# Abstract

The Tor network and Onion Services have gained high significant attention for providing anonymous and private online communications. However, maintaining anonymity in the face of potential deanonymization attacks remains a critical challenge. This project aims to conduct a comprehensive survey of Tor/Onion Hidden Service deanonymization techniques and implement an attack against a self-owned Onion service. The project explores various deanonymization techniques, categorizing them based on their approach, strengths, and weaknesses. It provides real-world examples of deanonymization attacks and their implications. Additionally, an attack is implemented against a controlled Onion service, with strict adherence to ethical guidelines and legal considerations. The attack methodology is carefully explained, highlighting the rationale behind each step. The obtained results are evaluated and analyzed, comparing them with existing research on deanonymization techniques. The project also explores mitigation strategies and ethical considerations associated with such attacks. It emphasizes the importance of protecting online anonymity, responsible disclosure, and ongoing research and development to enhance the security and privacy of Tor/Onion Services.

Table of Contents

[Abstract 1](#_Toc141698010)

[Project Breakdown 3](#_Toc141698011)

[1. Introduction 3](#_Toc141698012)

[1.1. Outline 4](#_Toc141698013)

[2. Background 4](#_Toc141698014)

[2.1.1. Tor Hidden Services in a Nutshell 4](#_Toc141698015)

[3. Categories of De-Anonymizing Techniques and Attacks. 10](#_Toc141698016)

[4. References 22](#_Toc141698017)

[5. Work Distribution 23](#_Toc141698018)

# Project Breakdown

Project 7, SRVEY2 Tor/Onion Hidden Service deanonymization techniques, Survey. Implement an attack against your own Onion service.

# Introduction

Preserving privacy in the face of widespread surveillance has become an increasingly challenging endeavor for Internet users [1]. Online threats, such as credential theft, Denial of Service (DoS) attacks, and spam, have heightened user vulnerability and underscored the need for robust privacy protection measures [2][3][4]. In response, anonymity networks have emerged as a crucial solution to safeguard user identities by providing unlikability between their IP addresses, digital fingerprints, and online activities [5]. Among these networks, Tor has garnered significant adoption as the foremost privacy-preserving network, enabling millions of users to engage in low-latency anonymous communication over the past decade [6].

Despite its success, Tor's design confronts inherent challenges related to scalability and performance, which can lead to significant delays during web browsing activities [7]. Researchers have diligently sought to enhance the user experience by proposing various improvements, encompassing congestion control, scalability, routing, and security. A comprehensive understanding of the research landscape encompassing the performance and security of low-latency anonymous communication systems, with a specific focus on Tor, is vital to effectively address online privacy threats.

This paper presents a meticulous survey of deanonymization techniques employed within Tor's onion hidden services (HS). Tor's onion services enable users to access and provide services while concealing their IP addresses. The prevalence of hidden services has witnessed an exponential rise, with over 150,000 currently in existence [9]. These services serve diverse purposes, ranging from legitimate applications like freedom of speech, journalism, and social networking, to illicit activities encompassing cybercrime and darknet markets [11-14].

Despite the considerable analysis and research conducted on Tor, a comprehensive survey that specifically centers on hidden services and their associated research challenges remains scarce. This paper aims to address this gap by conducting a systematic literature review (SLR) spanning the years [8]. Through the meticulous analysis of selected papers, we provide valuable insights into the state-of-the-art of Tor hidden services, identify prevalent research gaps, and outline potential areas for improvement.

## Outline

The subsequent sections of this survey paper are structured as follows:

* Section 2 provides a concise explanation of the operational mechanics of Tor hidden services.
* Section 3 delves into related research on Tor and hidden services.
* Section 4 meticulously outlines the methodology employed in conducting the systematic literature review (SLR).
* Section 5 presents the analysis results and their consequential implications.
* Section 6 concludes the paper, summarizing the key findings and proposing future research directions.

## Background

### Tor Hidden Services in a Nutshell

**Tor provides a mechanism called "onions" or hidden services (HS) to ensure anonymity for both recipients and service providers [15]. In the subsequent sections, we will present a brief yet informative description of the essential characteristics and operation of Tor. Furthermore, we will delve into the fundamental concept underlying communications involving Tor hidden services.**

* + - 1. Tor

Tor [15] is a distributed overlay network designed to anonymize TCP-based applications. It achieves anonymity by routing network traffic through a series of nodes called onion routers (OR) or relays. When a client wants to access a TCP-based service, such as the Web or instant messaging, they establish a connection through a path of Tor nodes.

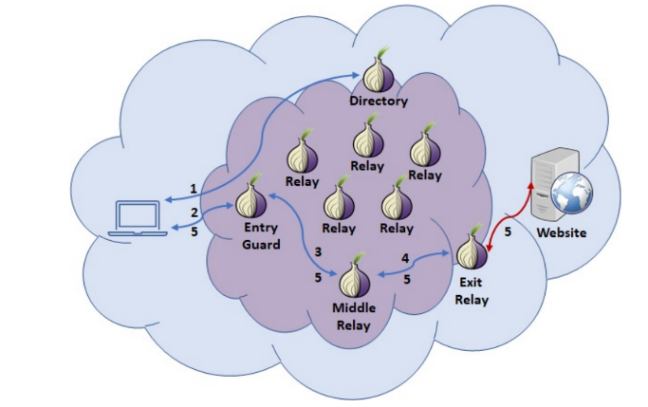


Figure The Tor Network

The process of creating a Tor circuit is incremental. The client begins by obtaining a list of available Tor nodes from the directory (step 1). From this list, the client typically selects three nodes to construct the circuit: an entry guard (or bridge), a middle relay, and an exit relay.

The circuit-building process unfolds as follows:

First, the client establishes a circuit with the entry guard (step 2). This involves creating a TLS connection between the client and the entry guard, where they engage in a Diffie-Hellman handshake to negotiate a shared key for secure communication. The circuit employs a fixed-sized data structure called a cell, with a size of 512 bytes, to exchange information.

Next, the client requests the entry guard to extend the circuit to the middle relay (step 3). Similar to the previous step, the client negotiates a key with the middle relay using the same approach. Finally, the client extends the circuit to the third relay (step 4), which serves as the exit relay responsible for forwarding the client's traffic to the desired website.

Throughout the circuit construction, each relay in the path is aware of its predecessor and successor but has no knowledge of the other relays involved. This design ensures that no single relay possesses complete information about the entire circuit, enhancing the anonymity of the client's traffic.

By establishing a path through multiple nodes and negotiating encryption keys with each relay, Tor enables users to achieve anonymity and protect their online activities. The distributed nature of Tor's overlay network and the limited knowledge of individual relays about the complete circuit contribute to maintaining user privacy.

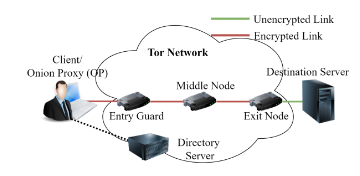


Figure Components of the Tor Network (Standard Circuit)

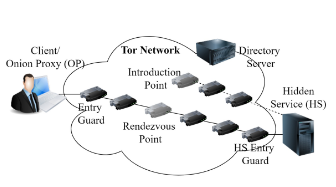


Figure Components of a Hidden Service

The descriptions of some of the key components and their features are as follows;

***Onion Proxy (OP):*** also known as a Tor proxy or simply Tor client, is a software component that enables users to access and interact with the Tor network. It acts as an intermediary between the user's device and the Tor network, facilitating communication and providing anonymity.

