



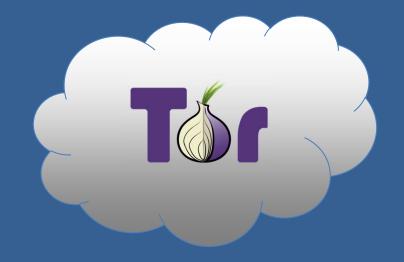
Users Get Routed: Traffic Correlation on Tor by Realistic Adversaries Aaron Johnson¹ Chris Wacek² Rob Jansen¹ Micah Sherr² Paul Syverson¹ ¹U.S. Naval Research Laboratory, Washington, DC ² Georgetown University, Washington, DC



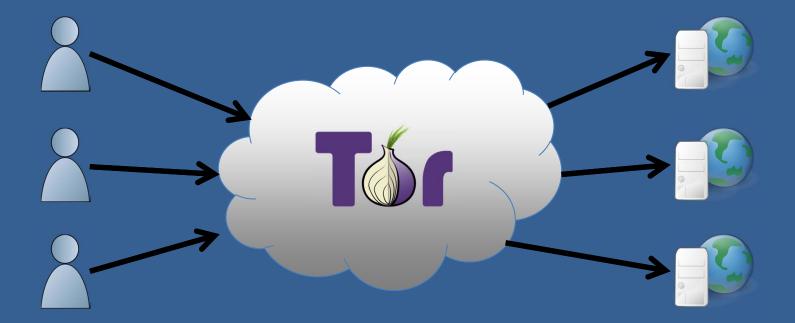
MPI-SWS July 29, 2013



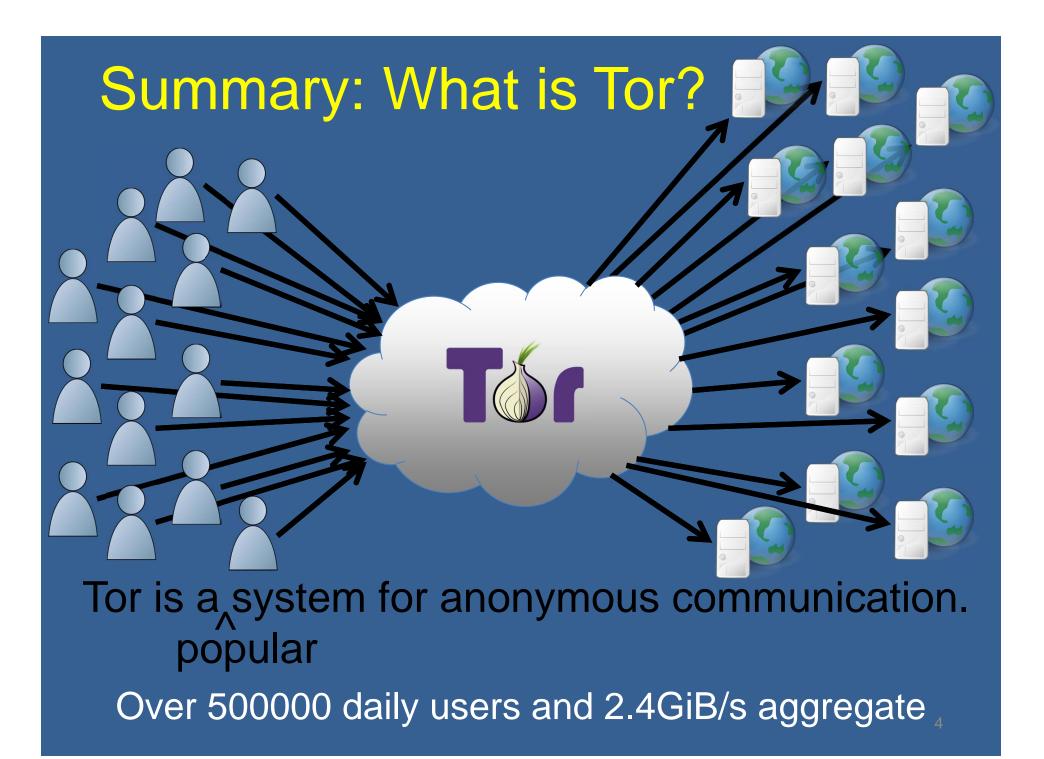
Summary: What is Tor?



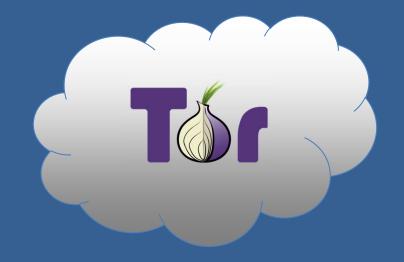
Summary: What is Tor?



Tor is a system for anonymous communication.

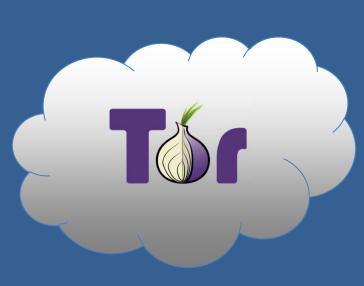


Summary: Who uses Tor?

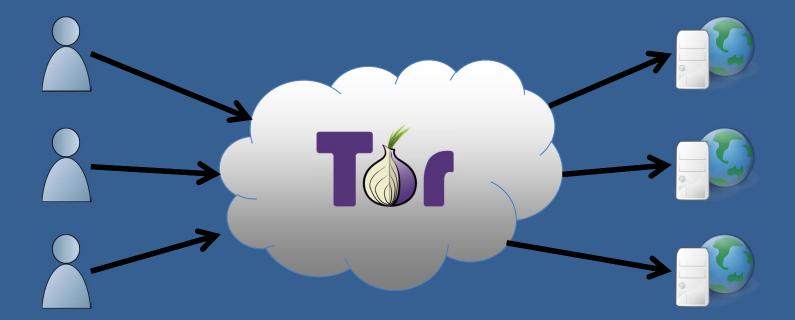


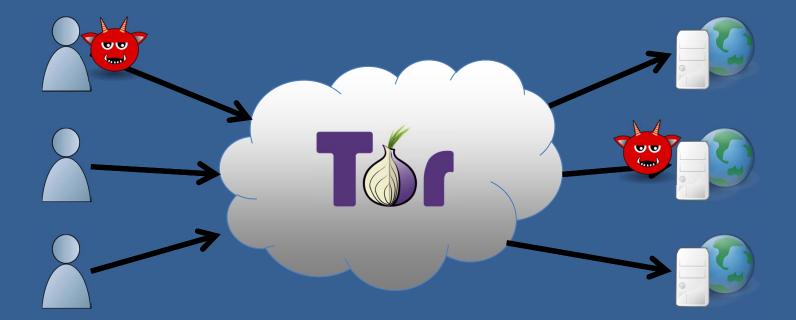
Summary: Who uses Tor?

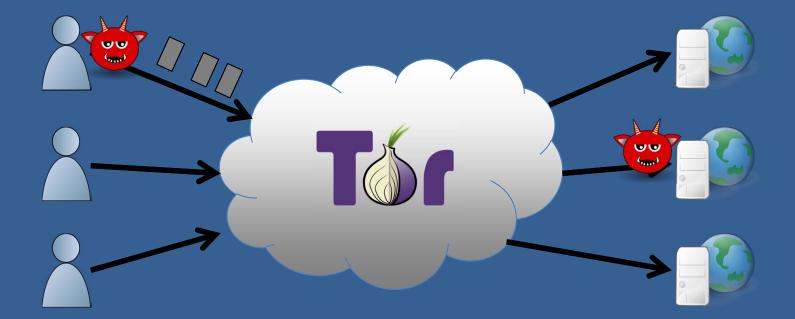
- Individuals avoiding censorship
- Individuals avoiding surveillance

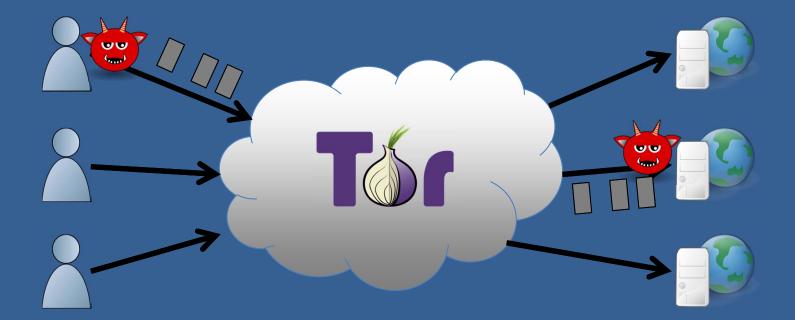


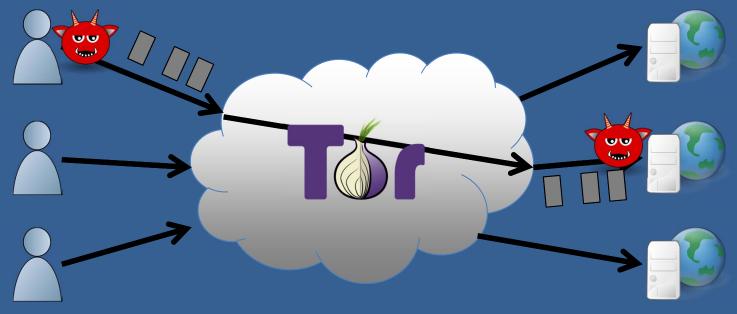
- Journalists protecting themselves or sources
- Law enforcement during investigations
- Intelligence analysts for gathering data



