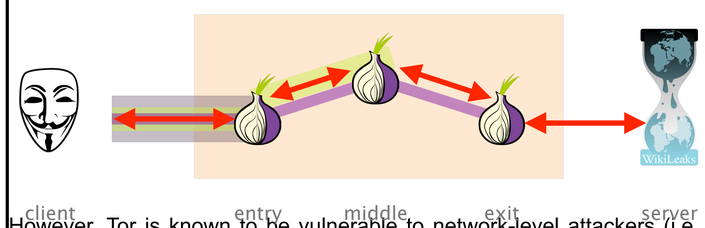
RAPTOR: Routing Attacks on Privacy in Tor

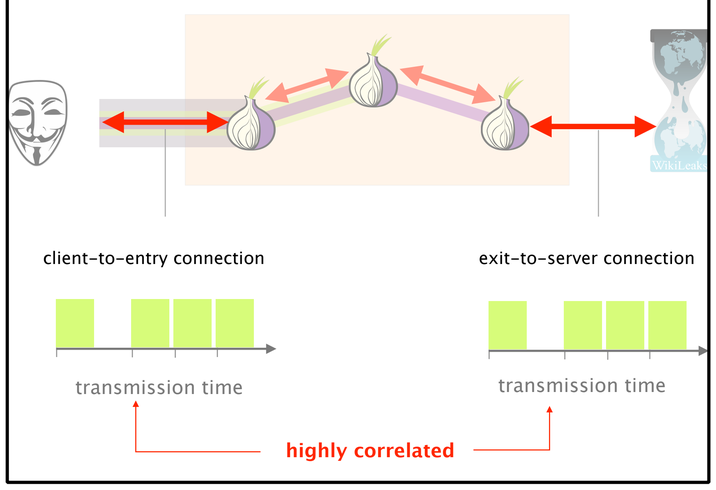
Yixin Sun
Anne Edmundson
Laurent Vanbever
Oscar Li
Jennifer Rexford
Mung Chiang
Prateek Mittal

Tor Background

The Tor network is a widely used system for anonymous communication. Tor protects user's anonymity by bouncing traffic through a network of relays between the client and the destination server. The Tor client initiates the circuit by selecting three relays, one of each type: guard/ entry relay, middle relay and exit relay.



However, Tor is known to be vulnerable to network-level attackers (i.e., Autonomous Systems) who can observe both ends of the communication path. The traffic entering and leaving Tor are highly correlated, and thus are vulnerable to traffic correlation attacks.



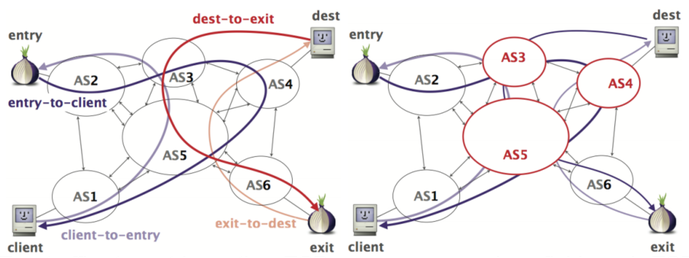
Raptor Attacks

We present three new attacks, called Raptor attacks, that deanonymize Tor users more effectively than previously thought possible. To do so, Raptor leverages the dynamic aspects of the internet routing protocol, i.e., the Border Gateway Protocol (BGP). Raptor attacks are composed of three individual attacks whose effects are compounded, as shown below.

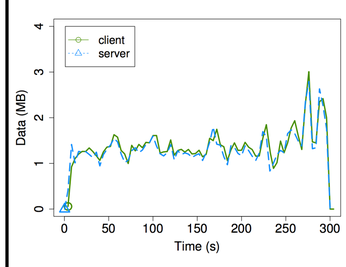
	Asymmetric Traffic Analysis	*BGP Churn*	*BGP Hijack/ Interception*
Symmetric	Known	***Novel***	***Novel***
Asymmetric	***Novel***	***Novel***	***Novel***

Asymmetric Traffic Analysis

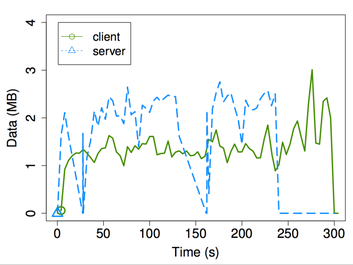
Conventional traffic analysis considers only one scenario where adversaries observe data traffic in one direction. However, Internet paths are often asymmetric: the path from the exit relay to the Web server may be different than the path from the Web server to the exit relay. Routing asymmetry increases the number of ASes who can observe at least one direction of traffic at both endpoints.



Tor traffic to retrieve the TCP sequence number field and TCP acknowledgment number field, and analyzes the correlation between these fields of both ends over time. Thus, it allows an adversary to deanonymize users as long as the adversary is able to observe any direction of the traffic, at both ends of the communication.