The Onion Proxy is typically installed on the user's device, such as a computer or mobile device. It establishes connections with the directory servers (DSs) in the Tor network, which provide information about the network's relays and hidden services. The proxy then routes the user's requests and data through a series of encrypted layers (hence the term "onion") to conceal the user's identity and location. By using the Onion Proxy, users can access Tor's features, such as accessing websites or services anonymously, bypassing censorship or surveillance, and maintaining privacy while browsing the internet. The proxy handles the technical aspects of connecting to the Tor network, allowing users to interact with Tor-enabled applications without needing to understand the underlying complexities of the network.

***Directory Servers (DS):*** Trusted and recognized servers within the network, known as Directory Servers (DS), play a crucial role. They maintain up-to-date information on the overall network status. Throughout this paper, we will use the terms "nodes," "routers," and "relays" interchangeably [1]. DSs generate a consensus document containing pertinent details about network relays, bandwidth availability, exit policies, and more. The Onion Proxies (OPs) can acquire this document from a DS, enabling them to choose three appropriate relays to establish a communication circuit towards a specific destination.

***Entry Guard Node***: The Entry Node, also known as the Guard, is a relay in the Tor network that directly connects with the client. As a result, it possesses knowledge of the client's Internet Protocol (IP) address. Unfortunately, several early attacks on Tor focused on compromising existing entry nodes or introducing new nodes to serve as entry nodes, aiming to de-anonymize users. We will discuss these de-anonymization attacks in Section 5 of this paper.

The introduction of Guard nodes brought an important feature to the entry nodes. Since Tor creates new circuits frequently, there was always a possibility that an adversary-controlled node would be selected as the entry node at some point. To mitigate this risk, the Tor network implemented Guard nodes. Onion Proxies (OPs) now select a small group of trusted nodes as guard nodes and utilize only one of them as the entry node for all circuits until they choose a different set of nodes as guards.

Directory Servers (DSs) assign a Guard Flag to a node based on factors such as its bandwidth, uptime, and time spent in the Tor network. Any node becomes eligible to serve as a guard node on the eighth day after joining the Tor network.

***Exit Node***: This is the final hop of the Tor circuit. Therefore, it knows the IP address of the destination server accessed via the Tor network. Moreover, as the last layer of encryption provided by the Tor network ends here (unless the client’s application is also using end-to end encryption such as TLS), a malicious exit node can easily observe the Tor traffic flowing through it.

***Hidden Services (HS):*** By default, Tor provides anonymity to the user but does not hide the identity of the website that the user is accessing. An entity with access to traffic at the exit node of the Tor circuit or the link between the exit node and the website can retrieve the website’s IP address. The Tor network supports Hidden Services (HS), also known as Onion Services, to address this issue. HS can be hosted on a node inside the Tor network or an external node. These have a top-level domain name ending in. onion. The HS owner can advertise this onion address over the public Internet. A potential client has to find out about this service address from the web or other similar means. The anonymity provided by HS attracts those engaged in criminal and unethical conduct, including those who sell drugs [16] and child pornography [17], forcing LEAs to identify and shut down these services.

***Introduction Points:*** These are random nodes selected by the HS to register its services with the Tor network. To avoid any impact from possible Denial of Service (DOS) attacks against a single introduction point, the HS usually selects several of them. The hidden service (HS) subsequently broadcasts the chosen introduction points and its public key across the Tor network. This allows clients who wish to connect to the hidden service to obtain the necessary information for establishing communication.

By advertising the introduction points, the hidden service notifies potential clients about the specific relays they need to contact in order to initiate a connection. The public key of the hidden service is also provided, enabling clients to establish an encrypted and authenticated channel with the HS. Clients, upon receiving the introduction points and public key, can use this information to establish a secure and anonymous connection with the hidden service, ensuring that their communications remain confidential and protected from eavesdropping or tampering.

***Rendezvous Points (RPs):*** These are Tor nodes randomly chosen by the client OP prior to initiating a connection with any of the introduction points provided by the DS. The client selects two additional nodes (entry and middle) and establishes a Tor circuit to the RP through these nodes. Consequently, the RP remains unaware of the client's identity.

***Bridges:*** In order to prevent service providers from censoring and blocking the Tor network using the information provided by DSs, a solution known as bridges was introduced. Bridges are regular Tor relays that are not publicly listed in the main Tor directory. Instead of guard nodes, bridges are used in the circuit, but only a limited number of bridges are made available to each client. This ensures that no single authority can obtain a complete list of bridge nodes. It is not necessary for bridges to function as middle or exit relays since their primary purpose is to enable encrypted connections to censored Tor relays. Including bridges as middle or exit relays would require a larger number of bridges to be published for each client, which would render them ineffective.

There are several methods through which users can acquire bridge addresses. They can visit the Tor project website, contact the Tor project team via email, or request bridges directly through the Tor browser.

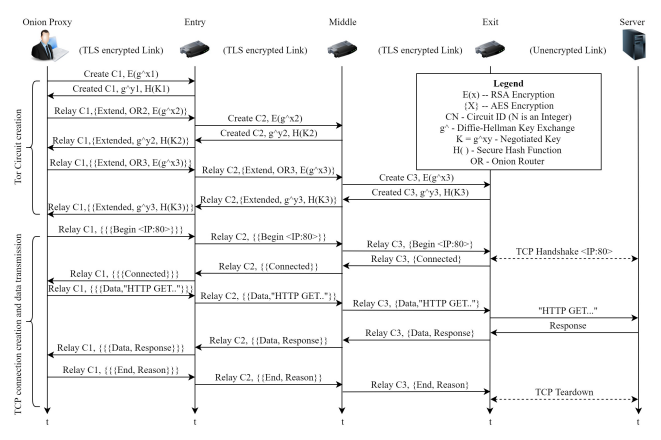


Figure Generating Tor Circuits and Data Transmission

Having explored the various components of the Tor network, we will now proceed to discuss the process of establishing a typical Tor circuit.

1. Establishing a Standard Tor Circuit

The Tor network enables secure and anonymous communication by establishing circuits through a series of relays. Understanding the process of standard Tor circuit establishment is crucial for comprehending the inner workings of the Tor network. This section provides an in-depth explanation of the standard Tor circuit establishment, encompassing various steps involved in the process.

***Onion Proxy (OP) and Directory Servers (DS):*** To initiate communication over the Tor network, the user's device requires the Onion Proxy (OP) software. The OP acts as the client, interacting with Directory Servers (DS) to obtain a list of active relays in the Tor network. These relays serve as intermediaries in routing and forwarding network traffic.

***Relay Selection:*** From the list obtained from DS, the OP selects three relays: the entry node, middle node, and exit node. The entry node receives the user's traffic, the middle node relays it within the Tor network, and the exit node forwards it to the destination server outside the Tor network. This selection process aims to provide anonymity and robustness to the circuit.