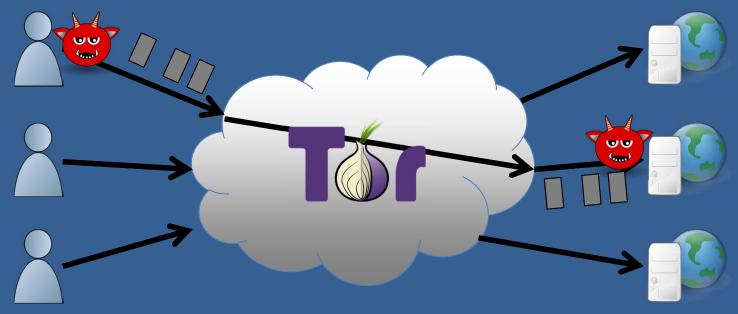








Traffic Correlation Attack

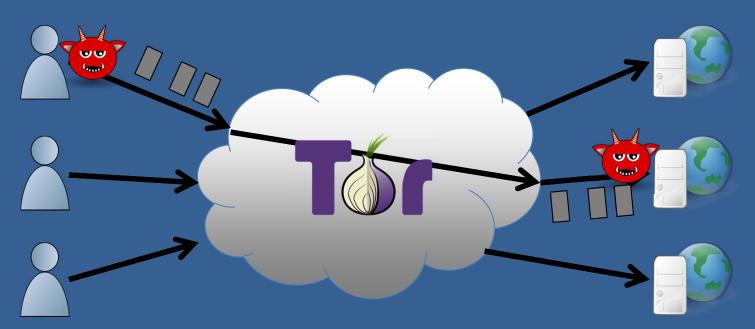


Traffic Correlation Attack

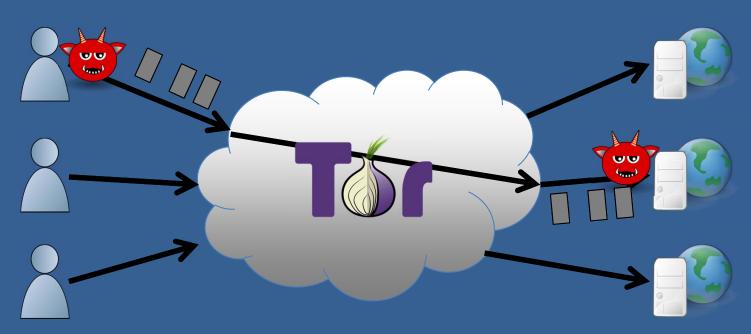
- Congestion attacks
- Throughput attacks
- Latency leaks

- Website fingerprinting
- Application-layer leaks
- Denial-of-Service attacks

Summary: Our Contributions



Summary: Our Contributions



- 1. Empirical analysis of traffic correlation threat
- 2. Develop adversary framework and security metrics
- 3. Develop analysis methodology and tools

Overview

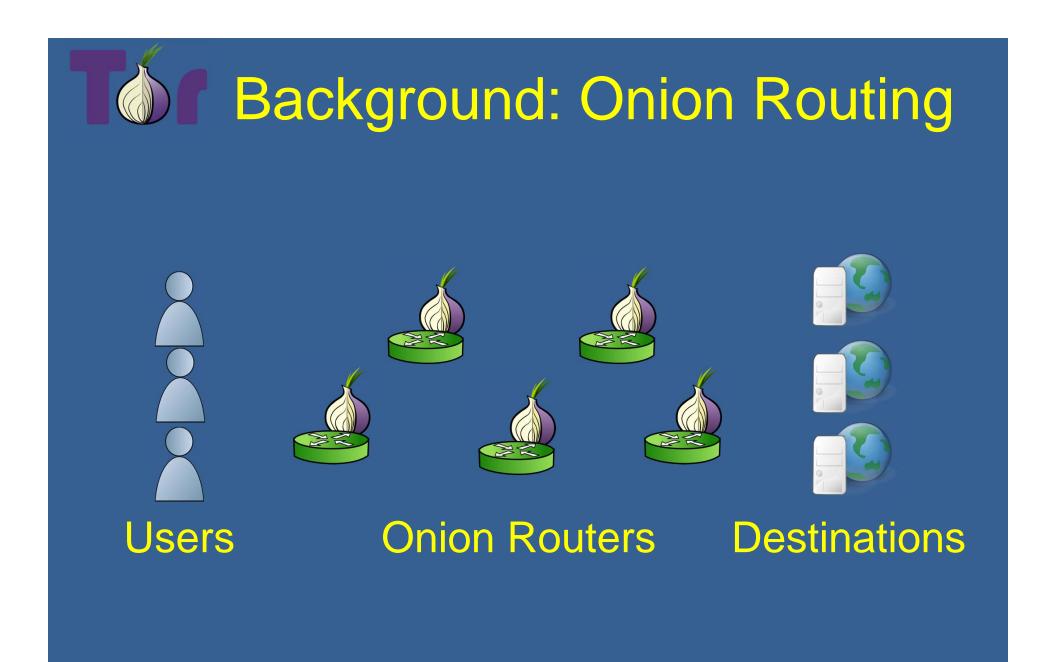
- Summary
- Tor Background
- Tor Security Analysis

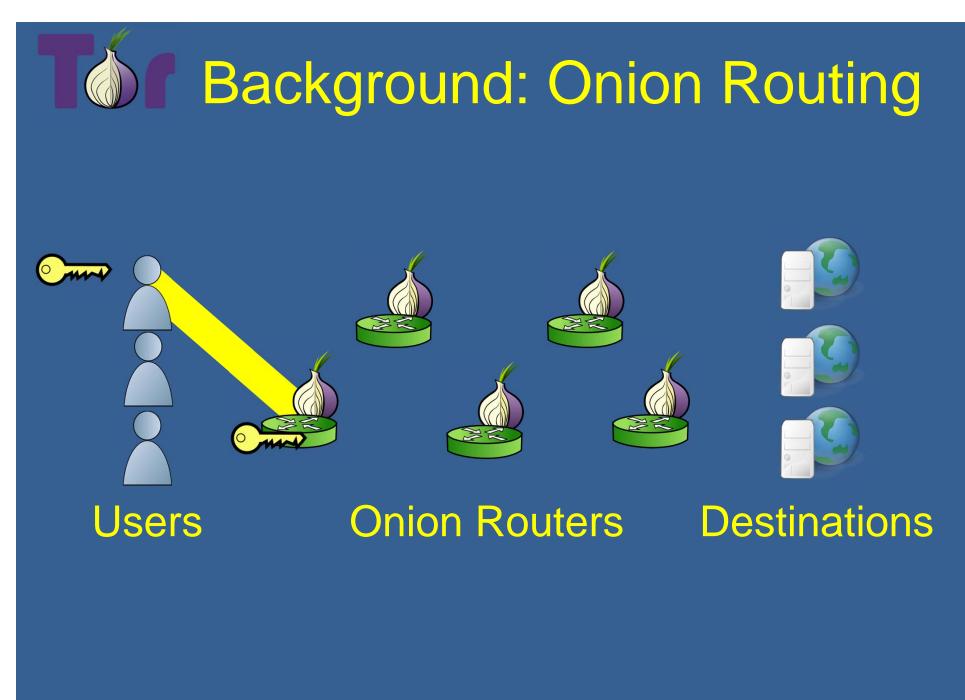
 Adversary Framework
 Security Metrics
 Evaluation Methodology
 Node Adversary Analysis
 Link Adversary Analysis
- Future Work

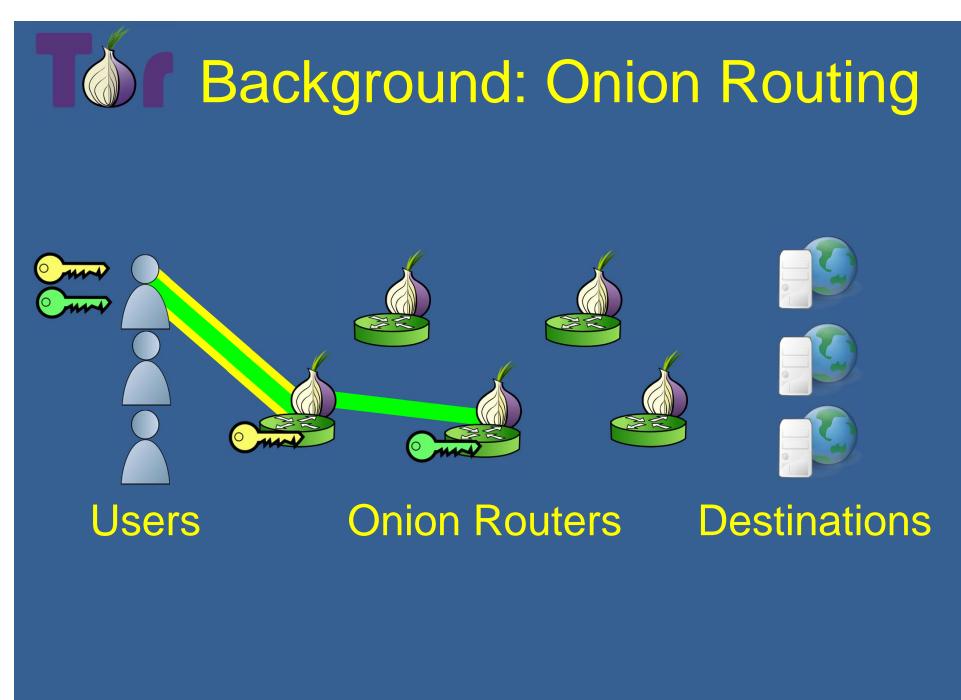
Overview

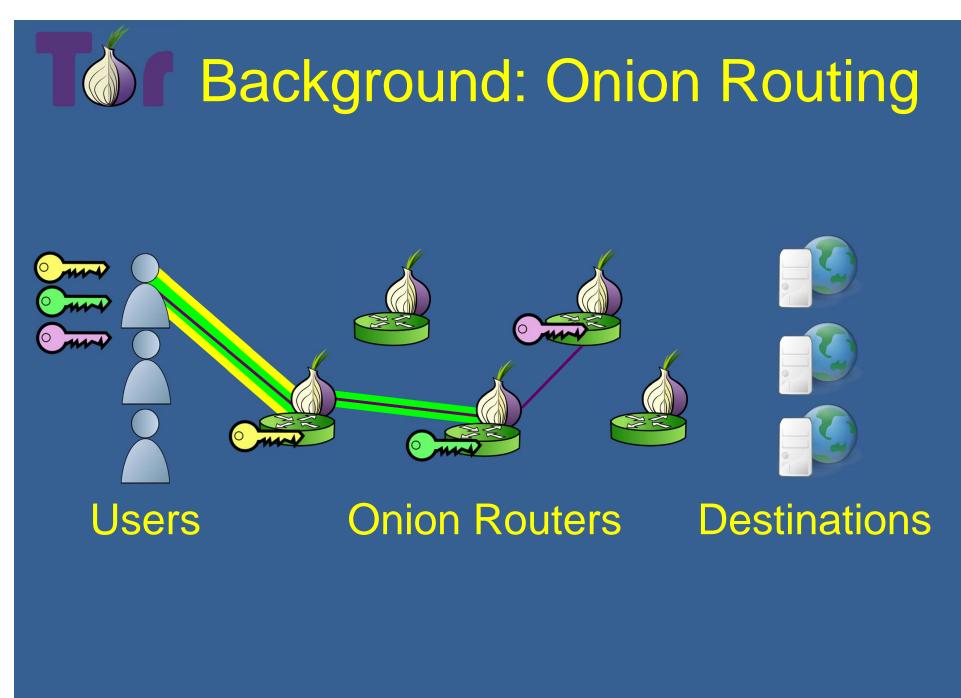
- Summary
- Tor Background
- Tor Security Analysis

