A Summary of An Anonymity Vulnerability in Tor

Damian Franco
Department of Computer Science
CS-585 - Computer Networks

Title: An Anonymity Vulnerability in Tor

Authors: Qingfeng Tan, Xuebin Wang, Wei Shi, Jian Tang, Zhihong Tian

Published: IEEE/ACM TRANSACTIONS ON NETWORKING, VOL. 30, NO. 6, DECEMBER 2022

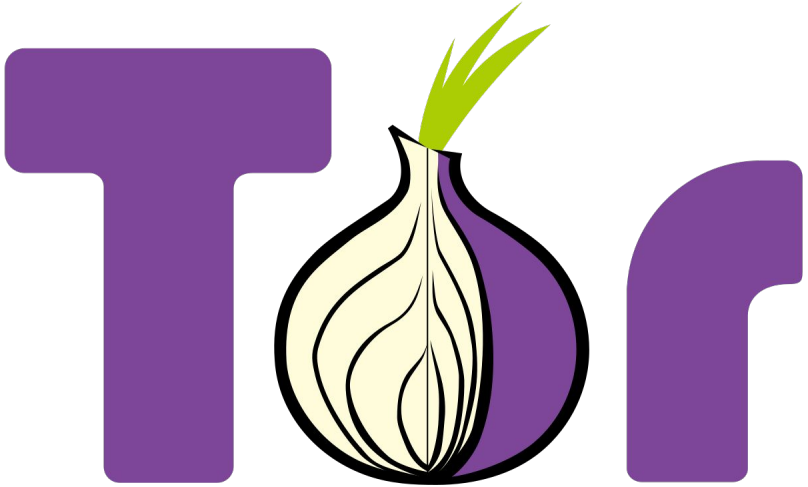
https://ieeexplore.ieee.org/document/9778273

Introduction

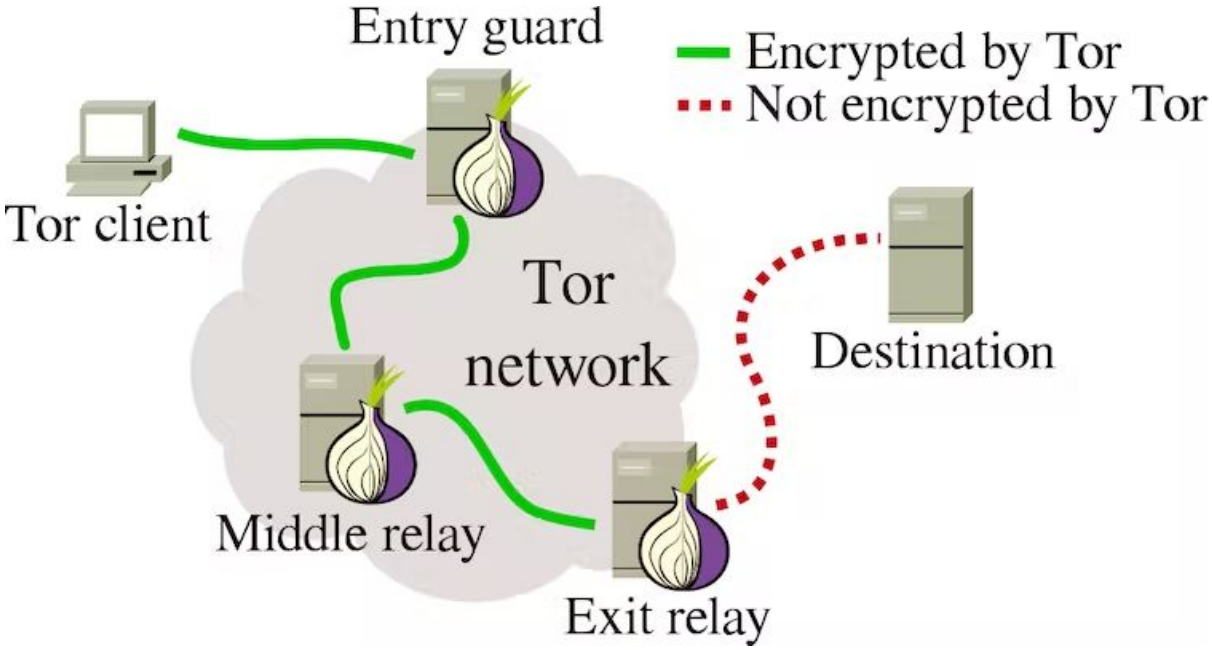
Privacy is and will continue to be one of the most concerning issues in cyberspace.

