Impact Analysis of Sybil Attacks in the Tor Network

Christoph Sendner[1][0000−0003−3766−783X], Dominik Schreider[1], and Alexandra Dmitrienko[1][0000−0001−5637−7016]

Julius-Maximilians Universität, Würzburg, Germany
{firstname.lastname}@uni-wuerzburg.de

Abstract. Tor, the most widely used anonymity network with around 2 million users, relies on volunteer-operated relays to route traffic through circuits while preserving anonymity. However, it faces a significant challenge, the *diversity problem*, where power is concentrated among a few dominant relay operators. It increases the risk of Sybil attacks—where attackers introduce multiple relays without clear attribution, potentially enabling traffic monitoring and de-anonymization. This paper examines the diversity problem and Sybil attack implications through data aggregation and simulations, revealing significant vulnerabilities. Notably, the top 10 relay families control over 50% of exit traffic, and an attacker operating two separate relay families with 120 and 178 relays could deanonymize 0.8% of all circuits. Additionally, combining Sybil attacks with bandwidth inflation could allow an attacker to observe up to 13.44% of exit and 6.45% of entry traffic for just $400 per month. The study also highlights weaknesses in Tor's Sybil detection mechanisms, emphasizing the need for mitigation strategies to safeguard user privacy.

1 Introduction

Internet censorship affects individuals worldwide, with governments and ISPs blocking content for various reasons, including political control and security concerns [15]. The Russian invasion of Ukraine intensified such efforts, notably restricting access to the Tor Project to suppress independent reporting and promote state propaganda. Tor, an open-source anonymity network, enables users to bypass censorship and evade surveillance by routing traffic through volunteer-operated relays using onion routing [29]. However, its open nature makes it vulnerable to attacks aimed at de-anonymization, including website fingerprinting [3,14,6], correlation attacks [20,13,8,7,19], bandwidth inflation [25,12,32], and Sybil attacks [4,2]. Understanding these threats is crucial to preserving Tor's integrity and ensuring user anonymity.

Diversity Problem. Since Tor relays are volunteer-operated, some may be malicious. To prevent circuits from including relays controlled by the same entity, Tor enforces relay families, requiring operators to declare their relays as a group. However, the *diversity problem* [16,22] arises from resource concentration among

a few dominant operators, increasing their ability to monitor network traffic and raising the likelihood of multiple relays from the same operator being selected for a circuit.

Sybil Attacks. A Sybil attack occurs when a single entity controls multiple relays to manipulate Tor's network traffic [17]. Attackers can either declare their relays as a legitimate family (scenario 1) or obscure their affiliation (scenario 2), increasing the chance of controlling critical circuit positions and compromising anonymity. As a countermeasure, defenders can deploy "white hat" Sybils [40] to limit the influence of malicious relays (scenario 3).

Sybil Defenses. Sybil attacks are well-studied [17] but remain difficult to mitigate in trustless networks like Tor. Existing defenses [4] use social constraints and computational puzzles to deter multiple identities but fail against well-resourced adversaries, such as state actors. Sybilhunter [43], a detection tool used by the Tor project, analyzes network churn, uptime patterns, and relay configurations to identify Sybil relays. However, attackers can reintroduce excluded relays at no cost, maintaining Tor's vulnerability. This paper does not propose new defenses but instead investigates the diversity problem, the real-world impact of Sybil attacks, and weaknesses in Tor's current detection methods.

Our Goals and Contributions. Our goal is to improve the security of the Tor network by studying the diversity problem and the potential impact of Sybil attacks. In particular, our contributions are as follows:

- **Tool development.** We developed an Analysis Tool (AT) for the Tor Network, designed to gather critical information about the network's composition. Additionally, we created a Sybil Injection Tool (SIT) that enables us to introduce Sybils into Tor network simulations effectively.
- **Studying Diversity Problem.** To shed light on the diversity problem in Tor, we analyze the current Tor network composition using our Analysis Tool. We find a critical concentration of resources and power in a few big relay operators and define their share in the network. We find out that, collectively, the top 10 largest relay families have the ability to monitor over 50% of the exit traffic within the Tor network.
- **Impact on Anonymity.** We analyze the Sybil attack in Tor using simulations based on Shadow [10] and the Sybil Injection Tool we developed. Our analysis specifically focuses on the implications for users' anonymity and for the overall Tor network. Our findings unveil that the magnitude of the impact on user anonymity, coupled with the ease of execution, is substantial. For example, by strategically operating two separate relay families of 120 Guard and 178 Exit relays, an attacker could potentially deanonymize 0.8% of all circuits in Tor.
- **Amplification Impact for Bandwidth Inflation.** We delve into the combination of the Sybil attack with bandwidth inflation attacks and investigate how the effects of both can be substantially amplified. For instance, we show that an attacker investing around $400 a month for cloud hosting could observe up to 13.44% of exit and 6.45% of entry user traffic, severely undermining Tor's anonymity guarantees.