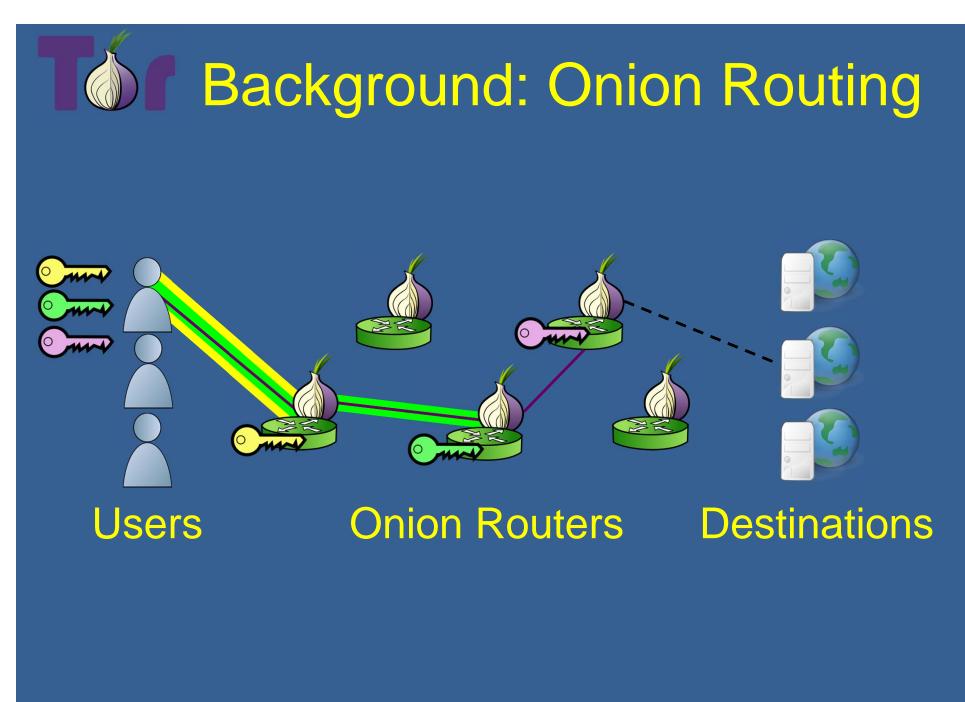
 Adversary Framework
 Security Metrics
 Evaluation Methodology
 Node Adversary Analysis
 Link Adversary Analysis
- Future Work

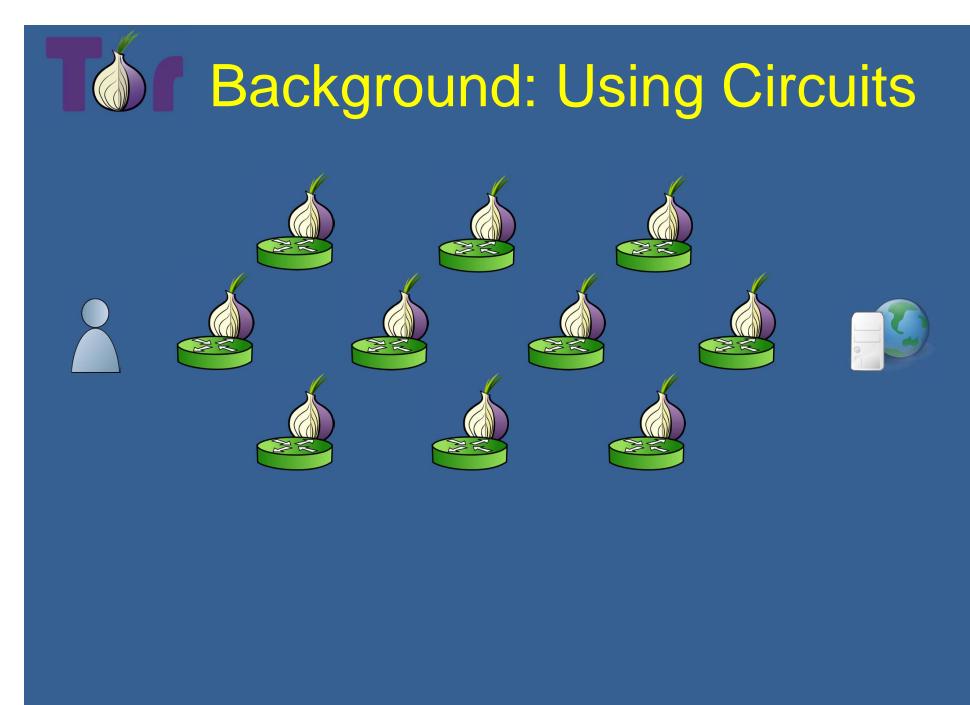


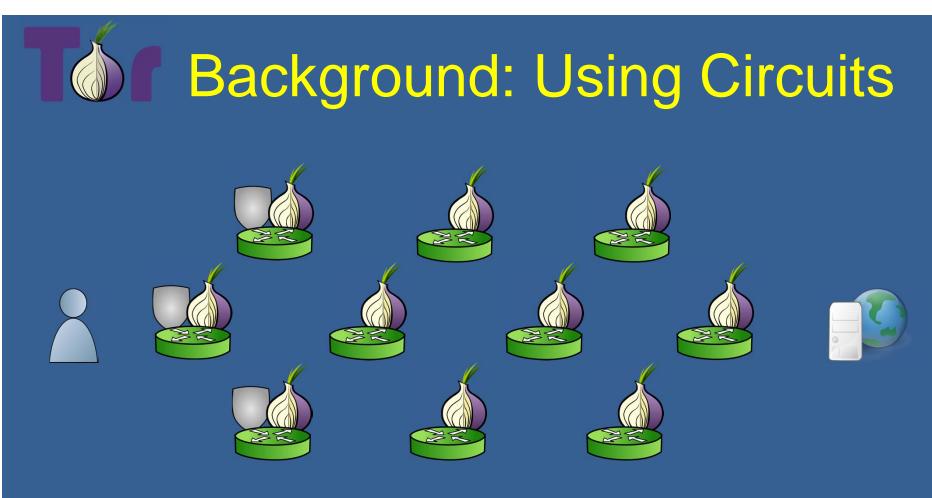




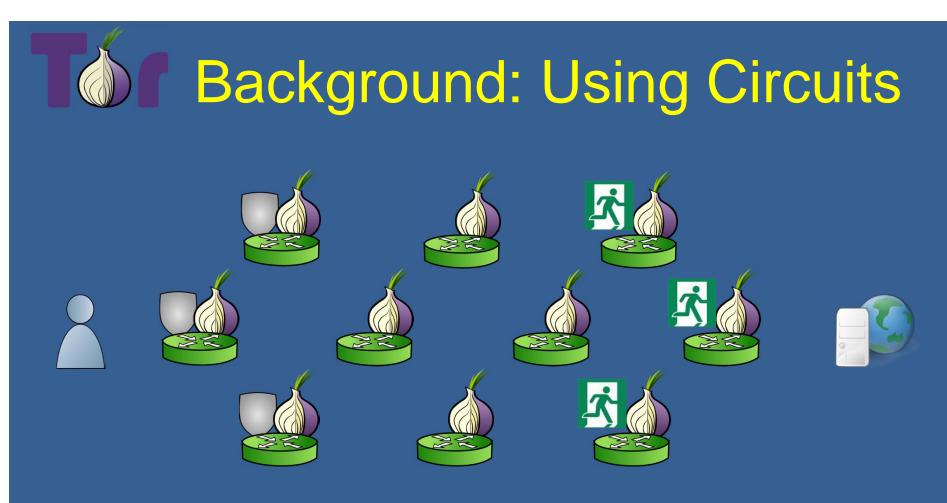




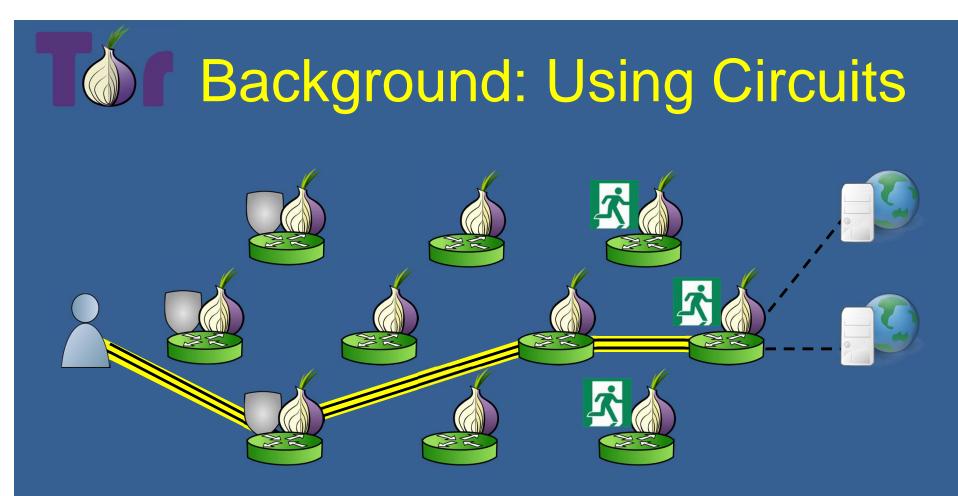




1. Clients begin all circuits with a selected guard.

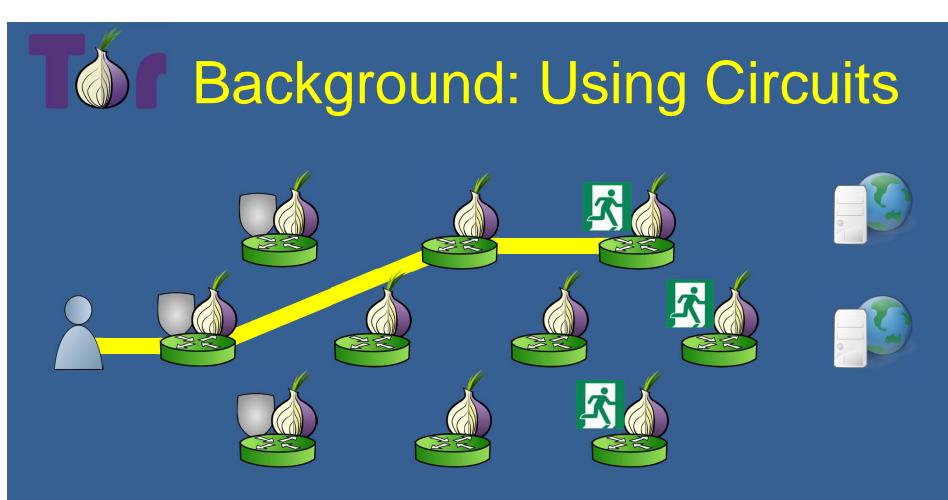


Clients begin all circuits with a selected *guard*.
 Relays define individual *exit policies*.