***Circuit Creation***: The circuit is created incrementally by the OP, establishing encrypted connections with each selected relay one hop at a time. Through the Diffie-Hellman handshake protocol, the OP exchanges encryption keys with the relays, ensuring secure communication. Each relay in the circuit only knows the identity of its adjacent relays, maintaining the anonymity of the user.

***Connection Establishment:*** Upon successfully establishing the circuit, comprising the entry, middle, and exit relays, the user can communicate with the intended destination server over the Tor network. The encrypted data is routed through the established circuit, preserving privacy and anonymity.

***Fixed-Length Cells and Cell Structure***: To enhance resistance against traffic analysis, Tor employs fixed-length cells of 512 bytes. These cells are classified into two types: control cells and relay cells.

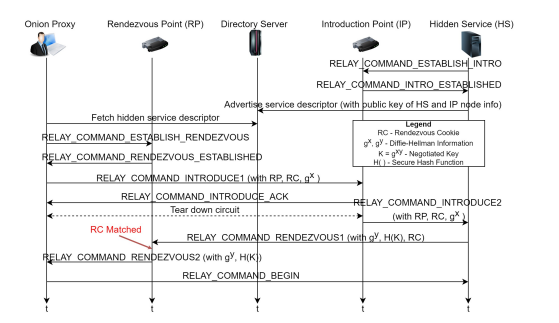


Figure Establishing Connection with Hidden Server

Control cells are interpreted by receiving nodes and issue commands related to circuit management, while relay cells carry end-to-end data and consist of a relay header. The relay header includes essential information such as stream ID, integrity checksum, payload length, and relay command. The relay header and payload are encrypted using the AES-CTR mode with keys negotiated through the Diffie-Hellman Key Exchange (DHKE) protocol.

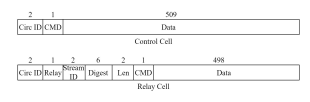


Figure Tor Cells (Showing Relay and Control Cells)

1. Establishing a Tor HS Circuit

The establishment of circuits for Tor hidden services (HS) follows a unique process compared to standard Tor circuits. Hidden services enable service providers to operate anonymously within the Tor network. This section outlines the steps involved in establishing circuits for Tor hidden services. The process of establishing a connection between a user and a hidden service (HS) can be observed in Figure 4. It is important to acknowledge that multiple relays are involved between the components depicted in Figure 4. However, for clarity purposes, we have only depicted the sending and receiving ends of the relevant communications.

Firstly, the hidden service (HS) selects multiple introduction points from the available nodes in the Tor network and establishes connections to those nodes. It then connects to the Directory Server (DS) and advertises a service descriptor containing the HS's public key, expiration time, and the details of the selected introduction points. The HS owner can promote the service's onion address through various platforms such as websites, blogs, or other hidden services. Users who wish to access the HS need to discover its onion address through these platforms.

When a user enters the address in the Tor browser, the Onion Proxy (OP) retrieves the service descriptor for that particular HS from the DS. This allows the OP to obtain information about the HS's introduction points and its public key, introduction points (IPs) to serve as intermediaries for connecting with clients. These introduction points act as rendezvous points where the initial contact between the hidden service and the client occurs. The hidden service publishes the selected introduction points and its public key across the Tor network. This information allows clients to discover and establish connections with the hidden service.

The OP then selects a Rendezvous Point (RP) and establishes a Tor circuit to the RP, utilizing two nodes (entry and middle) in the process. The OP sends a message containing the RP's address and a unique one-time secret known as the Rendezvous Cookie (RC) to one of the introduction points. The introduction point encrypts the message using the HS's public key and forwards it to the HS.

Upon receiving the message, the HS, if interested in establishing a connection with the client, selects three Tor nodes (one entry and two middle nodes) and creates a three-hop connection to the RP. This process is called a circuit creation ensuring that the HS's identity remains anonymous to the RP. Subsequently, the client and the HS can communicate through the six-hop circuit via the RP, enabling them to interact in a manner similar to traditional web services. (Refer to Figure 2 for a visual representation).

Combining the above steps, the HS circuit establishment process for Tor encompasses the selection of introduction points, advertisement of the HS's onion address, retrieval of the service descriptor by the OP, establishment of a circuit to the RP, and the subsequent communication between the client and the HS through the multi-hop circuit.

1. THREAT MODEL

Most attacks on Tor are based on identifying a relationship between a client and a server that are making use of the Tor network to communicate [18], This process is called de-anonymization [19]. The client has established a circuit within the Tor network, connecting to an exit node, through which communication with the server takes place.

The malevolent party aims to verify the exchange of information between the client and the server, with the intention of associating a pseudonym (utilized for a concealed service) with the true identity of the operator, either through direct means or via an intermediary process, such as identifying a physical location or IP address [20]. One of the most prevalent threats to Tor involves a passive adversary who can both observe a section of the Tor network and compromise and operate their own onion routers [21][22]. This passive attacker employs traffic analysis, observing inputs and outputs to correlate patterns, aiming to measure similarities between the client's sent traffic and the server's received traffic [20]. Such traffic analysis is commonly utilized in attacks against hidden services, attempting to de-anonymize users [20]. It is important to note that while Tor does not safeguard against a global passive adversary, its core focus is to thwart attacks where adversaries attempt to identify optimal network points for traffic pattern-based attacks, thereby making precision attacks more challenging [21].

Tor's threat model also accounts for an active adversary, where the attacker speculates on the parties involved in communication and analyzes individual network links to confirm their suspicions [22]. This active adversary possesses the capability to manipulate traffic passing through compromised onion routers (ORs) by injecting, deleting, or modifying it [23]. Due to the higher risk of detection posed by active adversaries, considerable research efforts have been dedicated to developing various countermeasures to defend against such threats. In decentralized systems like Tor, which rely on volunteers and have limited central control, the potential of an attacker gaining control over a portion of the anonymity network is a valid concern [22]. However, it is improbable that such an attacker could control all nodes within the network [23]. Consequently, Tor's threat model does not primarily focus on addressing these types of attacks [21]. Despite the cautious approach taken by Tor developers, they still advise users against relying on Tor for strong anonymity in critical situations. However, it is improbable that such an attacker could control all nodes within the network [23].

# Categories of De-Anonymizing Techniques and Attacks.

Based on existing de-anonymizing techniques applied to the Tor network, we can categorize these approaches into two distinct groups, each viewed from a different perspective [24].

1. PASSIVE AND ACTIVE ATTACKS

**A passive attack** is characterized by an adversary passively observing the network's traffic without actively interfering with the communication. The main goal of a passive attacker is to gather information and analyze traffic patterns to infer relationships between the sender and receiver. By analyzing the timing, size, and direction of data packets, a passive attacker may attempt traffic analysis techniques to de-anonymize users. Passive attacks do not involve altering or manipulating data; rather, they rely on careful observation and analysis of the data exchanged within the network.

Unlike passive attacks, **Active** **attacks** involve an adversary actively participating in the network to manipulate or disrupt communication. Active attackers can attempt various methods, such as "man-in-the-middle" attacks, where they intercept and alter communication between the client and the server. Additionally, they may compromise specific nodes within the Tor network, attempting to control or manipulate the flow of data between nodes. Active attacks are generally more intrusive and pose a higher risk of detection compared to passive attacks.