– **Analysis of Anti-Sybil Protection in Tor.** We analyze the effectiveness of Sybilhunter [43], which is currently employed for Sybil detection in the Tor network. Our results show that its mechanisms are insufficient to detect a well-prepared Sybil attack.

In summary, this work enhances Tor's security by analyzing network composition, simulating Sybil attacks, and exposing vulnerabilities. It reveals risks from the diversity problem, amplifies bandwidth inflation effects, and highlights weaknesses in Sybil detection.

Outline: In Section 2, we provide background information on the Tor network. Section 3 introduces the diversity problem in Tor, outlines the Sybil attack, and describes the attacker model. Our methodology is detailed in Section 4. In Section 5, we evaluate the analysis conducted using our methodology. We review related work in Section 6 and conclude our study in Section 7.

2 Background

This section provides essential information on Tor, covering Directory Authorities, onion routing, circuit composition, relay families, and self-reporting by untrusted relays.

Directory Authorities and Consensus. Tor's Directory Authorities (DAs) are a small group of trusted servers (typically 9-10) responsible for maintaining and distributing the network consensus [30]. They collect relay information (IP addresses, public keys, bandwidth) and periodically vote to create a signed consensus document, which is then distributed to clients.

Onion Routing and Communication Circuits. Tor ensures anonymity via *Onion Routing*, where data is encrypted multiple times as it passes through relays. Communication circuits consist of three relay types: *Entry/Guard Relays*, which know the user's IP and are selected for stability; *Middle Relays*, which obscure traffic further; and *Exit Relays*, which forward traffic to its destination but face the highest legal risks. The *Path Selection Algorithm* [30] prevents relays from the same family or /16 subnet from appearing in a single circuit. Despite these safeguards, attacks such as correlation, website fingerprinting, and Sybil attacks can still compromise anonymity.

Relay Families. To prevent circuits from including multiple relays controlled by the same entity, Tor uses *Relay Families*, where operators declare related relays. However, the lack of robust verification allows attackers to manipulate relay selection. Efforts like Authenticated Relay Operator ID (AROI) [23] aim to improve trust in relay operators.

Self-Reporting by Relays. Upon joining the network, relays self-report details such as nickname, contact info, public keys, and family status, which are stored in Server Descriptors (SDs) and included in the network consensus. Since most SD data is unverified to avoid centralization risks, attackers can exploit this by inflating bandwidth claims or misreporting relay families, affecting performance and anonymity. Tor mitigates this with Bandwidth Scanners, but attacks like bandwidth inflation [25,12,32] can bypass these defenses.

3 Problem Statement

This section examines three key security aspects of Tor. First, the *Diversity Problem*, where low relay diversity weakens anonymity and increases attack risks. Second, the *Sybil Attack*, where attackers control multiple relays to compromise circuits. Finally, the *Attacker Model* outlines adversary capabilities, objectives, and strategies.

3.1 Diversity Problem

A diverse set of relays is important for maintaining anonymity in Tor. However, reliance on a few major operators weakens security, increasing the likelihood of multiple relays from the same entity appearing in a circuit. While relays vary in resources, location, and performance, power remains concentrated, heightening attack risks. Even a single compromised relay can threaten anonymity. **(1) Compromised Guard.** Clients use the same guard relay for at least 30 days before rotating. If an attacker controls a guard, they can monitor user traffic patterns, potentially profiling activities despite not seeing the final destination. **(2) Compromised Exit.** Exit relays forward traffic to its destination. A malicious exit relay can profile accessed websites, inject attacks, and intercept unencrypted traffic, exposing sensitive data. **(3) Compromised Guard and Exit.** Controlling guard and exit relays enables traffic correlation attacks, fully de-anonymizing users by linking IP addresses to visited sites. This poses the most severe threat to Tor's privacy guarantees.

3.2 Sybil Attack

In Tor, an attacker's consensus weight determines how much traffic they can observe, impacting the success of attacks like website fingerprinting and traffic correlation. Sybil attacks increase this weight by introducing multiple relays, raising the chances of adversarial relays appearing in circuits and facilitating further exploits [5]. Executing a Sybil attack is simple, as attackers can add relays with or without declaring a family attribute. Omitting this attribute makes relays appear independent, increasing their likelihood of selection. Beyond enabling other attacks, Sybil relays degrade user experience by dropping traffic and weakening anonymity guarantees. They also worsen the *Diversity Problem* by concentrating traffic among fewer operators, reducing network diversity, and exposing users to surveillance. Due to Tor's minimal relay requirements [27] and low-cost hosting, attackers can scale attacks easily. A notable case involved *KAX17*, an operator who controlled over 900 relays across 50+ Autonomous Systems (AS) since 2017, achieving a 16% Guard probability, 35% Middle, and 5% Exit [24]. The Sybil attack is particularly dangerous when combined with bandwidth inflation [12,25,32], making it a cost-effective yet potent threat to Tor's anonymity.