Tor

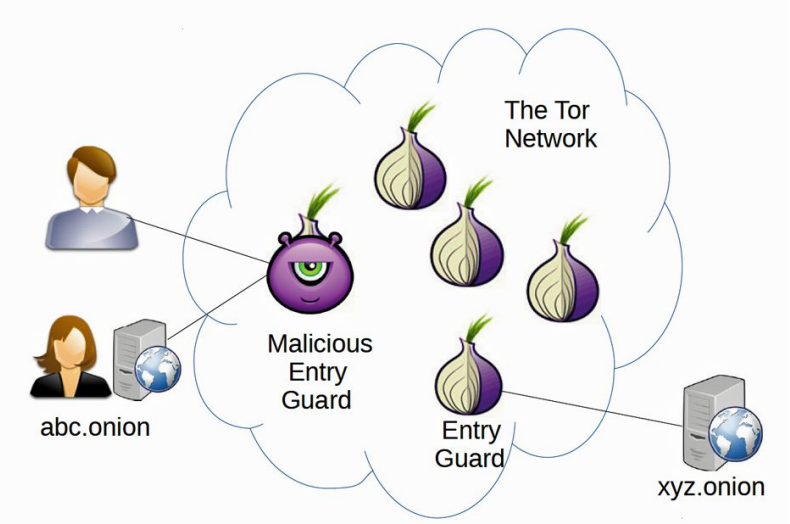
Tor or “The Onion Router” is a free and open-source software that allows users to browse the internet anonymously by routing their traffic through a network of volunteer-operated servers. Tor has become increasingly popular for protecting their communication anonymity. Tor and other anonymous communication techniques have good intentions in design, but are heavily exploited for many criminal activities.



Tor Network


Entry guard
Encrypted by Tor
Not encrypted by Tor
Tor client
Tor
network
Destination
Middle relay
Exit relay

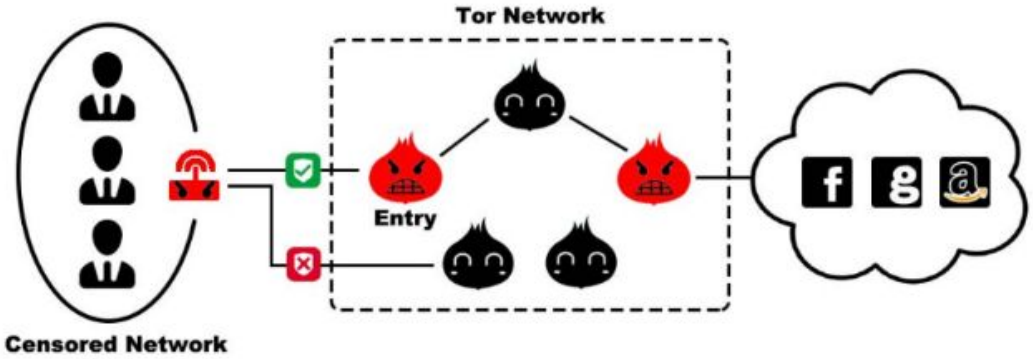
Deanonymization



A data mining strategy in which anonymous data is cross-referenced with other data sources to re-identify the anonymous data source. Most attacks on Tor focus on identifying a relationship between a client and a server that are using the Tor network to communicate

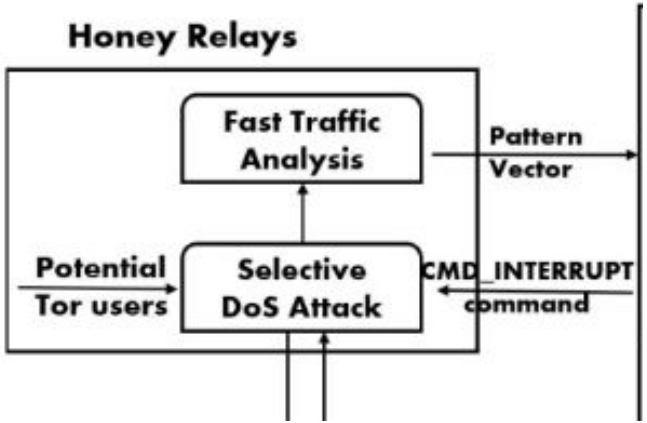
Trapper Attacks

We present a set of new attacks, called Trapper Attacks, that can provide an adversary the ability to selectively control Tor's routing path. Such an attack degrades the Tor path selection of a Tor user from probabilistic node selection to deterministic node selection.