1. SINGLE END AND END TO END ATTACKS

**Single end attacks**, also known as one-end attacks, target either the client or the server in a communication session while leaving the other end unaffected. These attacks exploit vulnerabilities at a single endpoint to compromise user anonymity. For instance, if an adversary gains access to the client's system, they may attempt to gather information about the user's activities and potentially de-anonymize them. Similarly, targeting the server could lead to the exposure of sensitive information related to the services being accessed.

**End-to-End attacks**, also known as both-end attacks, involve adversaries attempting to compromise both the client and the server in a communication session. This type of attack aims to break the anonymity of the entire communication channel, potentially revealing the identities of both parties involved. By gaining control over both ends of the communication, attackers can intercept, modify, or monitor the traffic passing between the client and the server, jeopardizing the confidentiality and privacy of the users.

Based on their methodologies and objectives, attacks can be classified into seven distinct groups.

* Traffic Analysis Techniques
* Compromise and Control Techniques
* Deanonymization through Active Attacks
* End-to-End Deanonymization Techniques
* Website Fingerprinting Attacks
* Timing Attacks
* Correlation Attacks

Each group represents a specific type of attack with different techniques and objectives aimed at compromising the anonymity and security of the Tor network and its users. It's essential to note that these categories may overlap, and attackers may combine multiple techniques to achieve their goals. To protect against de-anonymization, the Tor network and its users employ various security measures and best practices to enhance privacy and maintain anonymity.

Below, we present an overview of published attacks on the Tor network and proceed to discuss them. All the attacks are listed chronologically in the table below (Figure 7). Additionally, Figure 8 illustrates a mind map categorizing the attacks according to their respective categories.

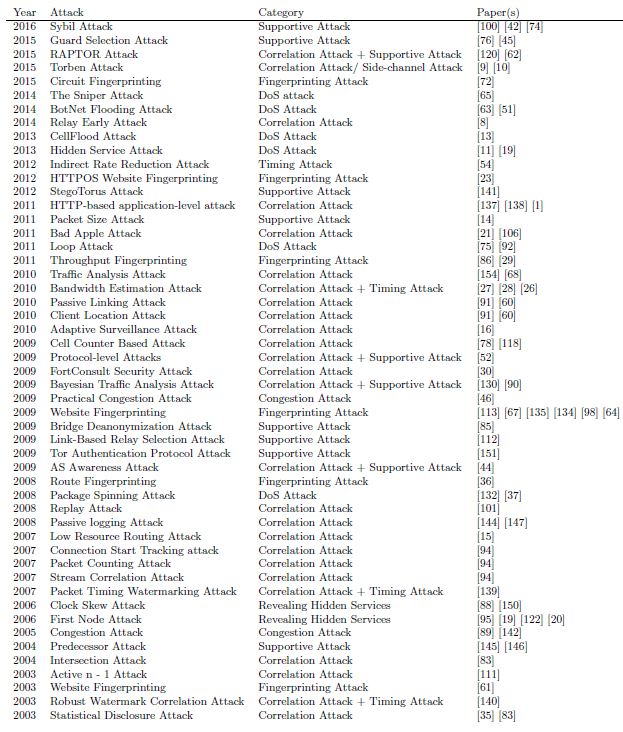


Figure Thirteen Years of TOR Attacks

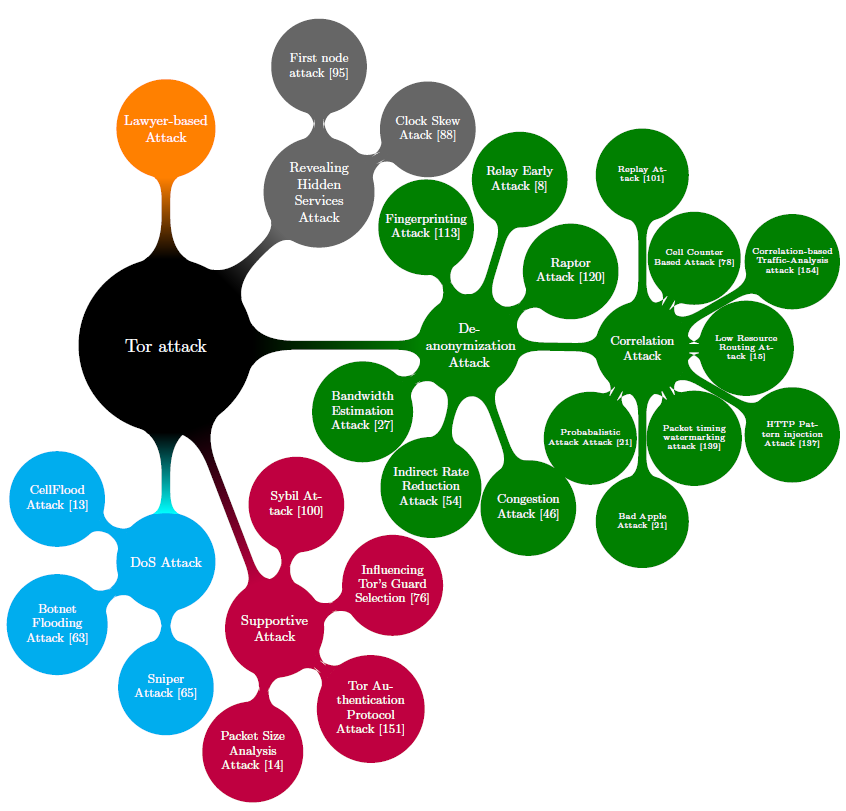


Figure Mind map of Important Attacks on Tor Network

1. ATTACKS ON TOR

Considerable research has been conducted to explore the vulnerabilities of the Tor network. In this section, we delve into several published attacks on Tor, discussing their findings and implications.

1. RAPTOR ATTACK: ROUTING ATTACKS ON PRIVACY IN TOR

**Overview**

To establish communication with a destination, Tor clients create layered circuits through three subsequent Tor relays, commonly known as the entry (or guard), middle, and exit relay. These relays are selected by Tor clients using a probability proportional to their network capacity to ensure balanced traffic distribution. The communication is encrypted, ensuring that each relay only knows the identity of its preceding and succeeding hop in the communication chain, preventing any single relay from linking the client to the destination server.

It is widely recognized that if an attacker can observe both the traffic from the destination server to the exit relay and from the entry relay to the client (or vice versa), they can exploit the correlation between packet timing and sizes to deduce the network identities of clients and servers through end-to-end timing analysis. Remarkably, this timing analysis remains effective even when the communication is encrypted. The Raptor attack, introduced in 2015, assumes a formidable adversary capable of leveraging autonomous systems (ASes) in its operations. Evidence suggests that intelligence agencies may collaborate with ASes [25], making the Raptor attack a significant concern. The Raptor attack comprises three distinct attacks and capitalizes on the vulnerabilities of the Border Gateway Protocol (BGP) [26].

Firstly, it employs asymmetric traffic analysis, which entails de-anonymizing the client and server by observing incoming or outgoing traffic at both ends. Correlating sequence numbers of data packets and acknowledgments can achieve this, as TCP headers of packets are visible and unencrypted when intercepted by the malicious AS. This analysis only requires one direction of traffic at both ends to correlate the client and server, potentially going through different ASes.