3.3 Attacker Model

In a Sybil attack, the adversary aims to control relays in as many circuits as possible to monitor traffic. Controlling both Guard and Exit relays enables traffic

correlation, potentially compromising anonymity. Tor's Onion Routing assumes no global adversary can observe both ends, as Guard and Exit relays should be operated independently. Thus, Tor does not protect against such an attacker in terms of anonymity.

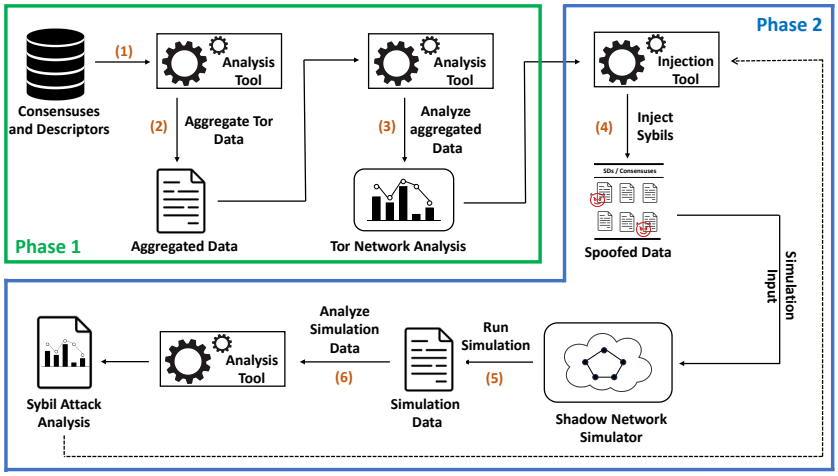
Adversarial Capabilities. We assume an attacker capable of deploying large-scale Sybil relays across multiple AS to obfuscate ownership [43]. The attacker can increase influence by adding new relays, conceal their identities, and execute correlation attacks but cannot compromise existing relays. Our focus is on adversaries who expand existing relay groups, create new ones, or form unlinkable families, resembling the known KAX17 threat actor [24].

Goals. The attacker's primary objective is to deploy Sybil relays undetected, blending in with legitimate relays to avoid removal. By increasing network share, they seek control over Tor's capacity, enabling user monitoring, data collection, and de-anonymization.

Attack Scenarios. We define three scenarios: (1) an attacker controls Sybil relays within a single family to increase circuit presence, maximizing traffic monitoring; (2) an attacker distributes Sybil relays across multiple families, bypassing Tor's path selection to control both Guard and Exit relays; (3) defenders deploy white-hat Sybils [40] within the same subnet as malicious relays, limiting the attacker's traffic observation.

4 Methodology

Fig. 1: Study Methodology



Our methodology, illustrated in Figure 1, has two key objectives. First, we quantify Tor's diversity problem (Section 4.1) by collecting and aggregating network data using our Analysis Tool (AT), which groups relays and computes metrics. This helps identify entities capable of executing effective Sybil attacks

with minimal effort. Second, we investigate Sybil attacks in a simulated Tor environment (Section 4.2) using our Sybil Inject Tool (SIT) and the Shadow simulator [10] to assess attack effectiveness and impact.

4.1 Studying Diversity Problem (Phase 1)

We collect, aggregate, and analyze network data to assess the impact of the diversity problem in Tor (Figure 1, Phase 1, green rectangle). This process consists of three steps:

Step 1: Information Gathering. Tor network data is recorded in hourly consensus files containing relay details, configurations, and metrics. Additional files, such as Server Descriptors (SDs), provide further insights. We obtain and archive Consensuses and SDs from Tor Metrics [29] for analysis.

Step 2: Aggregation. Using our Analysis Tool, we aggregate network data to examine relay grouping and network share distribution. This helps visualize the *Diversity Problem* (Section 3.1) by sorting relays based on attributes like family, AS, country, and AROI. We compute key statistics, including exit/guard probabilities, consensus weight fractions, and total bandwidth, to identify traffic concentration points.

Step 3: Analysis. We identify dominant relay groups that control large portions of traffic and analyze their exit/guard probabilities, consensus weight, and bandwidth distribution. Examining relay relationships and affiliations provides insights into Tor's network composition and the influence of different entities.

4.2 Sybil Attack (Phase 2)

The second phase of our study focuses on the Sybil attack, following an iterative experimental approach (Figure 1, Phase 2, blue area). Each iteration begins by selecting a suitable relay group for the attack.

Step 4: Injecting Sybils. We assess attack effectiveness by increasing the probability of controlling Guard or Exit positions. Using our Sybil Injection Tool (SIT), we inject Sybils into modified consensus and descriptor files from Phase 1. These modifications maintain a consistent network state for Shadow simulator [10] execution. Concealing relay family status allows attackers to occupy both Guard and Exit positions in a single circuit. Candidate relay groups are selected based on high Guard/Exit probability from Phase 1 analysis.