Clients begin all circuits with a selected *guard*.
 Relays define individual *exit policies*.

3. Clients multiplex streams over a circuit.

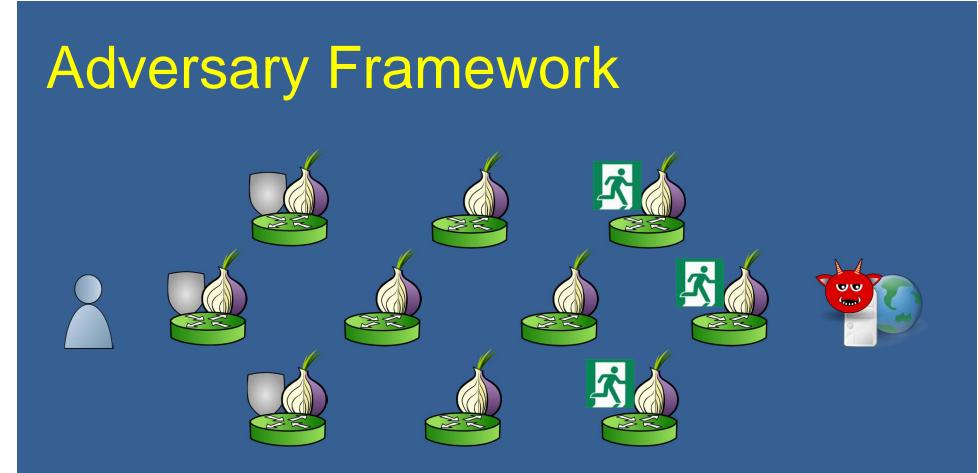


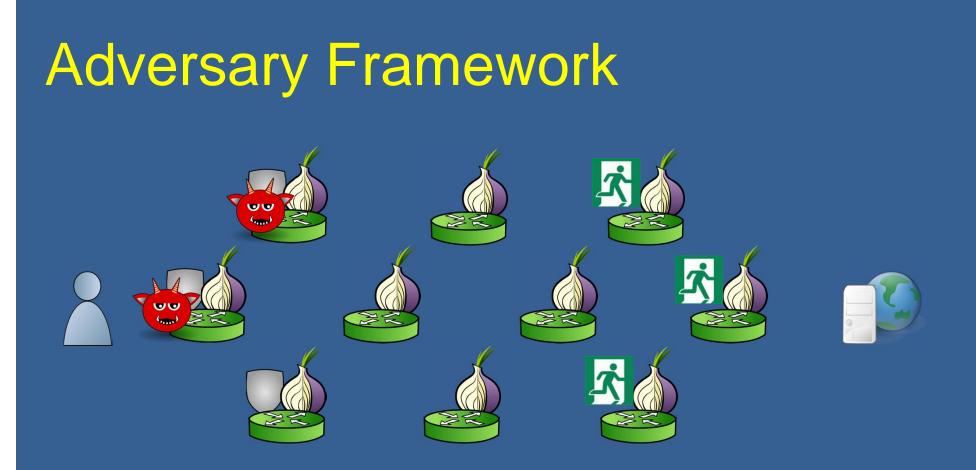
- 1. Clients begin all circuits with a selected guard.
- 2. Relays define individual exit policies.
- 3. Clients multiplex streams over a circuit.
- 4. New circuits replace existing ones periodically.

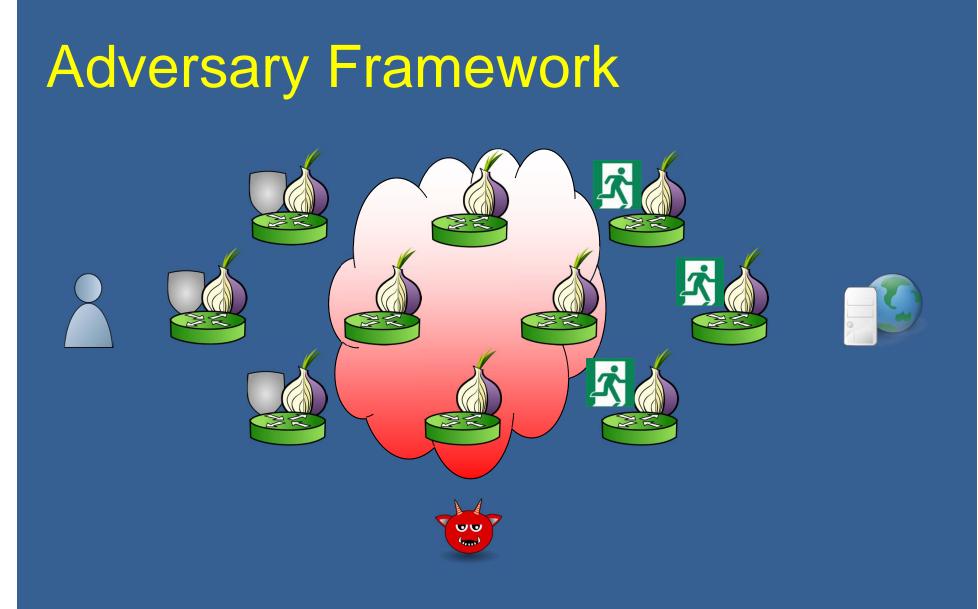
Overview

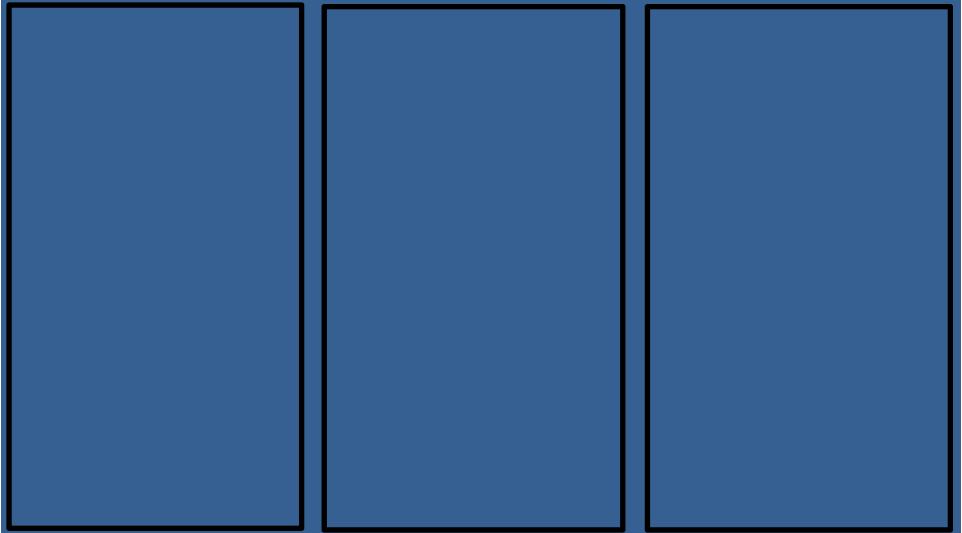
- Summary
- Tor Background
- Tor Security Analysis

 Adversary Framework
 Security Metrics
 Evaluation Methodology
 Node Adversary Analysis
 Link Adversary Analysis
- Future Work









Resource Types

- Relays
- Bandwidth
- Autonomous Systems (ASes)
- Internet Exchange Points (IXPs)
- Money

Resource Types

- Relays
- Bandwidth
- Autonomous Systems (ASes)
- Internet Exchange Points (IXPs)
- Money

Resource Endowment

- Destination host
- 5% Tor bandwidth
- Source AS
- Equinix IXPs

Resource Types

- Relays
- Bandwidth
- Autonomous Systems (ASes)
- Internet Exchange Points (IXPs)
- Money

Resource Endowment

- Destination host
- 5% Tor bandwidth
- Source AS
- Equinix IXPs

Goal

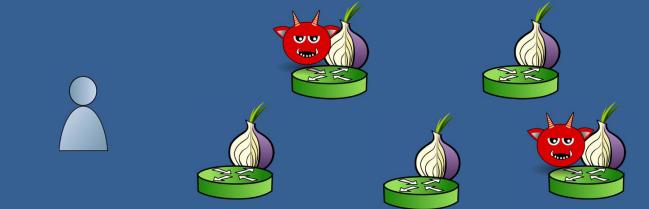
- Target a given user's communication
- Compromise as much traffic as possible
- Learn who uses Tor
- Learn what Tor is used for

Overview

- Summary
- Tor Background
- Tor Security Analysis

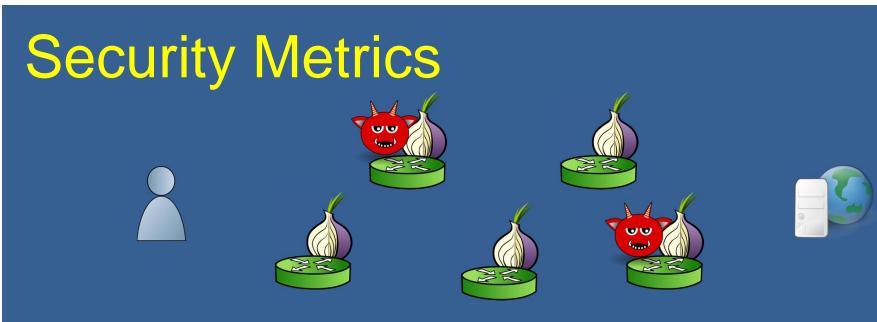
 Adversary Framework
 Security Metrics
 Evaluation Methodology
 Node Adversary Analysis
 Link Adversary Analysis
- Future Work