“Honey” Relay

A “honey” relay refers to a relay node that has been intentionally compromised by an adversary in order to monitor or manipulate Tor traffic. The term “honey” implies that the node is used to attract and deceive unsuspecting Tor clients, as opposed to being a legitimate node in the Tor network. By injecting honey relays, the Trapper Attack can compromise circuits and be misperceived as poor communication reliability, and combined with other attacks.



Main Goal

To investigates the Trapper Attacks on the Tor network, comprehensively studies its efficiency, accuracy, and feasibility, and validates it with experiments on a live Tor network

Main Contributions

1) Introduce Trapper Attacks which are a set of new attacks that can give an adversary the power to control Tor's routing path, leading to a shift from probabilistic to deterministic node selection for a Tor user.
2) Test their survivability against Sybilhunter and DannerDetector.
3) Proof-of-concept of an implementation of Trapper Attacks that results in communication deanonymization.
4) Large-scale simulations and real-world experiments on live Tor networks, and the evaluation results are presented.

Methods

Two sections of methods:

1) Threat Model
2) Proposed Trapper Attack

Methods: Threat Model

- The Tor network, which is designed for anonymous communication, is vulnerable to adversaries that can monitor network traffic entering and exiting the Tor communication channel.
- The adversaries in this model are assumed to be organizations such as internet service provider’s (ISPs) that can control Tor relays as honey relays and monitor all targeted Tor user's incoming and outgoing connections.
- It is assumed that the underlying network protocols are secure, but the types and amounts of honey routers that an adversary can control to launch a deanonymizing attack are limited.

Methods: Proposed Trapper Attacks

Basic Attack:

- Guard Relay Capture (GRC) Attack
- Controlling User's Exit Node

The Advanced Attacks With Anti-Attack-Detection Mechanisms:

- Sybilhunter
- DannerDetector

Deanonymization in Tor:

- The Trapper Attacks Prototype
- Circuit Controller
- Fast Traffic Analysis
- Traffic Correlation

Methods: Basic Attacks

Guard Relay Capture (GRC) Attack:

An ISP can use Tor's relay selection algorithms to manipulate the routing path of Tor by diverting from the actual guard relays. By hijacking Tor's relay selection, the traffic is automatically routed to the honey relays.

Steps:

1) Identify the entry guards of a Tor user.
2) Test their survivability against Sybilhunter and DannerDetector.
3) Repeat both steps above until the adversarial relay is selected as an entry guard.

This type of attack enables the adversary to control all responsible entry guards of a specific Tor client for a certain period.

Methods: Basic Attacks

Controlling User's Exit Node:

In order to conduct a successful deanonymization attack, an adversary must be able to monitor Tor users' traffic at both the entry and exit points of the communication channel.

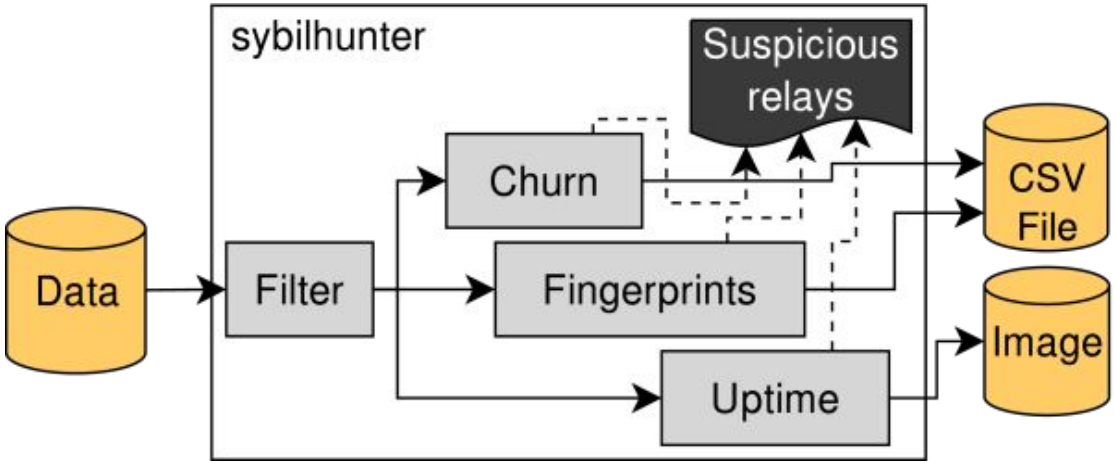
To achieve this, we exploit the ability of CMD_INTERRUPT command mechanism in Tor that can precisely destruct a Tor circuit using a given circuit ID.