Step 5: Attack Simulation. To ensure ethical testing (Section 4.3), we simulate Sybil attacks instead of deploying them in the live network. We use the Shadow Network Simulator [10,37] to model a realistic Tor environment, incorporating our modified network data with injected Sybils.

Step 6: Simulation Data Analysis. After execution, we evaluate the attack success rate, anonymity impact, and network performance using our Analysis Tool. Key metrics include circuit composition, Guard/Exit positions, and deanonymization probability.

Iterative Loop (Steps 4-6). We refine the attack by introducing new Sybils with varying configurations and running multiple simulations (dashed line in Figure 1). This iterative process deepens our understanding of Sybil attacks.

4.3 Ethical Considerations

To ensure safe research in Tor, the Tor Research Safety Board provides guidelines for assessing research safety and offers feedback on planned experiments [28]. Researchers are advised to use test networks whenever possible, limit attacks to their own relays or traffic, and avoid collecting sensitive user data. Non-compliance poses security risks to Tor and its users. Following these guidelines, we conduct our experiments in a simulated Tor environment. Jansen *et al.* evaluated the Tor simulator, Shadow with Tornettools [10,38], concluding that Shadow accurately models Tor's behavior and scales efficiently. They also highlight its realistic traffic model, replicating Tor conditions even at scale.

4.4 Analysis and Injection Tools

Our **Analysis Tool** comprises three modules: *Relay*, *Aggregation*, and *Simulation Analysis* modules, each supporting different analysis steps.
Relay module. This module consolidates consensus and server descriptor data (Steps 1/2) to analyze relay metrics, including path selection probabilities based on bandwidth weights and relay flags from the Tor Directory Specification [26]. It estimates the likelihood of a relay being chosen for a circuit position, reflecting network load balancing.
Aggregation module. This module groups relays by AROI, relay family, AS, and country (Step 3), providing insights into network composition and geographic distribution. By aggregating consensus weight and exit/guard probability, it evaluates the influence of different relay groups.
Simulation analysis module. After Sybil attack simulations (Step 6), this module processes logs to assess compromised circuits, data volume, and network diversity impact. It evaluates attack feasibility, considering cost and bandwidth requirements for achieving control over guard and exit probabilities.
Sybil Injection Tool. This tool parses Sybil configurations, generates RSEs and SDs, and integrates them into Tor consensuses to simulate various Sybil attack scenarios.

5 Study

In this section, we evaluate our approach across key areas. We describe our testbed and simulation setup (Section 5.1), analyze the diversity problem in Tor (Section 5.2), and assess the impact of Sybil attacks on anonymity (Section 5.3). We then examine bandwidth inflation attacks and their synergy with Sybils (Section 5.4) before evaluating Tor's Sybil detection tool, highlighting its strengths and weaknesses (Section 5.5).

5.1 Testbed

Machine. Our experimental testbed includes a host machine running Debian Buster with two Xeon(R) CPU E5-2690, resulting in 20/40 Cores/Threads, 542 GB of RAM, and 900 GB SSD disk space. Python is version 3.7.3.

Simulation Environment. We simulate a 15% scaled Tor network, including 1,000 relays, 1,100 clients, and 100 servers, reducing computational demands from 2,000 machines. The setup uses Shadow [37] (v2.4.0), Tornettools [11,38], Tor (v0.4.7) [39], and the March 2023 Tor state (Consensus & Descriptors).
Metrics. Our analysis evaluates the following metrics: the bandwidth of Sybil relays and their effectiveness in compromising circuits. Compromised circuits are classified into four categories: one-hop compromised, where either the guard, middle, or exit position is affected; guard-compromised, where only the guard position is controlled; exit-compromised, where only the exit relay is compromised; and bilateral-compromised, where the attacker controls both the guard and exit relays within the same circuit.

5.2 Diversity Problem Analysis

As discussed in Section 4.1, our study's first phase focuses on analyzing the diversity problem in the Tor network. For that, we collect certain characteristics of Tor relays from public SDs and consensus files.

Table 1: Top 10 Guard Probability Relay Families

Families	CW-Frac. (%)	Guard_prob (%)
prsv	1.9	2.4
for-privacy	2.4	1.8
tuxli	2.0	1.7
doedelkiste	0.7	1.1
746d70.github	0.6	1.0
lokodlare	0.5	0.8
we3point1	0.4	0.7
mutantcoder	0.4	0.7
hydra-family	0.4	0.6
BMTY90V...	0.4	0.6

Table 2: Top 10 Exit Probability Relay Families

Families	CW-Frac. (%)	Exit_prob (%)
nothingtohide	4.1	12.2
relayon	3.2	8.1
artikel10	2.4	7.2
quetzalcoatl	2.4	7.0
applied-privacy	2.1	6.4
cccs	1.4	4.1
quintex	1.3	4.0
for-privacy	2.4	3.9
f3netze	1.2	3.7
Nicholas Merril	1.2	3.6