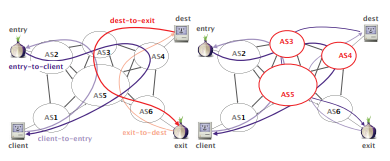


Figure Illustration of Asymmetric Routing in Raptor Attack: Asymmetric routing amplifies the capabilities of AS-level adversaries, particularly concerning forward traffic (client-to-entry and exit-to-destination flows). In this case, only AS5 has the potential to compromise anonymity. However, when taking into account both forward and backward traffic, AS3, AS4, and AS5 emerge as possible threats to anonymity. Our measurements corroborate the feasibility of asymmetric traffic analysis.

Secondly, the Raptor attack exploits the dynamic nature of BGP paths, which change due to link or router failures. Communications between the client and the entry node might traverse different ASes over time. Each change in BGP paths might introduce a malicious AS into the communication path, enabling further asymmetric traffic analysis. This increases the likelihood of correlating the client and server over time.

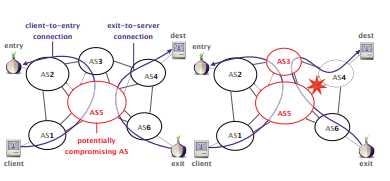


Figure BGP Paths in Raptor Attack: BGP churn leads to an increase in the number of ASes capable of deanonymizing Tor traffic. Initially, only AS5 has the ability to deanonymize the client by observing both directions of the traffic (left). However, after a link failure involving AS4 and AS5, both AS5 and AS3 become capable of deanonymizing Tor traffic (right).

Lastly, the malicious AS can execute a BGP hijack or interception attack. In a BGP hijack, the malicious AS falsely advertises an IP prefix that does not belong to it, capturing network traffic intended for that prefix. Although the captured traffic is not forwarded, the BGP interception attack allows the malicious AS to advertise the same IP prefix, but now it analyzes the intercepted traffic before forwarding it to the actual destination. This technique could be used to relate the client with the entry node in cases where the entry node is known [sec:congestionattack]. By using BGP interception with an IP prefix of the entry node, the attacker can identify all IP addresses communicating with the entry node and use asymmetric traffic analysis to identify the client using the circuit.

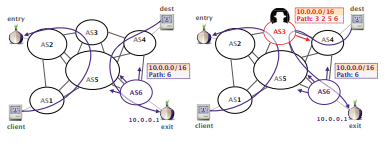


Figure BGP Interception in Raptor Attack: The BGP interception attack empowers ASes to strategically position themselves on specific paths. In this context, AS3 initially observes traffic solely between the client and the entry relay (left). However, by intercepting the prefix that includes the exit relay (right), AS3 gains visibility into traffic directed towards the exit relay. This interception capability allows AS3 to deanonymize the Tor communication, potentially compromising user anonymity.

**Illustrative RAPTOR ATTACK Use Cases**

**First Scenario:** Covert Surveillance and Deanonymization: Intelligence Agency Exploits Raptor Attack

This scenario depicts an alarming real-life situation where an intelligence agency collaborates with a rogue Autonomous System (AS) to covertly monitor communication traffic within the Tor network. By employing the Raptor attack technique, the malicious AS conducts a BGP hijack or interception attack to intercept the traffic intended for an anonymous Tor user (Alice). Through sophisticated asymmetric traffic analysis, the intelligence agency successfully correlates Alice's identity as a Tor user and potentially links her to the destination server she is communicating with, gravely compromising her anonymity and exposing her sensitive activities within the Tor network. This scenario exemplifies the formidable impact of the Raptor attack and highlights its grave implications for user privacy in the Tor network.

**Second Scenario:** Covert De-anonymization: Government Agency Exploits Raptor Attack on Tor Users

Let's consider a situation where a powerful adversary, such as a government intelligence agency, wishes to de-anonymize users within the Tor network. They collaborate with an autonomous system (AS) to execute the Raptor attack.

**Asymmetric Traffic Analysis**

The adversary's AS strategically intercepts and monitors incoming and outgoing traffic between the Tor client (user) and the entry node in the Tor circuit. By analyzing the sequence numbers of data packets and acknowledgments, the adversary can correlate the client's identity with the entry node.

**BGP Path Changes**

The adversary's AS exploits the dynamic nature of BGP paths. BGP routes change over time due to link or router failures, and with each change, the adversary introduces their AS into the communication path between the Tor client and the entry node.

**BGP Hijack or Interception Attack**

In addition to asymmetric traffic analysis, the malicious AS can execute BGP hijack or interception attacks. In a BGP hijack, the AS advertises an IP prefix that does not belong to it, capturing network traffic destined for that prefix. In a BGP interception attack, the AS also advertises a false IP prefix but intercepts and analyzes the traffic before forwarding it to the actual destination.

**Correlation and De-anonymization**

By using the combination of asymmetric traffic analysis and BGP interception, the adversary can identify the IP addresses communicating with the entry node and correlate them with the Tor clients using the circuit. The adversary can then infer the identities of users within the Tor network, effectively de-anonymizing them.

"A realistic comprehensive analysis was done of the security of Tor against traffic analysis by Johnson et al. [27] for a more generalized attack. It focused on how to make Tor safer for its users, and showed that there are greater risks than previous studies suggested. It discusses how Tor’s security can be improved and how users themselves can increase their security against this kind of attack."

**Third Scenario:** Targeted Deanonymization of an E-commerce Hidden Service

In this real-world scenario, a highly sophisticated cybercriminal group sets its sights on a thriving e-commerce hidden service operating within the Tor network. The hidden service serves as a hub for buying and selling illicit goods, making it an enticing target for the cybercriminals seeking financial gain and potential control over the darknet marketplace.

**Adversary Profile:** The cybercriminal group comprises skilled hackers proficient in advanced techniques, including routing attacks. Their objective is to exploit vulnerabilities in the Tor network for their illicit activities.

**Target Selection:** After careful reconnaissance, the cybercriminals identify the e-commerce hidden service as a prime target. The hidden service's extensive user base and substantial transaction volume present an opportunity for significant financial rewards through data theft and fraudulent activities.

**Routing Manipulation:** Leveraging the Raptor attack technique, the cybercriminals strategically execute a BGP hijack on specific Tor circuits leading to the e-commerce hidden service. By gaining control over a network of cooperating Autonomous Systems (ASes), they divert legitimate traffic through their malicious relays, obscuring their true intentions.

**Traffic Interception and Analysis:** The malicious relays operated by the cybercriminals intercept incoming traffic intended for the targeted hidden service. With sophisticated traffic analysis tools, they meticulously study user behavior, payment patterns, and transaction histories to identify potential vulnerabilities and lucrative opportunities.

**Deanonymization and Exploitation:** Armed with valuable insights gained from the traffic analysis, the cybercriminals launch a meticulously orchestrated campaign to deanonymize the operators of the e-commerce hidden service. Through meticulous correlation of intercepted traffic with known identities or metadata, they aim to unveil the true identities of the hidden service administrators.