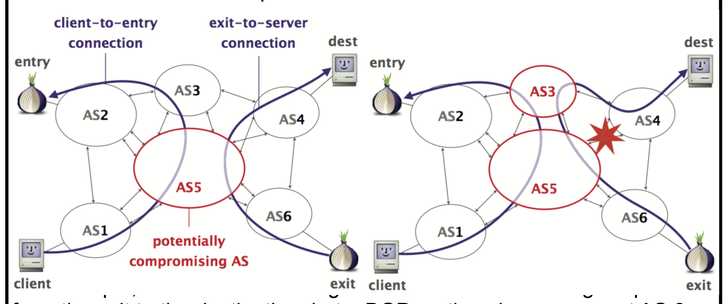
Left: Example of correlated client/server pair using TCP ACKs



Right: Example of uncorrelated client/server pair using TCP ACKs

BGP Churn

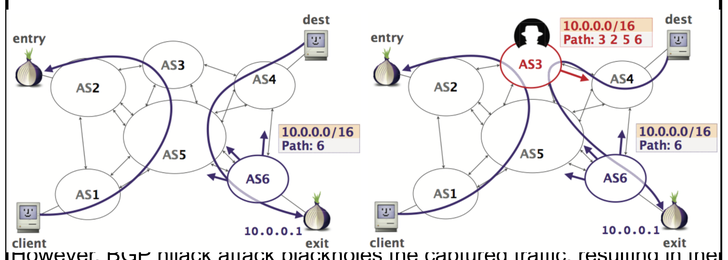
The underlying Internet paths between a client and guard relay vary over time due to changes in the physical topology (e.g., failures, recoveries, and the rollout of new routers and links) and AS-level routing policies (e.g., traffic engineering and new business relationships). These changes give a malicious AS surveillance power that increases over time.



from the exit to the destination, but a BGP routing change can put AS 3 on the path, allowing it to perform traffic analysis.

BGP Hijack/Interception

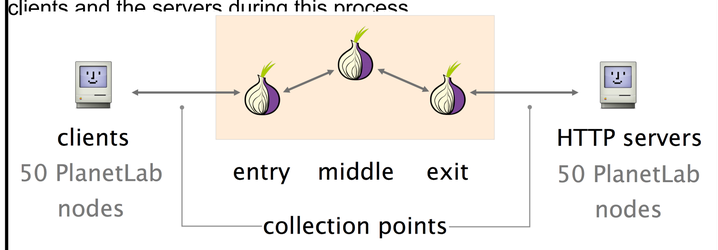
AS-level adversaries can hijack an IP prefix by advertising the prefix as its own. The attack causes a fraction of Internet traffic destined to the prefix to be captured by the adversary. Tor relay nodes can observe a large amount of client traffic. For example, a Tor guard relay observes information about client IP addresses. Thus, the IP prefixes corresponding to Tor guard and exit relays presents an attractive target for BGP hijack.



However, BGP hijack attack blackholes the captured traffic, resulting in the connection being dropped. BGP interception attack, on the other hand, allows the malicious AS to become an intermediate AS in the path towards the guard relay, i.e., after interception, the traffic is routed back to the actual guard relay. Such an interception attack allows the connection to be kept alive, enabling the malicious AS to exactly deanonymize the client via asymmetric traffic analysis.

Asymmetric Traffic Analysis Experiment

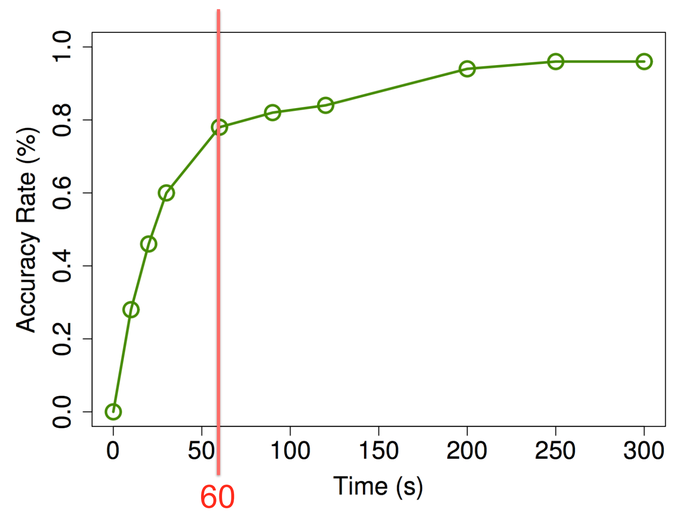
We set up our experiment on PlanetLab nodes, with 50 Tor clients and 50 HTTP servers. We launch wget requests on the 50 clients at the same time, each requesting a 100MB image file from one of the 50 web servers, respectively. We use tcpdump to capture data for 300 seconds at the clients and the servers during this process.



By correlating traffic captured at both ends, we were able to deanonymize ~95% of the pairs - with no false positives.