This is done by destroying the circuit and repeating the process until a honey relay is selected as the exit node. The attack enables the adversary to control both the entry and exit guards of the communication and launch an end-to-end traffic correlation attack.

Methods: Advanced Attacks With Anti-Attack-Detection Mechanisms

Sybilhunter:

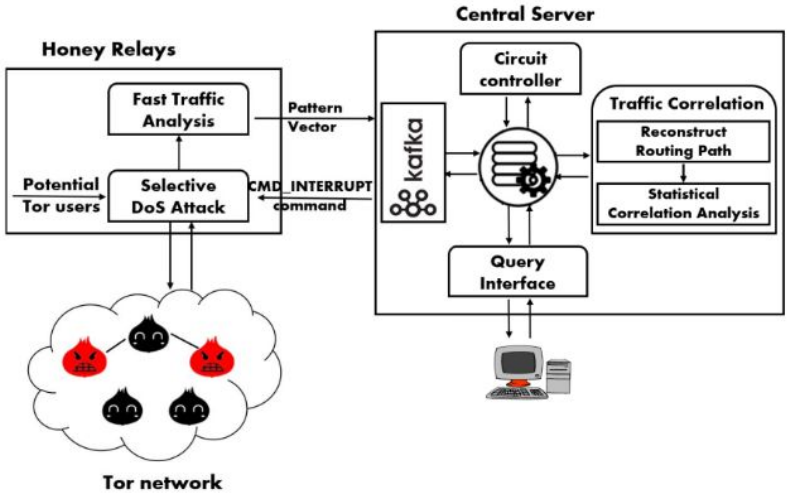
Sybilhunter is the system for detecting Sybil relays based on their appearance and behavior, such as configuration and uptime sequences. The more honey relays we inject into Tor network, the easier of Sybilhunter to find out our honey relays.



Methods: Advanced Attacks With Anti-Attack-Detection Mechanisms

DannerDetector:

DannerDetector aims to detecting selective denial of service (DoS) relays in Tor network. It could find out all honey relays within several rounds of execution. On each round, the detector established multiple testing circuit according to its policy through every relays for downloading a large file. If the circuit is interrupted when the file is downloading, the detector would suspect the existence of honey relays in the circuit.



Methods: Advanced Attacks With Anti-Attack-Detection Mechanisms

Obfuscating Deployment: To simplify the management of dozens of honey relays, the reported detection operators tend to adminisrate their relays simultaneously, such as reboot all of them at the same time or use the same configuration file

Deferral decision: DannerDetector can download a self-owned file through testing circuit which contains the honey relays. If the circuit is not compromised, the honey relay of basic Trapper Attacks would interrupt the circuit almost in a certain time. As a result, DannerDetector could easily discover the honey relays.

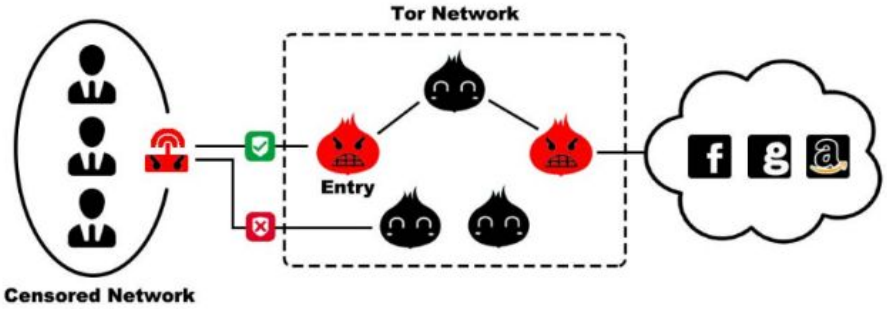
Collaborative Filtering: In order to find out each honey relays in Tor network, DannerDetector test many rounds to increase the accuracy of its algorithm.

Obfuscating deployment aims to bypass the Sybilhunter through varying the configuration of Tor instance. Deferral decision and collaborative filtering are proposed to conceal our honey relays from DannerDetector.

