exemption is obtained. The BBC blocks the IP addresses of all known Tor relays from its iPlayer service—including guards, relays, and exit nodes—regardless of geographic location. Bridge relays are not affected.

3 Bad apple attack

In March 2011, researchers with the Rocquencourt French Institute for Research in Computer Science and Automation documented an attack that is capable of revealing the IP addresses of BitTorrent users on the Tor network. The "bad apple attack" exploits Tor's design and takes advantage of insecure application use to associate the simultaneous use of a secure application with the IP address of the Tor user in question. One method of attack depends on control of an exit node or hijacking tracker responses, while a secondary attack method is based in part on the statistical exploitation of distributed hash tabletracking. According to the study:This attack against Tor consists of two parts: (a) exploiting an insecure application to reveal the source IP address of, or trace, a Tor user and (b) exploiting Tor to associate the use of a secure application with the IP address of a user (revealed by the insecure application). As it is not a goal of Tor to protect against application-level attacks, Tor cannot be held responsible for the first part of this attack. However, because Tor's design makes it possible to associate streams originating from secure application with traced users, the second part of this attack is indeed an attack against Tor. We call the second part of this attack the bad apple attack. The results presented in the bad apple attack research paper are based on an attack in the wild launched against the Tor network by the authors of the study. The attack targeted six exit nodes, lasted for 23 days, and revealed a total of 10,000 IP addresses of active Tor users. This study is particularly significant because it is the first documented attack designed to target P2P file-sharing applications on Tor.BitTorrent may generate as much as 40% of all traffic on Tor.Furthermore, the bad apple attack is effective against insecure use of any application over Tor, not just BitTorrent.

III I2P

The Invisible Internet Project (I2P) is an overlay network and dark net that allows applications to send messages to each other pseudonymously and securely. Uses include anonymous Web surfing, chatting, blogging and file transfers. The software that implements this layer is called an I2P router and a computer running I2P is called an I2P node. The software is free and open source and is published under multiple licenses. The name I2P is derived from Invisible Internet Project, which, in pseudo-mathematical notation, is represented as I^2P .I2P is beta software since 2003 Developers emphasize that there are likely to be bugs in the software and that there has been insufficient peer review to date.However, they believe the code is now reasonably stable and well-developed, and more exposure can help development of I2P.The network itself is strictly message-based (like IP), but there is a library available to allow reliable streaming communication on top of it (similar to TCP, although from version 0.6 there is a new UDP-based SSU transport). All communication is end-to-end encrypted (in total there are four layers of encryption used when sending a message), and even the end points ("destinations") are cryptographic identifiers (essentially a pair of public keys), so that neither sender nor recipient of a message need to reveal their IP address to the other side or to third-party observers.

Although many developers had been a part of the Invisible IRC Project (IIP)and Freenet communities, there are significant differences between their designs and concepts. IIP was an anonymous centralized IRC server. Freenet is a censorship-resistant distributed data store. I2P is an anonymous peer to peer distributed communication layer designed to run any traditional internet service(e.g. Usenet, email, IRC, file sharing, Web hosting and HTTP, Telnet), as well as more traditional distributed applications Many developers of I2P are known only under pseudonyms. While the previous main developer, random, is currently on hiatus, others, such as zzz, killyourtv, and Complication have continued to lead development efforts, and are assisted by numerous contributors.

2. Dark Web or Deep Web

Although all of these terms tend to be used interchangeably, they don't refer to exactly the same thing. An element of nuance is required. The 'Deep Web' refers to all web pages that search engines cannot find. Thus the 'Deep Web' includes the 'Dark Web', but also includes all user databases, webmail pages, registration-required web forums, and pages behind pay walls. There are huge numbers of such pages, and most exist for mundane reasons.The Deep Web refers to content hidden behind HTML forms. In order to get to such content, a user has to perform a form submission with valid input values The Deep Web has been acknowledged as a significant gap in the coverageof search engines because web crawlers employed by search engines rely on hyperlinks to discover new web pages and typically lack the ability to perform such form submissions There are two common approaches to offering access toDeep-Web content. The first approach is to create vertical search engines for specific domains In this approach we could create a mediator form for each domain and semantic mappings between individual data sources and the mediator form would need to be done in over 100 languages. The second approach is surfacing, which pre-computes themost relevant form submissions for all interesting HTML forms. The URL resulting from these submissions are generated off-line and indexed like any other HTML page. This approach enables leveraging the existing search engine infrastructure and hence the seamless inclusion of Deep-Web pages. User traffic is directed to Deep Web content when a user clicks on such a search result which he presumably already believes to be relevant based on its snippet. Onclicking,the user is directed to the underlying web site and hence will see fresh content. While on other hand the dark Web is the portion of the deep Web that has been intentionally hidden and is inaccessible through standardWeb browsers. Dark Web sites serve as a platform for Internet users for whom anonymity is essential, since they not only provide protection from unauthorized users, but also usually include encryption to prevent monitoring. A relatively known source for content that resides on the dark Web is found in the Tor network. The Tor network isan anonymous network that can only be accessed with aspecial Web browser, called the Tor browser.First debuted as The Onion Routing project in 2002by the US Naval Research Laboratory, it was a methodforcommunicating online anonymously. Another network,I2P, provides many of the same features that Tor does. However, I2P was designed to be a network within the Internet, with traffic staying contained in its borders. Tor provides better anonymous access to the open Internet andI2P provides a more robust and reliable “network withinthe network”.

IV CONCLUSION

Dark web is been in news for a couple of time but it has led an huge impact on internet from being a life saver to acriminal friend it has touch both good and bad aspect of society being hard to trace and crack is one of its great achievement but despite it a regular work and monitoring is required to make it more suitable and error free so it can help to sustain an anonymity of an individual also there is a need to check that it does not fall into the wrong hand and brings individualson the bright side instead of pushing him on dark side of internet.

V ACKNOWLEDGMENTS

The author would like to thank Department of Computer Engineering, R.N. Modi College of Engineering Kota for their valuable advice, moral support and technical guidance during the preparation of this paper.

REFERENCES

[1] Hsinchun Chen,WingyanChung, Jialun Qin,Edna Reid,Marc Sageman,Gabriel Weimann (2008) “Uncovering the dark Web: A case study of Jihad on the Web”,1347-59,

[2] MK Bergman,(2001)”White Paper: The Deep Web: Surfacing Hidden Value”.

[3] Gregory Fleischer (2009)“Attacking Tor at the Application Layer”.

[4] Roger Dingledine, Nick Mathewson, and PaulSyverson (2007) “Deploying low-latency anonymity:IEEE Security & Privacy”.

[5] Adam Back, Ulf Moller, and Anton Stiglic (2001) “Traffic analysis attacks and trade-offs in anonymity providing Systems”.

[6] Matthew Edman and Bulent Yener (2010) “On anonymity in an electronic society: A survey of anonymous communication systems”.

[7] Pocock, Zane (2014) “How to Navigate the Deep Web.”

[8] Siddiqui, Sameer Iqubal (2014) “Real Power of Deep Web and How to Harness It.”