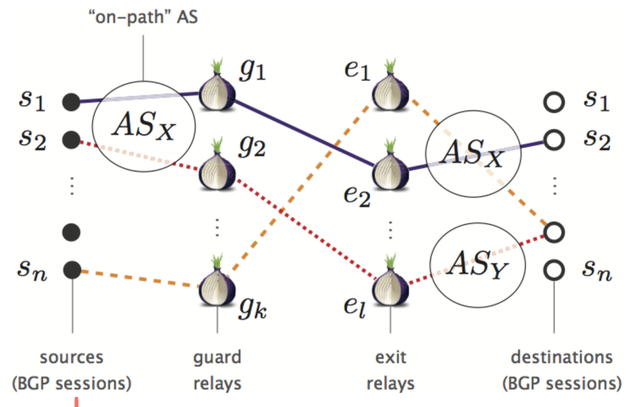
	Client ACK/ Server ACK	Client ACK/ Server Data	Client Data/ Server ACK	Client Data/ Server Data
Overall	96%	94%	96%	94%
False negative	4%	6%	4%	6%
False positive	0%	0%	0%	0%

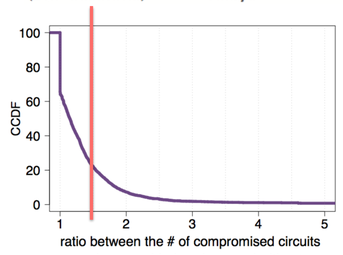
The detection accuracy rate quickly increases with time, reaching ~80% within only a minute.



BGP Churn Measurement

Churn significantly increases the number of compromising ASes. We measured the effect of churn by collecting BGP updates for 1 month in Jan 2015. We considered each BGP session as a Tor user or destination. On each session, we computed the ASes used to reach each entry and exit relays. An AS is compromising when it ends up simultaneously on a (src, entry) and (exit, dest) path.

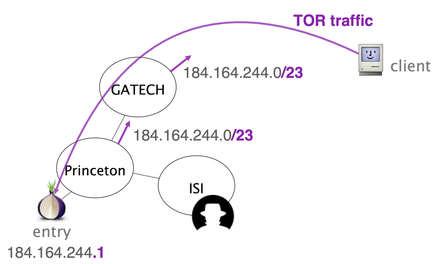




When considering churn, 60% of the pairs (src, dst) sees an increase of compromised circuits, and 20% of the pairs sees an increase of more than 50%.

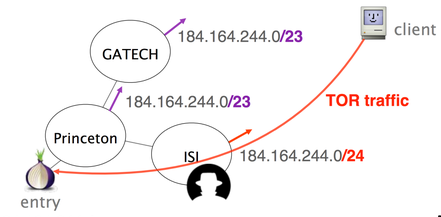
BGP Interception Experiment

We successfully performed a BGP attack on an existing Tor relay. Note that our experiments did not compromise the privacy or safety of real Tor users. We attacked our own traffic and our own relay.



We hosted an entry relay in Princeton, and advertised the covering IP prefix via GATECH. GATECH relayed on prefix to the entire Internet, and Tor traffic started to flow.

After 20 sec, we announced a more-specific prefix via ISI. As forwarding is based on the longest-match, all traffic soon started to enter via ISI.

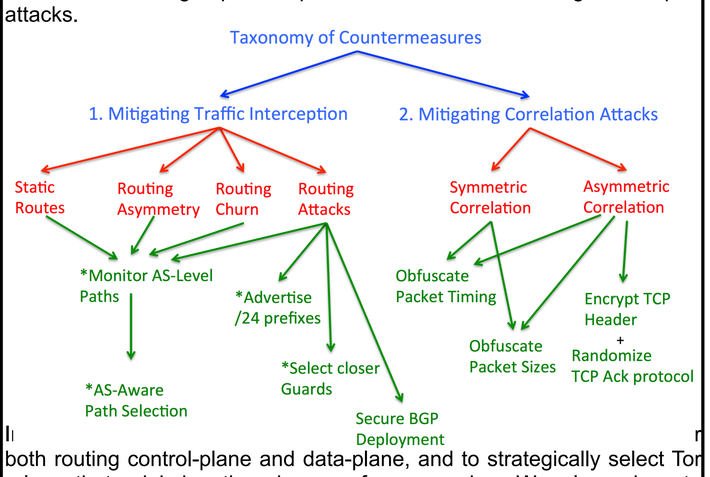


By correlating TCP ACK collected at the HTTP servers, we were able to deanonymize 90% of the pairs.

	Accuracy Rate	False Negative	False Positive
Client ACK/Server ACK	90%	8%	2%

Countermeasures

There are two main categories of countermeasures: migrating traffic interception and mitigating correlation attacks. The figure below illustrates the design space of potential countermeasures against Raptor attacks.



both routing control-plane and data-plane, and to strategically select Tor relays that minimize the chance of compromise. We also advocate defenses that aim to detect and prevent routing attacks. We do not focus on the class of approaches that aim to mitigate correlation analysis by obfuscating packet sizes and timings, as they are generally considered too costly to deploy. More specifically, we have been working on deploying real-time BGP monitoring system and Traceroute monitoring system to detect routing anomalies. Furthermore, we are working on optimization of guard relay selection to take into account AS-level topology.