Methods: Deanonymization in Tor

The Trapper Attacks Prototype:

This experiment aims to demonstrate how an adversary can selectively affect the reliability of Tor nodes, including guard and exit relays by using honey routers deployed in Tor network.



Existing traffic correlation analysis only considers scenarios where an adversary can monitor traffic in both directions between Tor clients and Internet destinations, but the paths between the client-to-entry may cross censored networks, making it difficult for adversaries to monitor both paths, and the use of fixed entry guards by Tor clients also decreases adversary visibility.

Methods: Deanonymization in Tor

Circuit Controller:

Uses control and relay cells to communicate between nodes and users, with control cells issuing commands and relay cells carrying end-to-end TCP stream data, while the Tor control protocol enables communication between the controller and the locally running Tor instance. The Circuit Controller tests for a Pearson correlation coefficient, and if it is below a certain threshold. It then sends an interrupt command to a honey relay to interrupt the circuit in one of two ways. This successfully destroys the circuit.

Relay Command Examples:

- RELAY_BEGIN
- RELAY_DATA
- RELAY_END
- RELAY_SENDME
- RELAY_DROP
- RELAY_RESOLVE

Methods: Deanonymization in Tor

Fast Traffic Analysis:

The fast traffic analysis techniques used by an adversary involve monitoring high-level protocol features and translating raw cells into pattern vectors to determine whether two circuits share a common path.

Feature	Value	Description
Cell Count	Integer	The amount of cells that a circuit sends or receives per unit time
Throughput	Float	The amount of data that a circuit transfers per unit time
Cell Type	Boolean	The *Command* field of a fixed-length cell
Direction	Boolean	Incoming or outgoing
Circuit ID	Integer	The *CircID* field determines which circuit

Fast traffic analysis translates each raw cell into pattern vectors that can facilitate further traffic correlation.

Methods: Deanonymization in Tor

Traffic Correlation:

To address the difficulty of correctly correlating the flows of Tor users, the traffic correlation module utilizes a clustering algorithm to group together Tor users who exhibit similar traffic patterns, thereby reducing the number of potential user-destination pairs to be correlated. The Pearson's correlation coefficient ρ between the pattern vector of circuit-based fingerprinting features over time to identify the closest matching flows for our correlation analysis.

Pearson Correlation Coefficient:

$$\rho = \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^{n}(x_i - \bar{x})^2}\sqrt{\sum_{i=1}^{n}(y_i - \bar{y})^2}}.$$

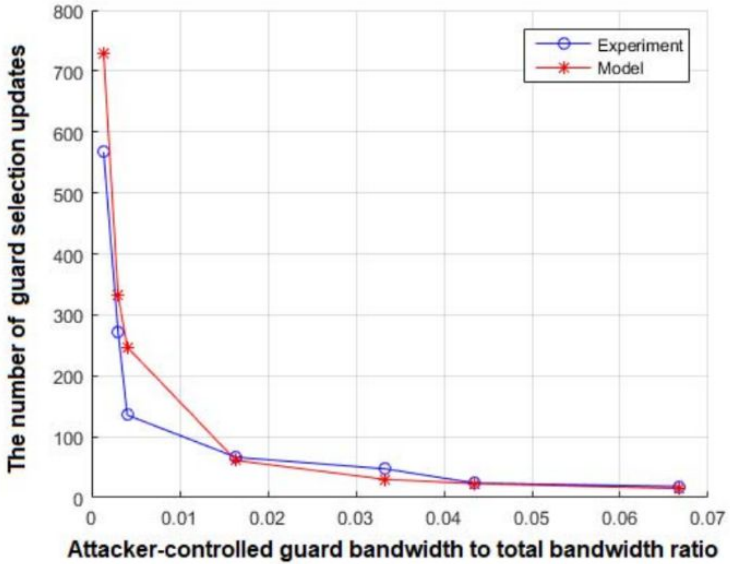
Results

Three sections of results:

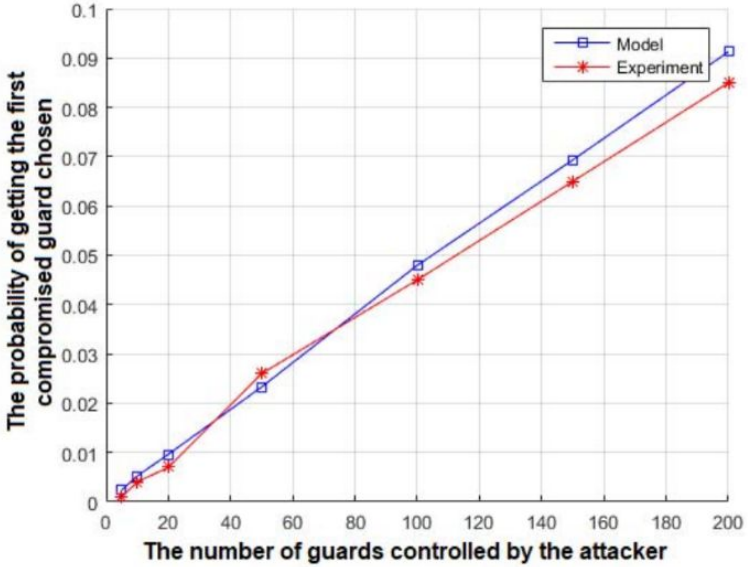
1) Effectiveness of Trapper Attacks
2) Advanced Trapper Attacks Survivability Evaluation
3) Accuracy of Deanonymization Attacks

Results: Effectiveness of Trapper Attacks

The probability of selecting the first compromised node as a guard node is affected by the number of guards under control of the attacker.



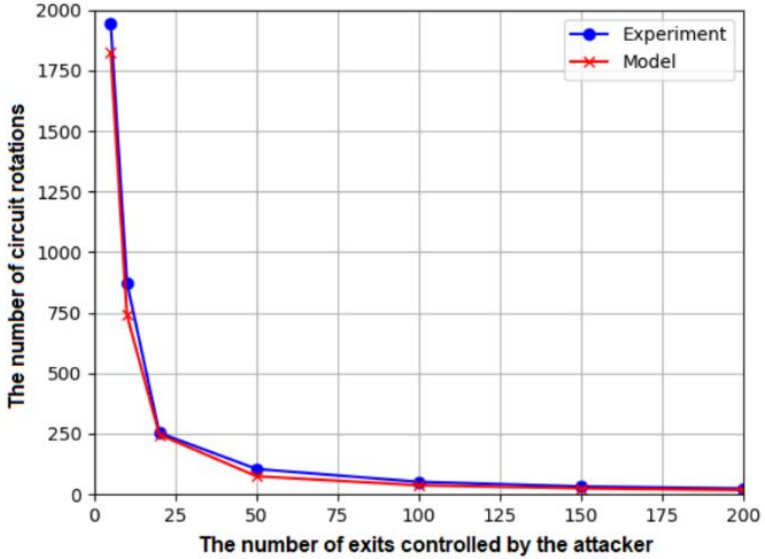
Results: Effectiveness of Trapper Attacks



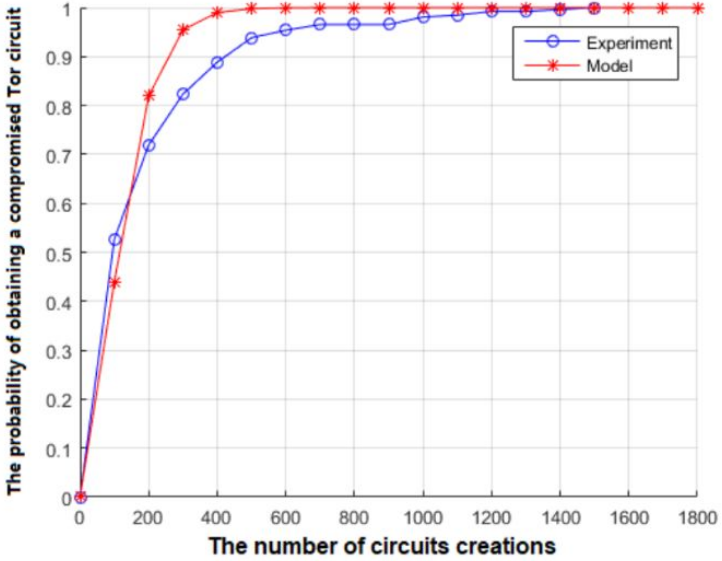
The ratio of attacker-controlled guard bandwidth affects the number of updates necessary to select the first compromised guard.