Table 3: Top 10 largest Families

Families	Consensus-Weight-Fraction (%)	Exit_Prob. ($\overline{\%}$)	Guard_Prob. ($\overline{\%}$)	Bandwidth (Gbit/s)	#Relay
relayon.org	3.2	8.1	0.0	22.97	268
quetzalcoatl-relays.org	2.4	7.1	0.0	18.71	182
nothingtohide.nl	4.1	12.2	0.1	26.26	178
artikel10.org	2.4	7.2	0.0	18.31	144
quintex.com	1.3	4.0	0.0	10.72	100
applied-privacy.net	2.1	6.4	0.0	14.18	92
for-privacy.net	2.4	4.0	1.8	16.71	86
prsv.ch	1.9	1.3	2.4	12.42	82
tuxli.org	2.0	2.6	1.7	15.98	77
cccs.de	1.4	4.1	0.0	8.32	48

The results of our analysis are presented in Tables 1 and 2. Table 1 lists the top ten relay families based on guard probabilities, while Table 2 ranks the top ten exit relay families. Each table contains three columns: (1) *Families*, listing relay family names; (2) *CW-Frac.*, representing the consensus weight fraction relative to the network; and (3) *Guard/Exit_prob*, showing the probability of a family being selected as a guard or exit relay. A key observation is that guard

resources are more evenly distributed across families, whereas exit resources are highly concentrated—the top 10 exit relay groups collectively handle over 50% of exit traffic. Notably, the *for-privacy* relay family has substantial control over guard and exit relays, raising the risk of full de-anonymization. Table 3 details the ten largest families in Tor, listing their name, consensus-weight-fraction, exit probability, guard probability, bandwidth (Gbit/s), and number of deployed relays. While most relay families operate only one relay type—reducing immediate correlation attack risks—our analysis reveals a concerning issue: four of the ten largest families (*for-privacy.net*, *prsv.ch*, *tuxli.org*, and *nothingtohide.nl*) control both guard and exit relays, posing a major anonymity threat. Despite Tor's design to prevent the same family from appearing in a circuit, clients can manually select multiple relays from the same family, increasing the risk of end-to-end correlation attacks. These families should, therefore, be considered potential threats, as their ability to control both circuit endpoints directly undermines Tor's fundamental goal of ensuring user anonymity.

5.3 Sybil Attack Analysis

We identified influential relay families with substantial network resources in the previous section. Here, we analyze the potential impact of Sybil attacks if these families were to launch them. We examine three attack scenarios, as detailed in Section 3.3: (i) a Sybil attack within a single family, (ii) a Sybil attack spanning multiple families, and (iii) the deployment of white-hat Sybils.

Sybil attack with single family. Table 4 presents the results of simulating a Sybil attack by introducing a legitimate relay family into the Tor network, including configurations for guards, exits, and both. The table contains six columns: (1) *# of Sybils*, indicating the number of Sybil relays deployed; (2) *Types of Relays*, specifying whether the Sybils act as guards, exits, or both (GUARD+EXIT); (3) *Bandwidth of Sybils*, measuring their available bandwidth; (4) *One Hop-compromised*, showing the fraction of circuits where at least one Sybil relay (guard, middle, or exit) is present; (5) *Guard-compromised*, denoting circuits with a compromised guard relay; and (6) *Exit-compromised*, referring to circuits with a compromised exit relay. The results suggest that adding only exit relays leads to a higher one-hop compromise probability than other configurations. Interestingly, Sybils acting as guards and exits achieve a 9.6% one-hop compromise with 120 relays, which is less effective than using only exits (13.2%) or only guards (11.9%).

Sybil attack with multiple families. This attack scenario is particularly impactful among those analyzed. It simulates a situation where an operator of a large relay family introduces a new, separate relay family into the network. For the Sybil attack to be effective, the new family is carefully configured so that all Sybil relays have a valid family attribute within the existing Tor network from Phase 1 (cf. Section 4.1). One family consists of guard relays, while the other operates exit relays. Since these families are in different subnets, the attacker can bypass Tor's family and subnet-based path selection constraints. Table 5 presents the impact of this attack configuration, using the same columns as Table 4 with

Table 4: Sybil Attack with Single Family.

# of Sybils	Types of Relays	Bandwidth of Sybils (Gbit/s)	One Hop-compromised (%)	Guard-compromised (%)	Exit-compromised (%)
40	GUARD	5.6	2.0	1.1	0.06
80	GUARD	11.5	6.2	3.6	0.2
120	GUARD	18.0	11.9	7.0	0.3
40	EXIT	6.1	4.1	0.0	4.1
80	EXIT	12.6	8.7	0.0	8.7
120	EXIT	19.0	13.2	0.0	13.2
40	GUARD+EXIT	8.7	3.0	0.6	0.6
80	GUARD+EXIT	19.2	8.2	2.5	3.4
120	GUARD+EXIT	24.7	9.6	3.2	3.3