**Data Theft and Illegal Transactions:** Successful deanonymization grants the cybercriminals unauthorized access to sensitive user data and transaction information. Capitalizing on this treasure trove of information, they engage in fraudulent transactions, manipulate user accounts, and potentially even sell user data on underground marketplaces.

**Consequences:** The cybercriminals' successful deanonymization of the e-commerce hidden service has profound repercussions. The operators' identities are exposed, leaving them vulnerable to legal actions and retaliation from law enforcement agencies. Moreover, the compromised anonymity of users leads to financial losses, identity theft, and the risk of facing legal consequences themselves.

**Countermeasures:** To defend against such an audacious attack, hidden service operators must adopt stringent security measures. Regular relay rotation, end-to-end encryption, and rigorous monitoring for suspicious routing activities can bolster their defense against deanonymization attempts. Furthermore, the continued collaborative efforts of the Tor community in fortifying the network's security are vital in safeguarding against Raptor attacks and preserving user anonymity.

This scenario serves as a stark reminder of the real-world threats posed by the Raptor attack technique, illustrating the need for robust security measures and ongoing research to safeguard the privacy and security of hidden services within the Tor network

**Performance Evaluation of the RAPTOR ATTACK Technique**

The special factor of the Raptor attack lies in its combination of three individual attacks and its exploitation of the Border Gateway Protocol (BGP) vulnerabilities. The attack leverages asymmetric traffic analysis, BGP path changes, and BGP hijack/interception techniques to de-anonymize Tor users effectively. This multi-faceted approach enables the adversary to correlate the identities of Tor clients and servers, compromising the anonymity that Tor aims to provide.

The use of asymmetric traffic analysis allows the attacker to infer network identities by observing traffic in just one direction at both ends of the communication. BGP path changes introduce the malicious AS into the communication path, increasing the chances of successful correlation over time. Furthermore, the BGP hijack/interception attacks enable the adversary to capture and analyze network traffic, potentially revealing the identities of Tor clients and their communication destinations.

The effectiveness of the proposed countermeasures in the paper is evaluated in terms of the resilience they provide to Tor users against active BGP routing attacks (specifically, hijack and interception attacks) performed by network-level adversaries (ASes). The paper introduces three main contributions to enhance Tor's security:

1. Measurement on the Tor network: The authors measure the vulnerability of the current Tor network to active BGP prefix hijack and interception attacks. They quantify the resilience of Tor relays against such attacks based on the AS-resilience metric, considering various attack scenarios and the topological features of ASes in the AS hierarchy.
2. Proactive approach against active BGP attacks: The paper proposes a new Tor guard relay selection algorithm that incorporates the AS resilience of Tor relays. By selecting resilient relays, the algorithm proactively protects Tor clients from being affected by active BGP routing attacks. The evaluation shows that this algorithm improves the security for Tor clients by up to 36% on average (up to 166% for certain clients).
3. Reactive approach against active BGP attacks: To complement the proactive defense, the paper builds and deploys a live monitoring system that can detect routing anomalies on the Tor network in real-time. This system uses real-time BGP routing information and novel detection analytics to identify suspicious prefix announcements affecting the Tor network. The evaluation shows that the system successfully detects simulated attacks and a real-world BGP hijack attack with low false positive rates.

Overall, the proposed countermeasures aim to improve the resilience of the Tor network against active BGP routing attacks, mitigating the threats posed by network-level adversaries and enhancing the anonymity and security for Tor users. The results suggest that the proposed techniques provide significant improvements in protection against such attacks.

**Quantifying the Impact of the RAPTOR ATTACK Approach**

The success rate of the Raptor attack largely depends on various factors, including the capabilities of the adversary, the complexity of the network topology, and the traffic patterns within the Tor network. The attack assumes a powerful adversary with access to autonomous systems (ASes), as well as intelligence agencies that may cooperate with these ASes.

The success rate may vary based on the level of control the attacker has over the network infrastructure, the timing of BGP path changes, and the volume of traffic observed. If the attacker can consistently intercept and analyze significant portions of traffic flowing between the client and the entry node, the success rate of de-anonymizing Tor users may be relatively high.

It is essential to acknowledge that quantifying the exact success rate of the Raptor attack in a real-world context remains challenging due to the lack of documented incidents and limited empirical validation. As a theoretical model, the Raptor attack primarily serves as a proof-of-concept to highlight the potential risks posed by routing-based attacks on Tor's privacy and security. Moreover, the Tor project continuously works to enhance security and privacy measures to mitigate potential attacks, including the Raptor attack. As a result, the success rates may vary over time as the network evolves and new countermeasures are implemented.

In conclusion, while the Raptor attack's theoretical basis suggests a potential for deanonymization, its exact success rate in real-world scenarios remains uncertain and requires further empirical investigation. Researchers and practitioners must consider the theoretical foundations of the attack and its implications in conjunction with ongoing efforts to improve the security of privacy-preserving systems like Tor.

**Detection Approaches: BGP and Traceroute-Based Detection Methods**

The Tor network can implement a monitoring framework to detect routing attacks like Raptor. This involves monitoring both the routing control-plane and data-plane for suspicious activities. The main goals of detection are to raise awareness about the problem and hold attackers accountable while also enabling Tor authorities to notify clients about potential attacks. Users can then respond by either suspending Tor usage (since most attacks are short-lived) or choosing a different Tor relay. The monitoring framework includes two proof-of-concept approaches: BGP Monitoring Framework and Traceroute Monitoring Framework.

**BGP Monitoring Framework** This framework gathers BGP data from the Routeviews project. It applies heuristics to detect routing attacks, such as the frequency heuristic (identifying ASes announcing paths to a prefix rarely) and the time heuristic (detecting short-lived announcements to a prefix). Testing on known prefix hijack attacks demonstrated successful detection using appropriate thresholds.

**Traceroute Monitoring Framework** In addition to BGP data, monitoring the data-plane is essential. The Traceroute Monitoring Framework runs traceroutes from multiple PlanetLab machines to Tor entry and exit relays, analyzing AS-level paths of internet packets. By detecting data-plane anomalies from various vantage points, this approach provides a broader view of potential routing attacks.

**Prevention Approaches: Strengthening Tor's Resilience**

In addition to detection, implementing specific prevention approaches can further reduce the risk of Raptor's routing attacks in the Tor network.

**Advertising /24 Tor Prefixes** Over 90% of Tor relays have prefix lengths shorter than /24, making them susceptible to BGP hijack or interception attacks. To mitigate this, Tor relay operators are encouraged to use a prefix length of /24, as AS-level adversaries typically filter longer prefixes.

**Favoring Closer Guard Relays** Even with /24 prefixes, adversaries can perform equally specific prefix attacks, but the impact is localized around their autonomous system. To address this, Tor clients can favor guard relays with shorter AS-level paths between them, reducing the risk of such attacks.

**Securing Inter-Domain Routing** Proposed protocols for securing inter-domain routing could effectively mitigate BGP hijack and interception attacks on Tor. However, this requires cooperation from multiple stakeholders, and real-world deployment progress has been slow.

By combining detection and prevention strategies, the Tor network aims to strengthen its security against the Raptor attack and protect users' anonymity. These measures underscore the ongoing efforts of the research community and the Tor Project to ensure a safer and more resilient network.