Results: Effectiveness of Trapper Attacks

The number of exits controlled by the attacker affects the amount of circuit rotations, i.e., destruction and reconstruction of circuits, needed to select the first compromised node as an exit

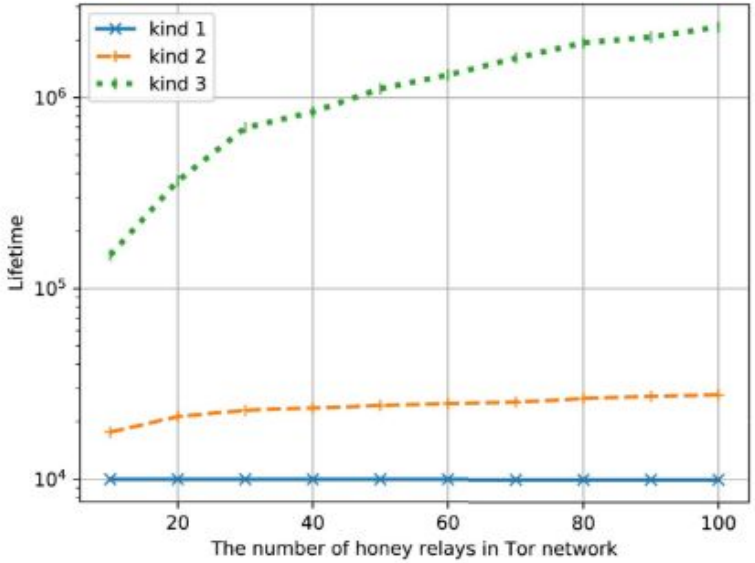


Results: Effectiveness of Trapper Attacks



The probability of successfully obtaining a compromised Tor circuit, including a guard and an exit node, while varying the number of circuit creations, given that 100 guard nodes and 20 exit nodes are under the control of attackers.

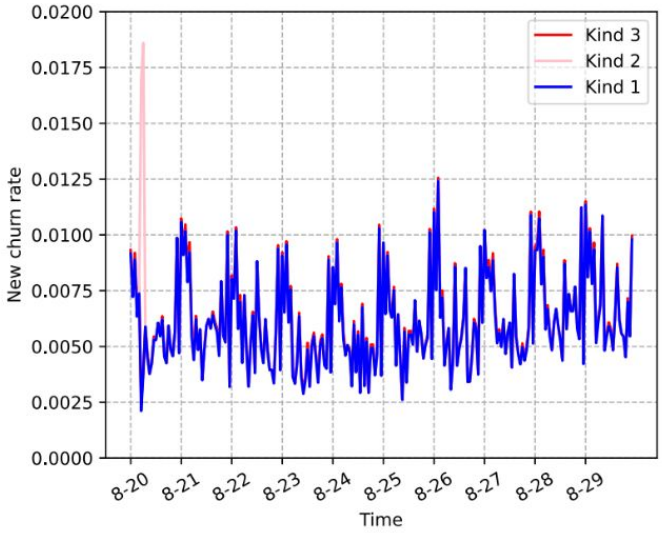
Results: Advanced Trapper Attacks Survivability

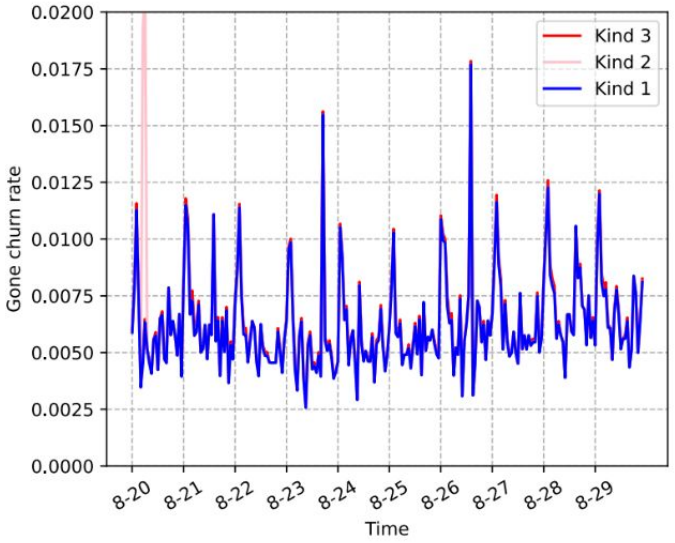


Lifetime of different policy under DannnerDetector.

Results: Advanced Trapper Attacks Survivability

Calculation of the network churn rate of the joining or leaving relays on three different data sets is done. Both plots illustrate the network churn rates during ten days in August 2020. We found an unexpectedly high churn rate without applying the obfuscating method in 2020-08-20. The high new or gone churn rate means that many relays joined or left the Tor network as also revealed in Sybilhunter.