an additional column, *Bilateral-compromised*, which shows the fraction of circuits where the attacker controls both the guard and exit relays. The results reveal that many configurations can compromise over 10% of circuits in at least one position, as shown in the "One-Hop Compromised" column. For instance, the large guard family *tuxli* and exit family *nothingtohide*, by introducing a new Sybil relay family with 18.0/19.0 GBit/s of bandwidth, manage to place one of their relays in nearly every fifth circuit, highlighting the severity of this attack. Furthermore, attacks such as website fingerprinting and Denial of Service do not require control of both guard and exit relays, making them highly effective when combined with this Sybil strategy. Across all experiments, the attacker maintains some probability of controlling both relay positions, which is critical for compromising user anonymity. Notably, the *nothingtohide* family, when combined with Sybil relays, can monitor both ends of a connection with a probability approaching 0.8%, significantly increasing the risk of de-anonymization. This demonstrates the potential for major relay operators or governments to de-anonymize a substantial fraction of the network over time by strategically deploying relays. The ease with which new relays can be introduced into the network further exacerbates this risk.

White-hat Sybils. In this scenario, a defender deploys white-hat Sybils [40] within an existing relay family on the same subnet, leveraging the Path Selection Algorithm [30], which prevents two relays from the same /16 subnet from appearing in the same circuit (cf. Section 2). The family status of the existing relays remains unchanged, ensuring that traffic concentration is reduced for a single relay operator since the white-hat Sybils are controlled by the defender rather than the original relay family. To evaluate this approach, we deployed guard and exit Sybils in four major relay families: *tuxli*, *for-privacy*, *nothingtohide*, and *artikel10*. Guard Sybils were placed in guard-heavy families (*tuxli* and *for-privacy*) to lower their likelihood of selection, while Exit Sybils were deployed in exit-heavy families (*nothingtohide* and *artikel10*). Tables 6 and 7 contain four columns similar to Table 3, but with an additional distinction: *Family* represents the guard/exit probability of the existing family, while *Sybils* reflects the guard/exit probability of the white-hat Sybils. This allows a direct comparison

Table 5: Probabilities of Exit and Guard Sybils in different Families.

# of Sybils + Family	Bandwidth of Sybils (Gbit/s)	One Hop-compromised (%)	Guard-compromised (%)	Exit-compromised (%)	Bilateral-compromised (%)
60 Guard Sybils + nothingtohide	8.4	15.2	2.0	11.8	0.3
120 Guard Sybils + nothingtohide	18.0	21.3	6.1	11.0	0.8
60 Guard Sybils + artikel10	8.4	10.9	2.0	7.4	0.2
120 Guard Sybils + artikel10	18.0	16.4	6.6	5.5	0.4
60 Exit Sybils + for-privacy	9.3	12.3	1.1	10.3	0.1
120 Exit Sybils + for-privacy	19.0	19.4	1.1	17.4	0.2
60 Exit Sybils + tuxli	9.3	11.5	1.8	8.0	0.2
120 Exit Sybils + tuxli	19.0	18.2	2.2	14.5	0.4

between simulations with white-hat Sybils and prior analysis in Section 5.2. Table 6 shows the results of deploying Guard Sybils within the *tuxli* and *for-privacy* families by selecting IPs from their /16 subnets. Compared to previous results (Tables 3 and 4), the effect on guard probability is minimal: *tuxli* decreases from 1.7% (Table 3) to 1.6%, while *for-privacy* drops from 1.8% to 1.5%. The Sybils performed similarly to when they were deployed as a separate relay family. Likewise, Exit Sybils were added to *nothingtohide* and *artikel10*, reducing exit probability from 12.2% to 11.5% for *nothingtohide*, while *artikel10* saw a slight increase from 7.2% to 7.6% (Table 7). The performance of the Sybils remained unchanged compared to when they operated independently. Ultimately, the introduction of white-hat Sybils did not significantly impact the performance of the Sybils or the targeted families as seen in other domains, white-hat Sybils [40] prove ineffective as a defense mechanism in the Tor network.

Table 6: White-Hat Guard Sybils in Pre-existing Family

# of Sybils + Family	Guard_Prob. Family (%)	Guard_Prob. Sybils (%)	Bandwidth Sybils (Gbit/s)
40 + tuxli	1.6	1.7	5.6
80 + tuxli	1.6	3.4	11.5
40 + for-privacy	1.5	1.7	5.6
80 + for-privacy	1.5	3.4	11.5

5.4 Inflation Attack with Sybil Attack

We combine the inflation attack called MirageFlow, introduced by Sendner *et al.* [31,32], with a Sybil attack, dramatically expanding an adversary's capacity to surveil Tor traffic. Its two variants, C-MirageFlow and D-MirageFlow, exploit co-resident relays or dedicated servers to inflate the apparent bandwidth of malicious nodes, thus capturing a larger fraction of user traffic. By leveraging