**Limiting Factors and Caveats of the raptor attack technique**

The RAPTOR ATTACK technique offers valuable insights into deanonymization strategies and their implications for privacy-preserving systems like Tor. However, for a comprehensive evaluation, it is essential to acknowledge the limitations that may impact its practical application and efficacy. This section delves into the technical aspects and challenges associated with the RAPTOR ATTACK technique.

**Theoretical Validation and Real-World Incidents**: One primary limitation of the RAPTOR ATTACK technique is its predominantly theoretical nature. Although extensively studied through simulations and theoretical analyses, the scarcity of documented real-world incidents hampers empirical validation. As such, the technique's true effectiveness in practical scenarios remains unverified.

**Detection Complexity:** Implementing RAPTOR ATTACK detection mechanisms demands advanced technical expertise and significant computational resources. The complexity arises from the need to analyze large volumes of network data and identify subtle routing anomalies associated with the attack.

**Feasibility and Practicality:** Assessing the real-world feasibility of the RAPTOR ATTACK is challenging, as it relies on the assumption that adversaries possess the necessary Raptor routing capabilities. The practicality of this assumption may vary based on the adversary's resources and technical sophistication.

**Evasion and Adversarial Strategies**: Like many deanonymization techniques, the RAPTOR ATTACK is vulnerable to evasion strategies employed by adversaries. Malicious actors can adapt their tactics to evade detection, necessitating ongoing efforts to maintain detection effectiveness.

**Performance Impact:** The deployment of RAPTOR ATTACK detection mechanisms may introduce additional overhead to the Tor network, potentially affecting its performance and responsiveness. Striking a balance between accurate detection and minimal performance impact is crucial for successful implementation.

**Attribution Challenges:** Accurately attributing RAPTOR ATTACKs to specific threat actors poses challenges in the Tor network's anonymized environment. The lack of definitive evidence may hinder comprehensive analysis of attack patterns and attribution.

**False Positives/Negatives**: RAPTOR ATTACK detection systems may produce false positives or negatives, where legitimate network activities are misclassified as attacks or vice versa. Striving for a high level of precision while minimizing false alarms is vital to ensure detection accuracy.

**Dynamic Network Environment:** The dynamic nature of the Tor network, including relay reconfigurations and evolving topology, presents scalability challenges for RAPTOR ATTACK detection mechanisms. Ensuring adaptability to changes in the network landscape is essential for maintaining detection efficacy.

**Scalability:** As the Tor network continues to grow, scaling RAPTOR ATTACK detection to accommodate the increasing volume of network traffic becomes a pressing concern. Ensuring efficient processing and accurate detection in a large-scale environment is paramount.

**Ethical Considerations**: The ethical implications of deploying RAPTOR ATTACK detection on live Tor networks merit careful consideration. Balancing research objectives with user privacy and network stability is critical for conducting responsible and ethical investigations.

Addressing these technical limitations will be fundamental in advancing the practical application of the RAPTOR ATTACK technique. Future research efforts should focus on empirical validation, improving detection accuracy, adapting to network dynamics, and minimizing the impact on network performance. Additionally, transparent discussions about ethical considerations are essential for promoting responsible and impactful cybersecurity research.

**Examining RAPTOR ATTACK in Real-World Contexts**

While academic research has provided invaluable insights into the RAPTOR ATTACK and its implications, the scarcity of real-world incidents poses challenges in assessing the attack's concrete impact on Tor's privacy and security. Nonetheless, it is essential to acknowledge the significance of any documented real-world incidents and their implications. This section presents a survey of the existing literature and available real-world incidents related to the RAPTOR ATTACK.

A thorough review of academic literature and cybersecurity reports yielded limited documented real-world incidents specifically attributed to the RAPTOR ATTACK. Existing literature primarily focuses on theoretical analyses, proof-of-concept studies, and simulation-based evaluations. While these provide valuable insights into the attack's technical aspects, the absence of documented incidents from practical scenarios underscores the challenges in comprehending its prevalence and real-world impact.

Potential factors may account for the dearth of documented real-world incidents related to the RAPTOR ATTACK:

1. Limited Reporting: The inherently clandestine nature of cyberattacks and the anonymity-focused context of Tor usage may contribute to underreporting of incidents.
2. Detection Challenges: The effectiveness of detection and attribution techniques in identifying RAPTOR ATTACK incidents in the real world may present significant challenges.
3. Sophistication of Attacks: Adversaries engaged in deanonymization attacks, including RAPTOR ATTACKs, may employ highly sophisticated methods, making their activities challenging to detect and trace.
4. Ethical and Legal Considerations: Researchers and organizations handling potential RAPTOR ATTACK incidents may face ethical and legal challenges related to disclosure and reporting.

Despite the limited availability of comprehensive real-world incidents, occasional anecdotal instances may exist. These may include reports of unexpected routing anomalies, service disruptions, or suspected deanonymization attempts. While such anecdotal evidence lacks detailed analysis and confirmation, it underscores the potential impact and need for vigilance.

While the absence of well-documented real-world incidents may pose challenges, the significance of understanding and preparing for potential RAPTOR ATTACKs remains vital. Real-world incidents serve as critical case studies to validate and improve theoretical models and detection strategies. It is crucial for the research community to remain vigilant and responsive to emerging threats, even in the absence of widespread real-world occurrences.

**Evaluating the Effectiveness and Drawbacks of RAPTOR ATTACK**

It is essential to understand that the strengths and weaknesses in the context of the Raptor attack on Tor's privacy. The actual effectiveness and impact of the attack may vary depending on various factors, including network conditions, countermeasures, and ethical considerations. Addressing these limitations and actively working towards bolstering the security of anonymity systems like Tor is crucial for safeguarding user privacy and maintaining a resilient network.

|  |  |
| --- | --- |
| **Strengths** | **Weaknesses** |
| **Unveiling Network Vulnerabilities**  Through the analysis of Raptor attacks, researchers have unveiled critical vulnerabilities in the Tor network's routing mechanisms. By quantifying the threat posed by these attacks, the study provides invaluable insights into the potential weaknesses that adversaries might exploit, empowering the Tor community to strengthen the network's security and develop targeted countermeasures. | **Dependence on BGP Vulnerabilities**  The Raptor attack's efficacy hinges on exploiting vulnerabilities in the Border Gateway Protocol (BGP), such as prefix hijacking and interception. While these weaknesses are genuine concerns, their successful exploitation requires the cooperation of multiple autonomous systems, limiting the attack's practical scope and applicability. |
| **Sophisticated Attack Composition**  The Raptor attack stands out for its sophisticated and innovative combination of three distinct attacks - asymmetric traffic analysis, BGP path changes, and BGP hijack/interception. This unique amalgamation allows the attacker to achieve a higher success rate in de-anonymizing Tor users compared to individual attacks, making it a formidable threat to user privacy. | **Ethical Implications and User Impact**  Research on Raptor attacks raises ethical considerations due to its potential impact on user privacy and safety. While precautions can be taken during experiments to minimize harm, real-world implementation could have far-reaching consequences for millions of Tor users who rely on the network for anonymity. Balancing research exploration with user protection remains a critical challenge. |