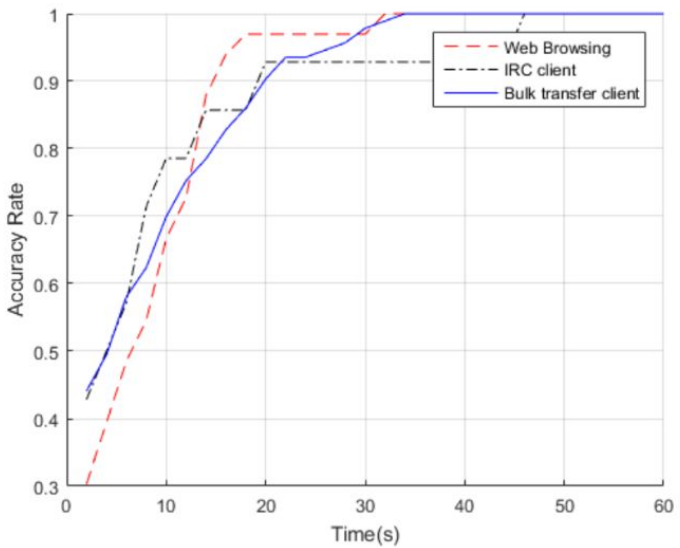
Results: Advanced Trapper Attacks Survivability

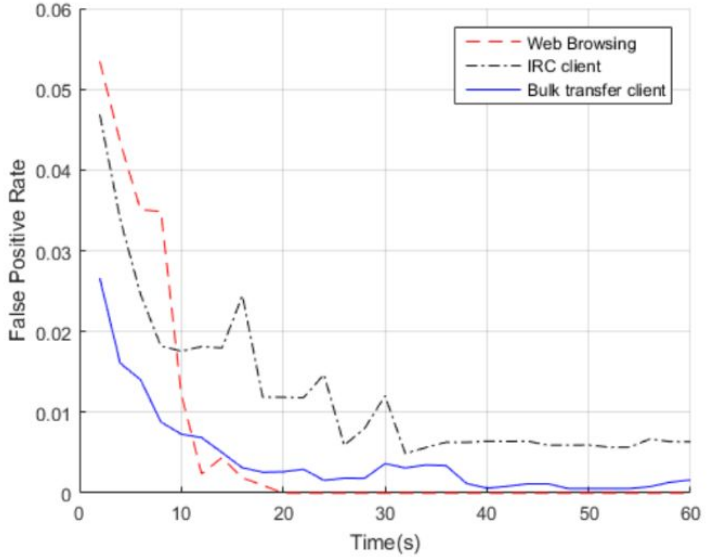
Month	Injected relays	Survival rate with Trapper Attacks	Survival rate without Trapper Attacks
2020-01	209	95.20%	5.74%
2020-02	195	92.80%	5.13%
2020-03	219	93.60%	6.85%
2020-04	220	95.00%	4.09%
2020-05	211	96.20%	6.16%
2020-06	197	94.90%	7.10%
2020-07	205	92.19%	5.37%
2020-08	213	94.36%	4.22%

Through running Sybilhunter on the origin datasets and honey datasets, we can evaluate the effectiveness of the Trapper Attacks before and after applying the obfuscating method. The result of detection shows that we can inject 209 honey relays per month on average with our obfuscating method and about 94.28% of them could not be detected by Sybilhunter. However, the survival rate is only at 5.58% without applying our proposed obfuscating method.

Results: Accuracy of Deanonymization Attacks

Both demonstrate that the source IP and destination IP of a user can be accurately detected within 30 seconds with a false positive rate of less than 1%.





Conclusion

- Trapper Attacks are introduced, which enable adversaries to identify anonymous communication over Tor accurately.
- Through experiments, the attacks are shown to allow adversaries to control the routing path by selectively affecting the reliability of Tor circuits or hijacking Tor's relay selection, thereby deanonymizing user communications.
- Trapper Attacks deny service on trustworthy onion routers to redirect users' data towards adversary-controlled honey relays, resulting in a significant increase in the adversary's knowledge.
- The feasibility, survivability, and effectiveness of the attacks are demonstrated through an experiment on a live Tor network, which shows that the attacks can deanonymize the network in near real-time with a high survival rate even in the presence of honey relay detection software.
- Presents a formal analysis framework to quantitatively assess the severity of the attacks and their security implications on the live Tor network.

Thank you for watching!