Table 7: White-Hat Exit Sybils in Pre-existing Family

# of Sybils + Family	Exit_Prob. Family (%)	Exit_Prob. Sybils (%)	Bandwidth Sybils (Gbit/s)
40 + nothingtohide	11.5	4.5	6.1
80 + nothingtohide	11.3	9.1	12.6
40 + artikel10	7.6	4.5	6.1
80 + artikel10	7.5	9.1	12.6

Tor's allowance of up to eight relays per IPv4 address, attackers can advertise disproportionately high bandwidth (e.g., 8 GB/s with only ten servers), potentially monitoring 8% of Tor traffic; deploying five exit and five guard families on dedicated servers can yield 13.44% exit and 6.45% entry traffic observation at a monthly cost of about $400. Although D-MirageFlow's larger-scale tactics require more substantial resources, they pose a serious threat by enabling even broader de-anonymization.

5.5 Sybilhunter Detection Analysis

Sybilhunter was introduced in 2016 and has not been updated in the past four years [21]. This tool has limitations, particularly in its use of Nearest Neighbor Search, as it focuses on relay configurations without considering relationships between them. While this static analysis can identify obvious patterns in relay setups, it is ineffective against a well-planned and deliberate Sybil attack. By introducing diversity in elements such as nicknames, contact information, IP addresses, uptimes, and onion router ports, we could evade Sybilhunter's detection.

5.6 Limitations

Our study relies on simulations, which may not fully capture Tor's dynamic behavior and defenses. While we analyze relay families as a risk factor, disclosed affiliations do not necessarily imply malicious intent, and a broader risk assessment considering adversary-controlled bandwidth or relay count could provide deeper insights. The real-world feasibility of Sybil attacks remains uncertain, as deployment is constrained by economic costs and detection mechanisms like human intervention. Variations in simulation parameters could also affect results, warranting further analysis with probabilistic models like TorPS [13]. Finally, we focus on technical vulnerabilities without addressing the ethical balance between privacy and law enforcement, an important area for future research.

6 Related Work

The diversity problem and Sybil attacks in Tor have received little attention, and our work is the first to address these issues. Previous efforts, such as Sybilhunter [43] and its enhancement by Sui *et al.* [33], focus on detecting Sybils but do not analyze large-scale attacks or their impact. Sybil attacks also occur in social networks [18] and blockchains [36], but defenses relying on social or computational constraints [2,17,40] do not translate well to Tor's decentralized

setting [43]. Related attack vectors include website fingerprinting [9,42], traffic correlation [8,7,19], and bandwidth inflation, with techniques from Øverlier *et al.* [25], Johnson *et al.* [12], and Sendner *et al.* [31,32] amplifying Sybil attacks. Additionally, routing attacks [1,35,34,41] exploit path selection vulnerabilities and can be combined with Sybil attacks. Unlike previous work, we systematically analyze the diversity problem, simulate Sybil attacks, and show how they facilitate other threats.

7 Conclusion

In conclusion, our study reveals significant vulnerabilities in the Tor network due to the diversity problem and the threat of Sybil attacks. By analyzing Tor data and conducting simulations, we demonstrated that the top 10 Exit-based relay families route over 50% of exit traffic, and strategic Sybil attacks could deanonymize 0.8% of circuits. We also showed how Sybil attacks, combined with bandwidth inflation, could further compromise user anonymity at a low cost. Our research highlights the critical need to address these issues and improve Sybil detection methods to enhance the privacy and security of Tor users.

References

1. Bauer, K., McCoy, D., Grunwald, D., Kohno, T., Sicker, D.: Low-resource routing attacks against Tor. In: ACM WPES (2007)
2. Borisov, N.: Computational puzzles as sybil defenses. In: IEEE P2P (2006)
3. Cai, X., Zhang, X.C., Joshi, B., Johnson, R.: Touching from a Distance: Website Fingerprinting Attacks and Defenses. In: ACM CCS (2012)
4. Douceur, J.J.: The Sybil Attack. In: IPTPS (2002)
5. Evers, B., Hols, J., Kula, E., Schouten, J., Den Toom, M., Van der Laan, R., Pouwelse, J.: Thirteen years of Tor attacks, `https://github.com/Attacks-on-Tor/Attacks-on-Tor/blob/master/README.md`
6. He, G., Yang, M., Gu, X., Luo, J., Ma, Y.: A novel active website fingerprinting attack against tor anonymous system. In: IEEE CSCWD (2014)
7. Houmansadr, A., Borisov, N.: SWIRL: A Scalable Watermark to Detect Correlated Network Flows. In: CCS (2018)
8. Houmansadr, A., Kiyavash, N., Borisov, N.: RAINBOW: A Robust And Invisible Non-Blind Watermark for Network Flows. In: NDSS (2009)
9. Jansen, R., Juárez, M., Galvez, R., Elahi, T., Díaz, C.: Inside Job: Applying Traffic Analysis to Measure Tor from Within. In: NDSS (2018)
10. Jansen, R., Newsome, J., Wails, R.: Co-opting linux processes for High-Performance network simulation. In: USENIX ATC (2022)
11. Jansen, R., Tracey, J., Goldberg, I.: Once is never enough: Foundations for sound statistical inference in Tor network experimentation. In: USENIX Security (2021)
12. Johnson, A., Jansen, R., Hopper, N., Segal, A., Syverson, P.: Peerflow: Secure load balancing in tor. In: PoPETs (2017)
13. Johnson, A., Wacek, C., Jansen, R., Sherr, M., Syverson, P.: Users get routed: Traffic correlation on tor by realistic adversaries. In: CCS (2013)
14. Juarez, M., Afroz, S., Acar, G., Diaz, C., Greenstadt, R.: A critical evaluation of website fingerprinting attacks. In: CCS (2014)