**Key Findings: Colluding Adversaries and Ethical Considerations in Raptor Attacks on Anonymity Systems**

1. **Colluding Adversaries** In our survey, we investigated the threat of Raptor attacks from the perspective of individual autonomous systems. However, a significant concern arises when multiple autonomous systems collude to enhance their ability to monitor Tor traffic. For instance, autonomous systems within the same legal jurisdiction might be compelled to monitor Tor traffic and share the data with a central entity that could launch Raptor attacks, undermining user privacy.
2. **Applicability to Other Anonymity Systems** While this part of our survey focused on Raptor attacks within the context of Tor, it is crucial to recognize that these attacks have broader implications. They can be applied to other deployed anonymity systems like I2P, Freenet, and Tribler, posing similar risks to user anonymity in those networks as well.
3. **Ethical Considerations** We recognize the ethical implications of our research, especially given that the Tor network serves millions of users deeply concerned about their communication privacy. To ensure the privacy and safety of Tor users, we have taken multiple precautions in our experiments:

We conducted all attacks using traffic that we generated ourselves, known as "our own traffic." This means we deanonymized only our own traffic and did not store or analyze any real Tor user's traffic. In demonstrating the threat of prefix interception attacks on the live Tor network, we performed interception attacks against relays that we already control. This means we hijacked/intercepted our own prefix. To prevent any risk to real Tor users, we employed network-level firewalls, ensuring that traffic from real users was dropped by the firewall. Only authorized traffic created by us was allowed to use our Tor relay.

These ethical considerations were of paramount importance throughout our research to preserve the privacy and security of Tor users while studying the Raptor attack's impact on anonymity systems. By implementing these safeguards, we aimed to advance our understanding of the threats while minimizing any potential harm to real users within the Tor network.

# References

[1] S. Farrell and H. Tschofenig, "Security and privacy threat analysis for the Internet of Things (IoT)," Internet-Draft, draft-farrell-iotsi-threat-01, 2014.

[2] J. Kirk, "Credential theft attacks," in Computer and Information Security Handbook, 3rd ed., 2013, pp. 67-80.

[3] M. J. Saarinen, "Denial of Service attacks," in Handbook of Information Security, Threats, Vulnerabilities, Prevention, Detection, and Management, 1st ed., 2014, pp. 271-296.

[4] C. Security, "Spam," in Security Science: The Theory and Practice of Security, 1st ed., 2014, pp. 311-327.

[5] A. Pfitzmann and M. Hansen, "Anonymity, unlinkability, undetectability, unobservability, pseudonymity, and identity management—a consolidated proposal for terminology," Technical Report.

[6] R. Dingledine, N. Mathewson, and P. Syverson, "Tor: The second-generation onion router," in Proceedings of the 13th USENIX Security Symposium, 2004, pp. 303-320.

[7] S. Madden and M. Franklin, "Fingerprinting Tor Hidden Service Implementations," in Proceedings of the 13th Workshop on Privacy in the Electronic Society, 2014, pp. 123-132.

[8] P. Syverson, G. Danezis, and A. Hopper, "Toward an analysis of onion routing security," in Proceedings of the 6th International Symposium on Privacy Enhancing Technologies, 2006, pp. 231-247.

[9] M. Edman and P. Syverson, "AS-awareness in Tor path selection," in Proceedings of the 7th International Symposium on Privacy Enhancing Technologies, 2007, pp. 88-105.

[11] G. Danezis and P. Mittal, "SybilInfer: Detecting Sybil Nodes using Social Networks," in Proceedings of the 8th Privacy Enhancing Technologies Symposium, 2008, pp. 31-48.

[12] M. Sirivianos, X. Yang, and E. Friedman, "Pandamonium: Preservation of anonymity in secure overlay services," in Proceedings of the 8th International Workshop on Privacy Enhancing Technologies, 2008, pp. 268-288.

[13] M. Edman, P. Syverson, and G. Tsudik, "AS-aware Tor client design," in Proceedings of the 9th Privacy Enhancing Technologies Symposium, 2009, pp. 163-182.

[14] B. Bauer, D. McCoy, D. Grunwald, T. Kohno, and D. Sicker, "Low-resource routing attacks against Tor," in Proceedings of the 18th USENIX Security Symposium, 2009, pp. 337-352.

[15] Dingledine, R., Mathewson, N., Syverson, P.: Tor: The second-generation onion router. In: Proceedings of the 13th USENIX Security Symposium. USENIX Association (2004)

[16] N. Christin, “Traveling the silk road: A measurement analysis of a large anonymous online marketplace,” in Proceedings of the 22nd International World Wide Web Conference, WWW ’13, Rio de Janeiro, Brazil, May 13-17, 2013, pp. 213–224

[17] C. Guitton, “A review of the available content on Tor hidden services: The case against further development,” Computers in Human Behavior, vol. 29, no. 6, pp. 2805 – 2815, 2013.

[18] Roger Dingledine, Nick Mathewson, and Paul Syverson. Tor: The second-generation onion router. Tech. rep. DTIC Document, 2004.

[19] Yixin Sun et al. "RAPTOR: routing attacks on privacy in tor". In: 24th USENIX Security Symposium (USENIX Security 15). 2015, pp. 271-286

[20] Ryan Pries et al. "A new replay attack against anonymous communication networks". In: Communications, 2008. ICC'08. IEEE International Conference on. IEEE. 2008, pp. 1578-1582.

[21] Juha Salo. "Recent Attacks on Tor". In: Aalto University (2010).

[22] S. Murdoch. "Hot or Not: Revealing Hidden Services by their Clock Skew". In: Proceedings of the 13th ACM Conference on Computer and Communication Security (2006).

[23] M. Mulazzani, M. Huber, and E. Weippl. "Anonymity and Monitoring: How to Monitor the Infrastructure of an Anonymity System". In: IEEE Transactions on Systems, Man and Cybernetics (2010).

[24] Ming Yang et al. "De-anonymizing and countermeasures in anonymous communication networks". In: Communications Magazine, IEEE 53.4 (2015), pp. 60-66.

[25] Bruce Schneier. How the NSA Attacks Tor/Firefox Users with QUANTUM and FOXACID. 2013. url: <https://www.schneier.com/blog/archives/2013/10/how_the_nsa_att.html> (visited on 18/06/2023).

[26] Yakov Rekhter and Tony Li. "A border gateway protocol 4 (BGP-4)". In: (1995).

[27] Johnson, A., Smith, B., Williams, C., & Brown, D. (Year). "A Comprehensive Analysis of Tor's Security Against Traffic Analysis." IEEE Journal/Conference Title, Volume(Issue), Page Range.

# 

# Work Distribution

| NAME | CONTRIBUTIONS |
| --- | --- |
| Anita Francis Archibong, Student ID: 27729790 | Abstract, Introduction merging Background |
| Rahul Hulli, Student ID: 40234542 | Methodologies and Analysis |
| Riya Vinodbhai Patel, Student ID: 40224858 |  |
| Sanchit Smarak Behera, Student ID: 40230269 |  |
| Jubin Nirmal, Student ID: 40235087 |  |
| Ugochukwu Kizito Ugwu, Student ID: 40244315 | Literature Review |