Security Metrics









Prior metrics

Probability of choosing bad guard and exit
 a. c² / n²: Adversary controls c of n relays
 b. ge : g guard and e exit BW fractions are bad

Prior metrics

Probability of choosing bad guard and exit

 a. c² / n²: Adversary controls c of n relays
 b. ge : g guard and e exit BW fractions are bad

 Probability some AS/IXP exists on both entry and exit paths (i.e. path independence)

Prior metrics

- 1. Probability of choosing bad guard and exit
 - a. c^2 / n^2 : Adversary controls c of n relays
 - b. ge : g guard and e exit BW fractions are bad
- 2. Probability some AS/IXP exists on both entry and exit paths (i.e. *path independence*)
- 3. g_t : Probability of choosing malicious guard within time t

Principles

- 1. Probability distribution
- 2. Measure on human timescales
- 3. Based on adversaries

Principles

- 1. Probability distribution
- 2. Measure on human timescales
- 3. Based on adversaries

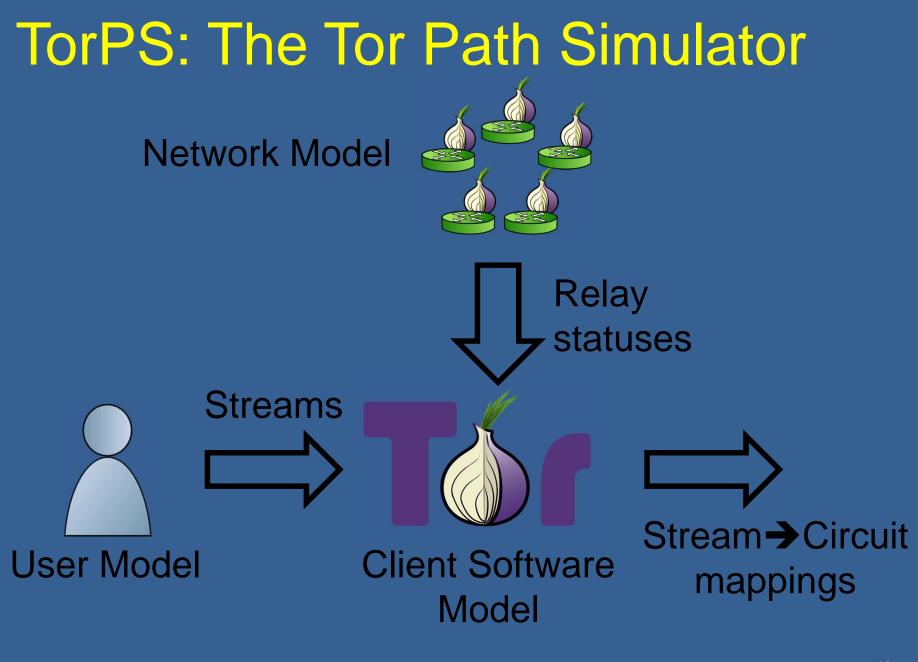
Metrics

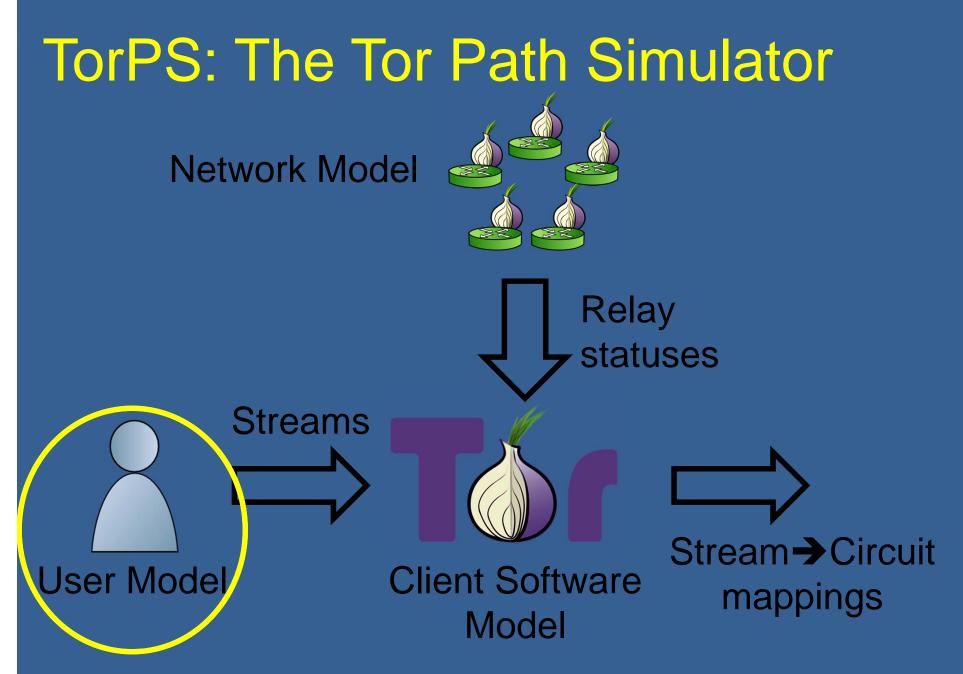
- 1. Probability distribution of time until first path compromise
- 2. Probability distribution of number of path compromises for a given user over given time period

Overview

- Background
- Onion Routing Security Analysis

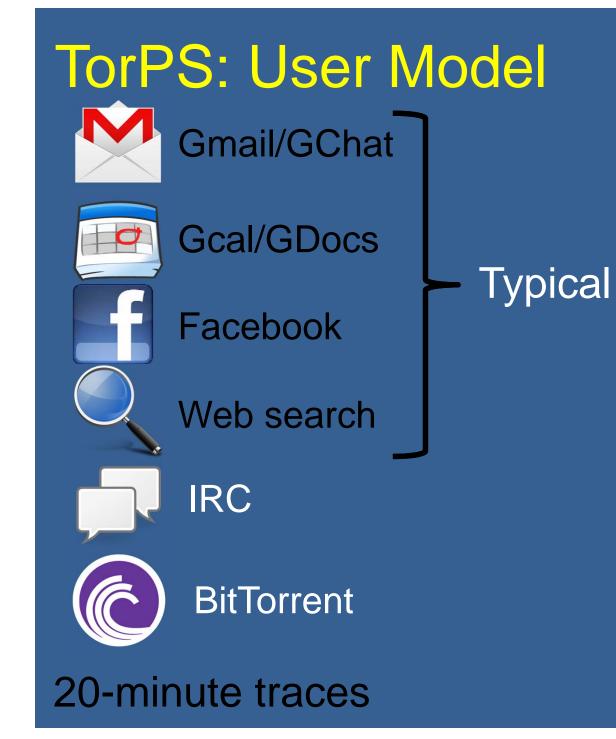
 Problem: Traffic correlation
 Adversary Model
 Security Metrics
 Evaluation Methodology
 Node Adversary Analysis
 Link Adversary Analysis
- Future Work

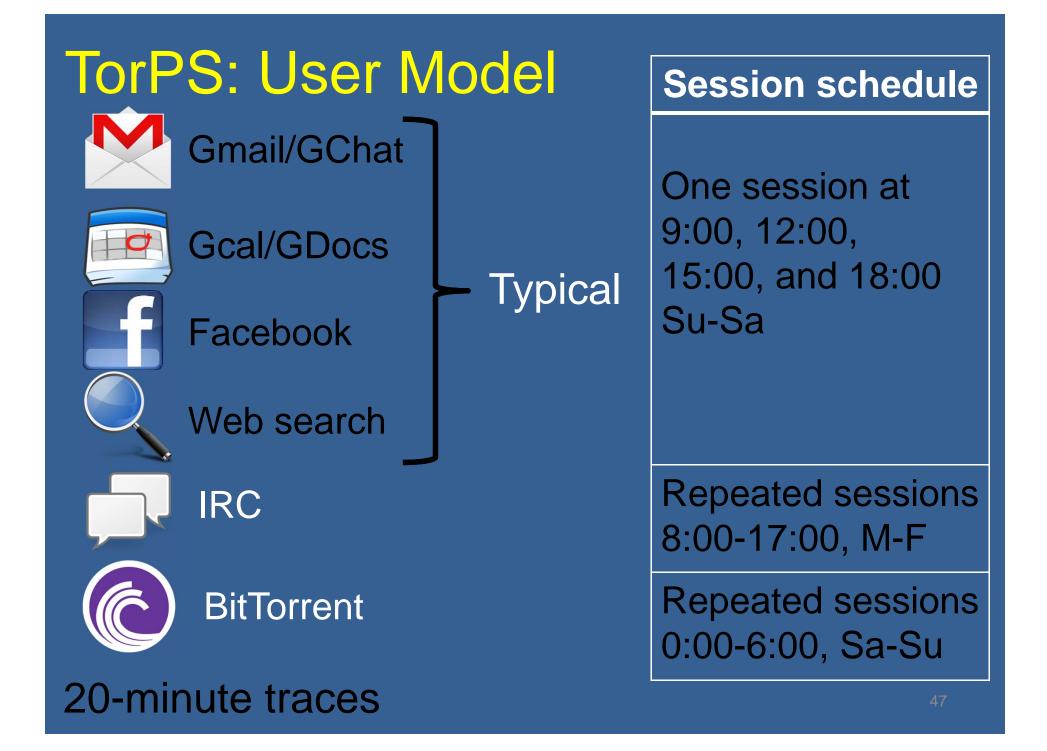




TorPS: User Model Gmail/GChat Gcal/GDocs Facebook Web search IRC **BitTorrent**

20-minute traces





TorPS: User N	lodel	Session schedule
Gmail/GChat		One session at
Gcal/GDocs	Typical	9:00, 12:00, 15:00, and 18:00
Facebook	 Typical Worst Port 	Su-Sa
Web search	(6523) Best Port	
IRC	(443)	Repeated sessions 8:00-17:00, M-F
BitTorrent		Repeated sessions 0:00-6:00, Sa-Su
20-minute traces		48