15. Klimas, M.: How do ISPs block sites & how to access them anyway (2022), https://surfshark.com/blog/how-do-isps-block-sites
16. Koppen, G., Færøy, A.: Anonymity loves Diversity: The Case of Tor (2020), https://foss-north.se/2020ii/slides/tor_diversity_foss2020.pdf
17. Levine, B., Shields, C., Margolin, N.: A survey of solutions to the sybil attack. In: Proceedings o Technical Report of Univ of Massachussets Amherst (2005)
18. Lobo, A., Mandekar, Y., Pundpal, S., Roy, B.: Detection of sybil attacks in social networks. In: CSoNet (2020)
19. Nasr, M., Bahramali, A., Houmansadr, A.: DeepCorr: Strong Flow Correlation Attacks on Tor Using Deep Learning. In: CCS (2018)
20. Nithyanand, R., Starov, O., Zair, A., Gill, P., Schapira, M.: Measuring and mitigating AS-level adversaries against Tor. In: NDSS (2015)
21. NullHypothesis: Sybilhunter GitHub Repository, https://github.com/NullHypothesis/sybilhunter
22. Nusenu: OrNet Stats, https://nusenu.github.io/OrNetStats/
23. Nusenu: Tor-Relay-Operator-IDs-Trust-Information, https://nusenu.github.io/tor-relay-operator-ids-trust-information/
24. Nusenu: Is KAX17 performing deanonymization Attacks against Tor Users? (2021), https://nusenu.medium.com/is-kax17-performing-de-anonymization-attacks-against-tor-users-42e566defce8
25. Overlier, L., Syverson, P.: Locating hidden servers. In: IEEE S&P (2006)
26. Project, T.: Directory Specification, https://gitlab.torproject.org/tpo/core/torspec/-/blob/HEAD/dir-spec.txt#L3145
27. Project, T.: Relay Requirements, https://community.torproject.org/relay/relays-requirements/
28. Project, T.: Research safety board, https://research.torproject.org/safetyboard/
29. Project, T.: Tor metrics, https://metrics.torproject.org
30. Project, T.: Tor Specification, https://github.com/torproject/torspec/blob/main/
31. Sendner, C., Stang, J., Dmitrienko, A., Wijewickrama, R., Jadliwala, M.: Tormult: Introducing a novel tor bandwidth inflation attack. In: preprint, arXiv (2023)
32. Sendner, C., Stang, J., Dmitrienko, A., Wijewickrama, R., Jadliwala, M.: Mirageflow: A new bandwidth inflation attack on tor. In: NDSS (2024)
33. Sui, J., Guo, W., Shi, X., Zhang, S., Hu, H.: Refining node similarity analysis: An optimized nearest-neighbor ranking algorithm. In: IEEE ICCECT (2024)
34. Sun, Y., Edmundson, A., Feamster, N., Chiang, M., Mittal, P.: Counter-RAPTOR: Safeguarding Tor Against Active Routing Attacks. In: IEEE S&P (2017)
35. Sun, Y., Edmundson, A., Vanbever, L., Li, O., Rexford, J., Chiang, M., Mittal, P.: *RAPTOR* : Routing attacks on privacy in tor. In: *USENIX* Security (2015)
36. Swathi, P., Modi, C., Patel, D.: Preventing sybil attack in blockchain using distributed behavior monitoring of miners. In: ICCCNT (2019)
37. Team, S.: Shadow - Network Simulator, https://shadow.github.io
38. Team, S.: Tornettools, https://github.com/shadow/tornettools
39. Tor: The Tor Project, https://www.torproject.org/
40. Trifa, Z., Khemakhem, M.: Sybil nodes as a mitigation strategy against sybil attack. Procedia Computer Science, ANT, SEIT (2014)
41. Wan, G., Johnson, A., Wails, R., Wagh, S., Mittal, P.: Guard Placement Attacks on Path Selection Algorithms for Tor. PoPETs (2019)
42. Wang, T., Goldberg, I.: Improved website fingerprinting on Tor. In: WPES (2013)
43. Winter, P., Ensafi, R., Loesing, K., Feamster, N.: Identifying and Characterizing Sybils in the Tor Network. In: USENIX Security (2016)