TorPS: User Model

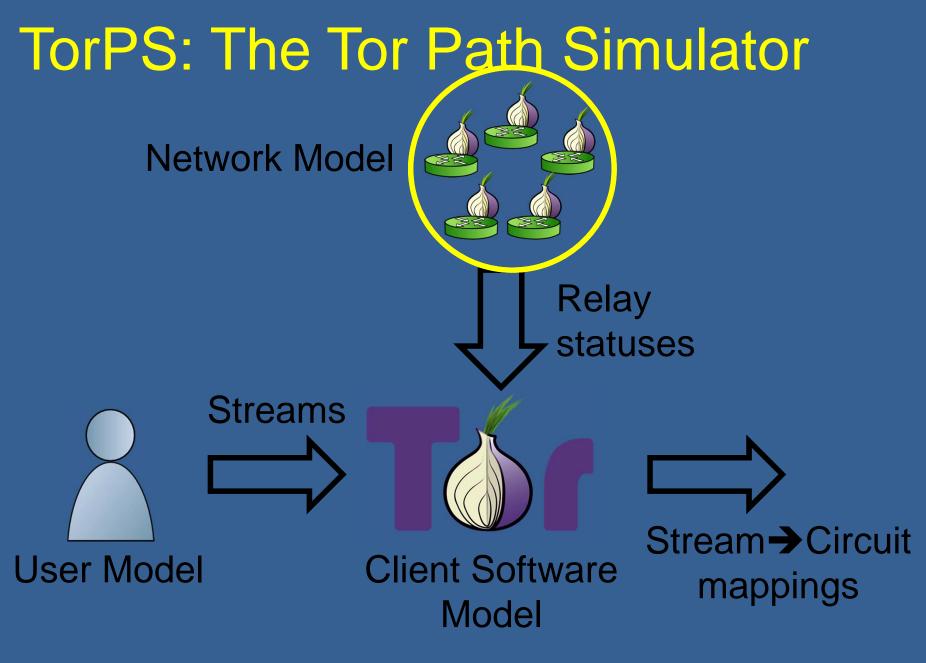
Rank	Port #	Exit BW %	Long- Lived	Application
1	8300	19.8	Yes	iTunes?
2	6523	20.1	Yes	Gobby
3	26	25.3	No	(SMTP+1)
65312	993	89.8	No	IMAP SSL
65313	80	90.1	No	HTTP
65314	443	93.0	No	HTTPS

Default-accept ports by exit capacity.

TorPS: User Model

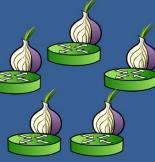
Model	Streams/wee k	IPs	Ports (#s)
Typical	2632	205	2 (80, 443)
IRC	135	1	1 (6697)
BitTorrent	6768	171	118
WorstPort	2632	205	1 (6523)
BestPorst	2632	205	1 (443)

User model stream activity



TorPS: The Tor Path Simulator

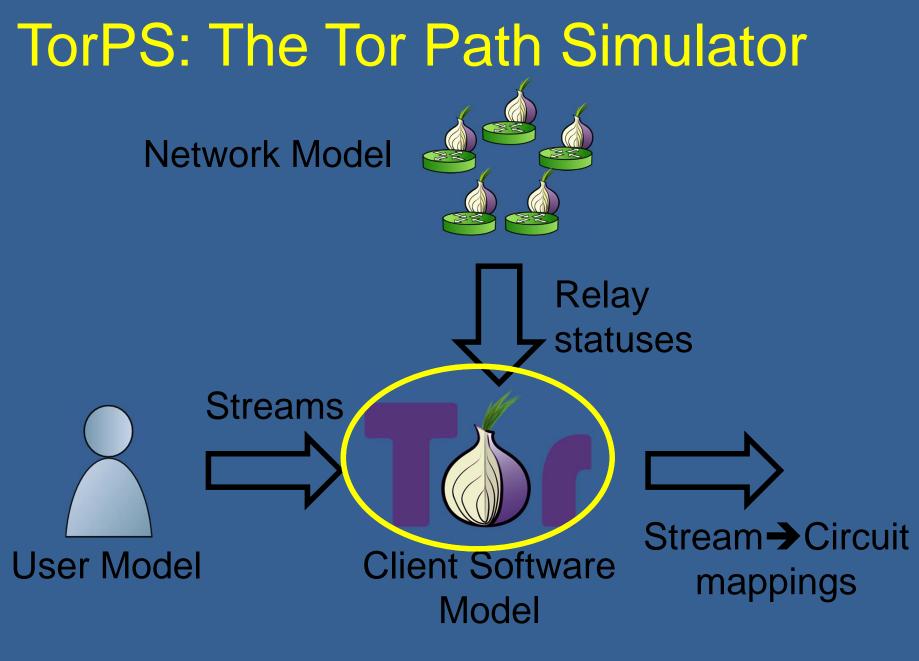
Network Model



metrics.torproject.org



Hourly consensuses Monthly server descriptors archive



TorPS: The Tor Path Simulator

Client Software Model

- Reimplemented path selection in Python
- Based on current Tor stable version (0.2.3.25)
- Major path selection features include
 - Bandwidth weighting
 - Exit policies
 - Guards and guard rotation
 - Hibernation
 - /16 and family conflicts
- Omits effects of network performance

Overview

- Background
- Onion Routing Security Analysis

 Problem: Traffic correlation
 Adversary Model
 Security Metrics
 Evaluation Methodology
 Node Adversary Analysis
 Link Adversary Analysis
- Future Work

Node Adversary

و و

100 MiB/s total bandwidth

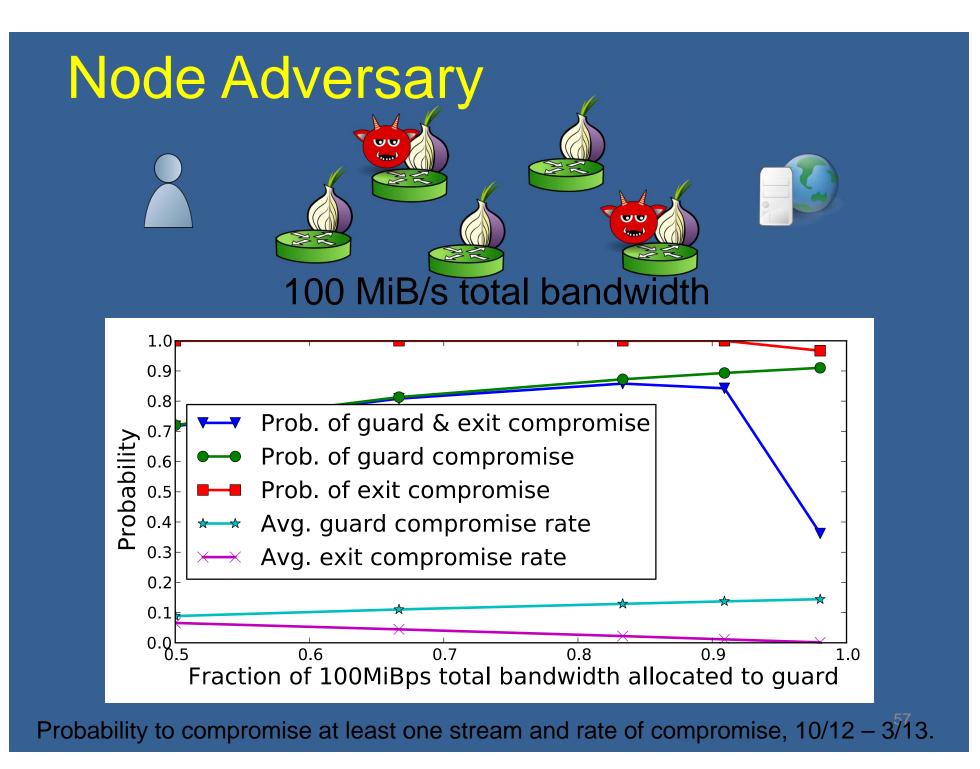
Relay Type	Number	Bandwidth (GiB/s)
Any	2646	3.10
Guard only	670	1.25
Exit only	403	0.30
Guard & Exit	272	0.98

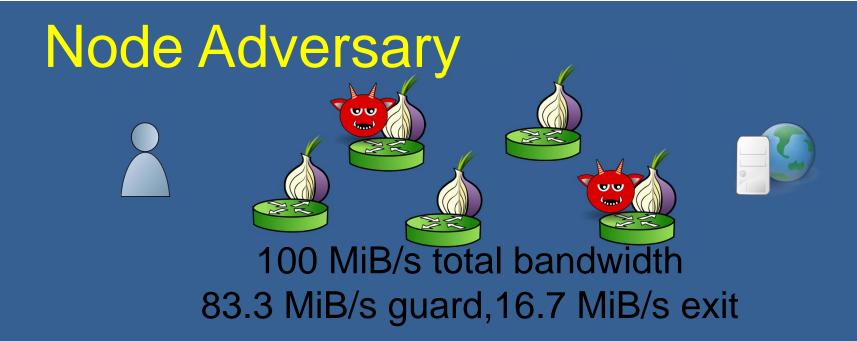
Tor relay capacity, 3/31/13

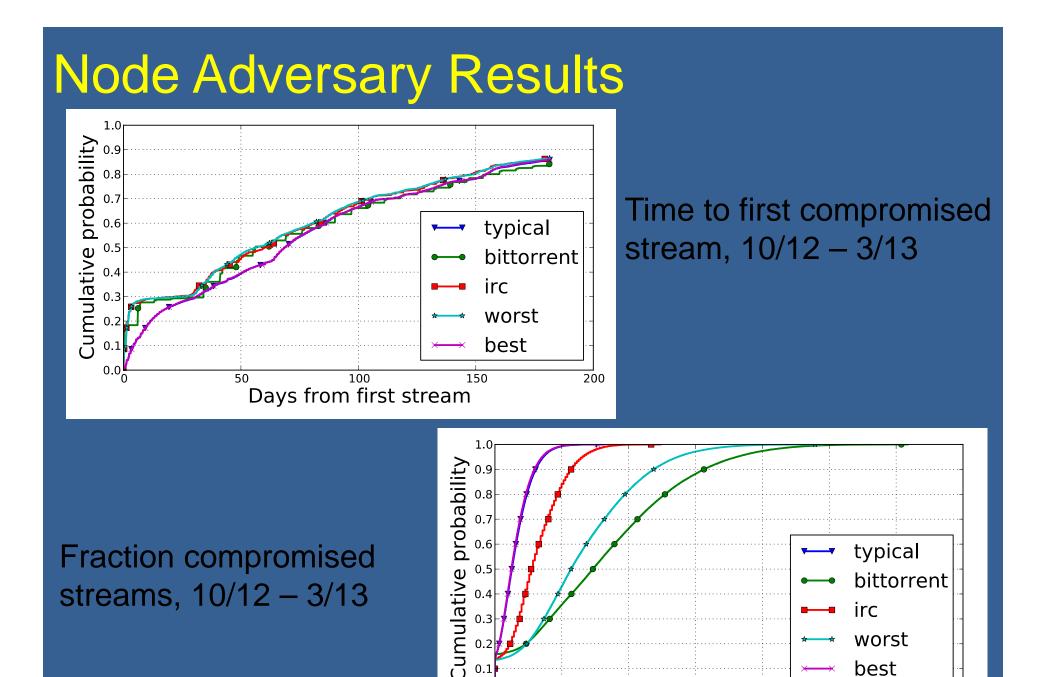
Rank	Bandwidth (MiB/s)	Family
1	260.5	torservers.net
2	115.7	Chaos Computer Club
3	107.8	DFRI
4	95.3	Team Cymru
5	80.5	Paint

 $\mathbf{\Theta}\mathbf{\Theta}$

Top Tor families, 3/31/13







0.1

0.0

0.01

best

0.06

0.07

59

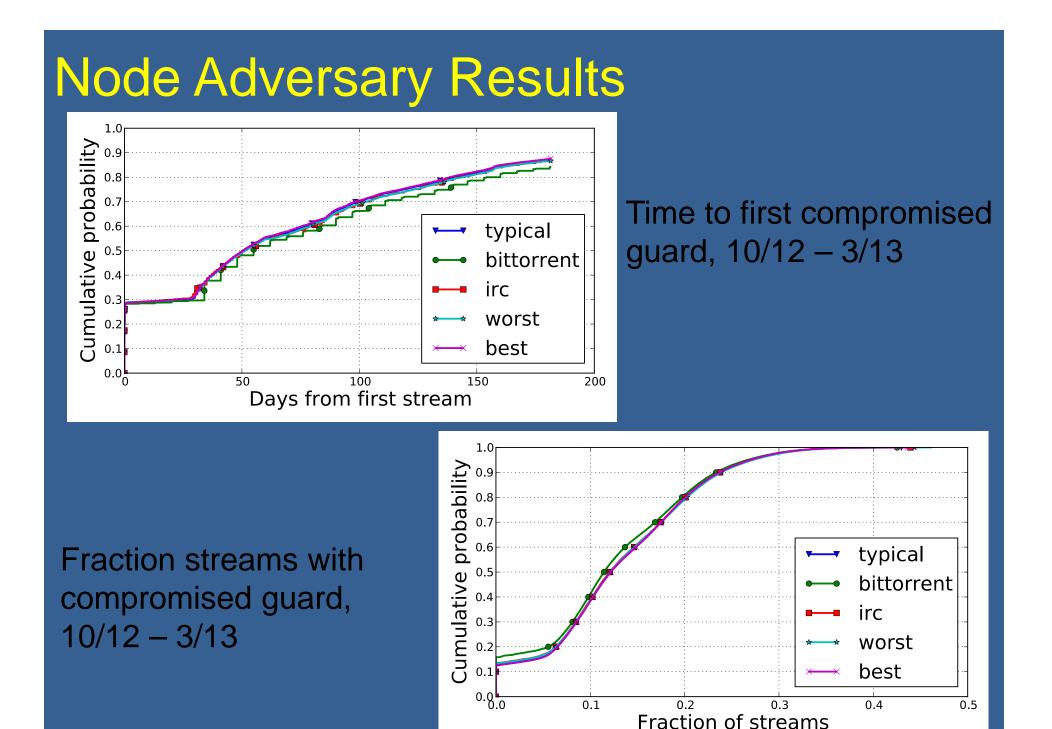
0.05

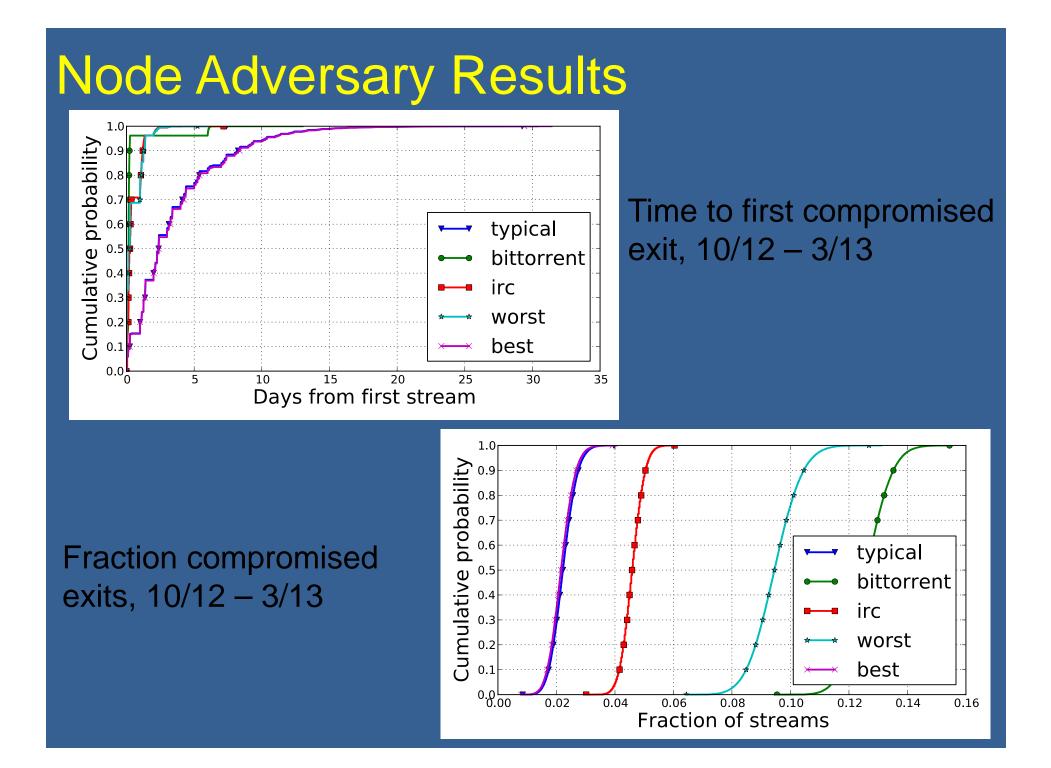
0.03

0.04

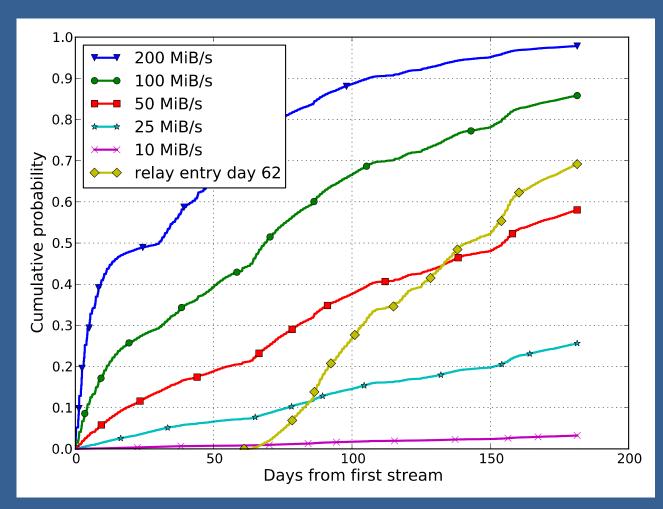
Fraction of streams

0.02





Node Adversary Results



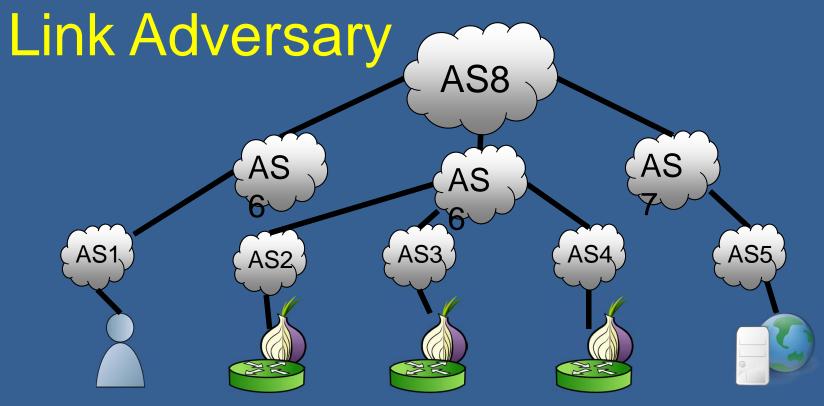
Time to first compromised circuit, 10/12-3/13

Overview

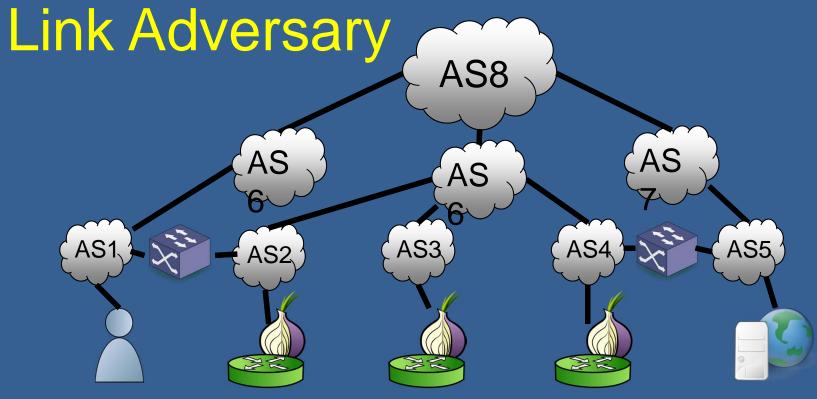
- Background
- Onion Routing Security Analysis

 Problem: Traffic correlation
 Adversary Model
 Security Metrics
 Evaluation Methodology
 Node Adversary Analysis
 Link Adversary Analysis
- Future Work

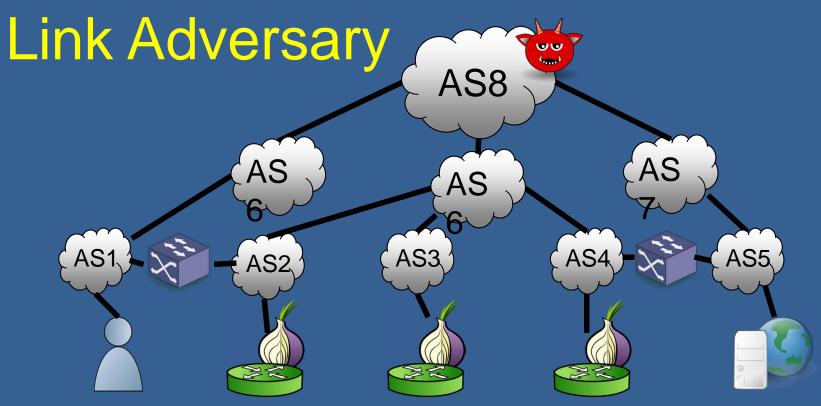




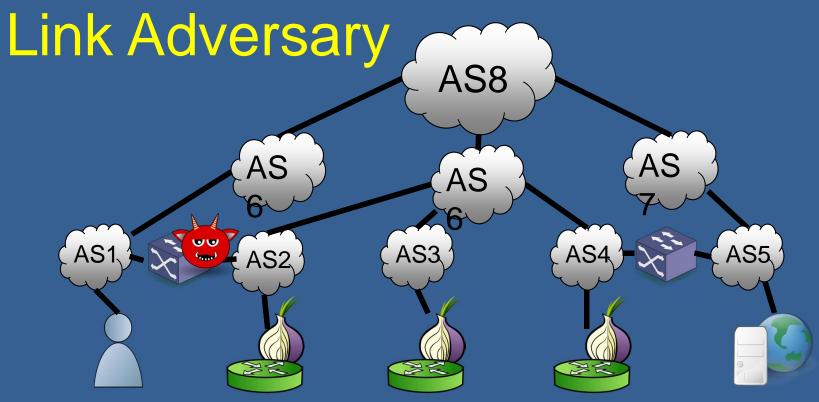
1. Autonomous Systems (ASes)



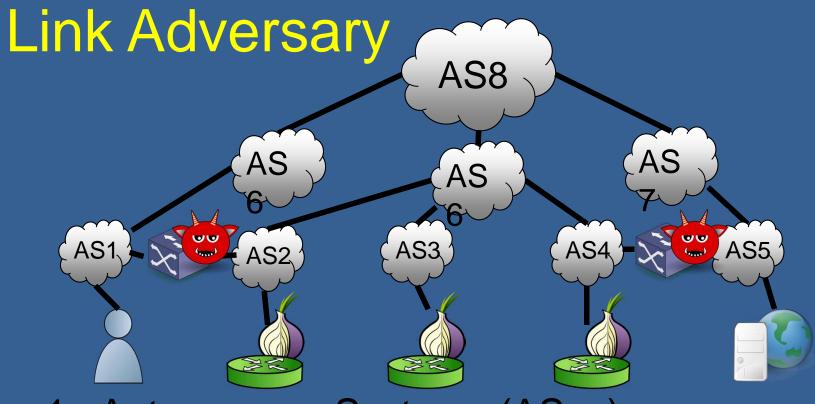
- 1. Autonomous Systems (ASes)
- 2. Internet Exchange Points (IXPs)



- 1. Autonomous Systems (ASes)
- 2. Internet Exchange Points (IXPs)
- 3. Adversary has fixed location



- 1. Autonomous Systems (ASes)
- 2. Internet Exchange Points (IXPs)
- 3. Adversary has fixed location



- 1. Autonomous Systems (ASes)
- 2. Internet Exchange Points (IXPs)
- 3. Adversary has fixed location
- 4. Adversary may control multiple entities
 - a. "Top" ASes
 - b. IXP organizations

Client locations

 Top 5 non-Chinese source ASes in Tor (Edman&Syverson 09)

AS/IXP Locations

- Ranked for client location by frequency on entry or exit paths
- Exclude src/dst ASes
- Top k ASes /top IXP organization

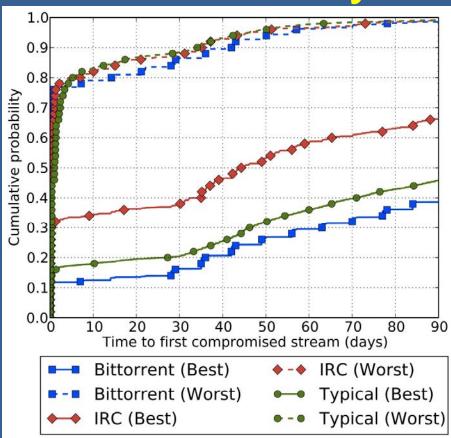
AS#	Description	Country
3320	Deutsche Telekom AG	Germany
3209	Arcor	Germany
3269	Telecom Italia	Italy
13184	HanseNet Telekommunikation	Germany
6805	Telefonica Deutschland	Germany

Туре	ID	Description
AS	3356	Level 3 Communications
AS	1299	TeliaNet Global
AS	6939	Hurricane Electric
IXP	286	DE-CIX Frankfurt
IXP Org.	DE-CIX	DE-CIX

Example: Adversary locations for BitTorrent client in AS 3320

IXP organizations obtained by manual clustering based on PeerDB and PCH.

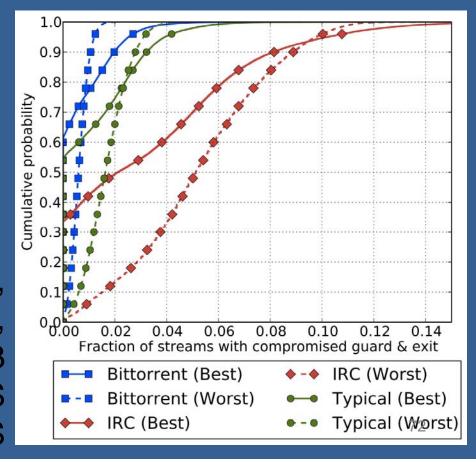
#	IXP Organization	Size	Country
1	Equinix	26	global
2	PTTMetro	8	Brazil
3	PIPE	6	Australia
4	NIXI	6	India
5	XChangePoint	5	global
6	MAE/VERIZON	5	global
7	Netnod	5	Sweden
8	Any2	4	US
9	PIX	4	Canada
10	JPNAP	3	Japan
11	DE-CIX	2	Germany
12	AEPROVI	2	Equador
13	Vietnam	2	Vietnam
14	NorthWestIX	2	Montana, US
15	Terremark	2	global
16	Telx	2	US
17	NorrNod	2	Sweden
18	ECIX	2	Germany
19	JPIX	2	Japan
IXP organizations ranked by size ⁷¹			

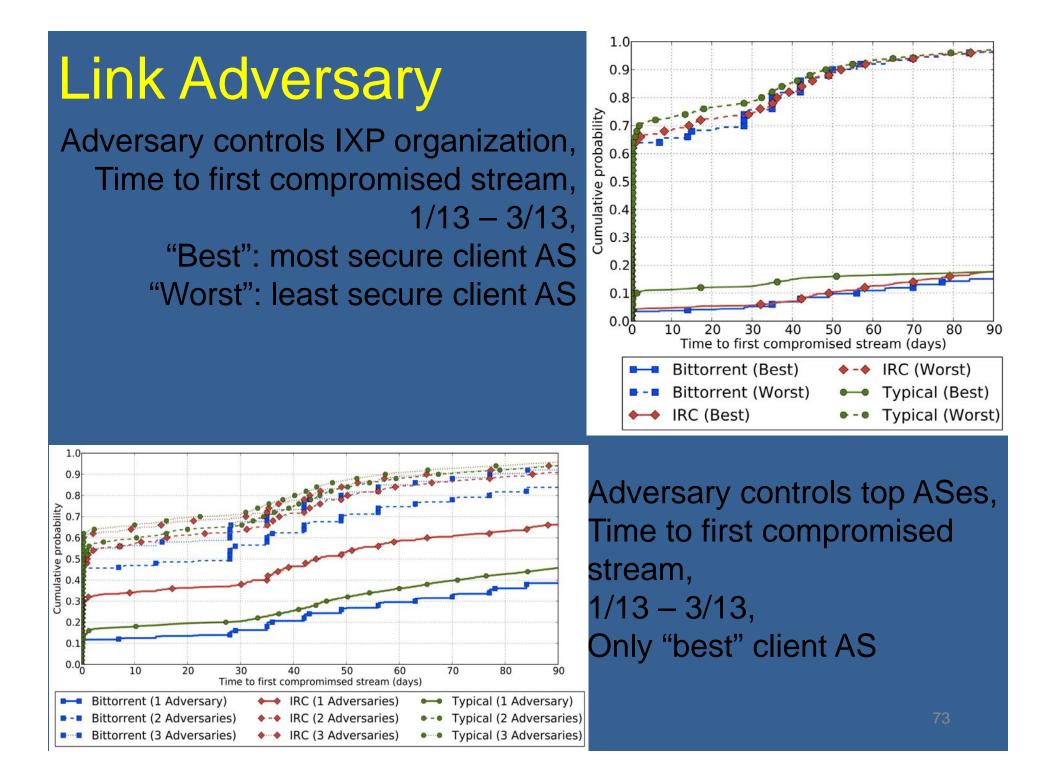


Adversary controls one AS, Fraction compromised streams, 1/13 – 3/13

"Best": most secure client AS "Worst": least secure client AS

Adversary controls one AS, Time to first compromised stream, 1/13 – 3/13 "Best": most secure client AS "Worst": least secure client AS





Overview

- Background
- Onion Routing Security Analysis Problem: Traffic correlation Adversary Model Security Metrics Evaluation Methodology Node Adversary Analysis o Link Adversary Analysis • Future Work

Future Work

- 1. Extending analysis
- 2. Improving guard selection
- 3. Using trust-based path selection to protect against traffic correlation
- 4. Dealing with incomplete and inaccurate AS and IXP maps
- 5. Include Tor's performance-based pathselection features in TorPS