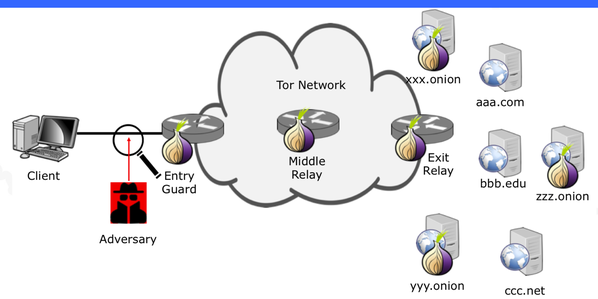
A Pilot Study on Website Fingerprinting Vulnerability of Tor Onion Services & General Websites

Loc Ho[1], Young-Ho Kim[2], Won-gyum Kim[3], Donghoon Kim[1], Doosung Hwang[2]
[1]Arkansas State University, [2]Dankook University, [3]AiDeep

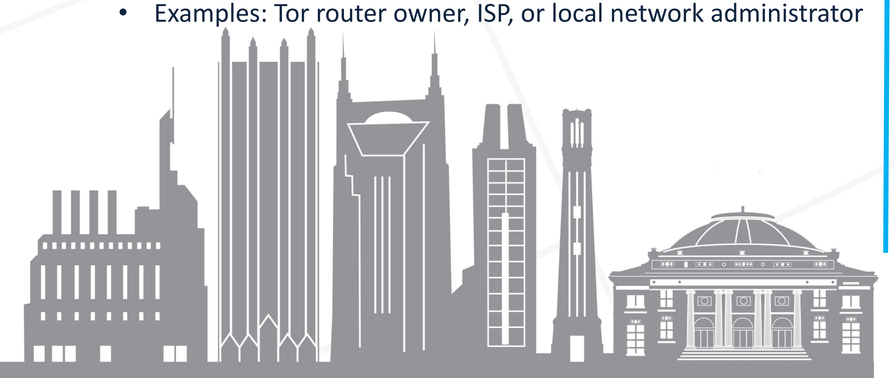
Motivation

- Website fingerprinting attacks have exposed a vulnerability in Tor network.
- Tor browsers can access both onion services and general websites.
- The onion (hidden) services are services that can only be accessed over Tor while the general websites (non-hidden services) that can also be accessed with regular web browsers.
- However, there has not been much research conducted on how secure the onion services are in the website fingerprinting attack compared to the general websites.

Threat Model



- An adversary is able to observe the network traffic from a client to the entry Tor router (entry guard) and the traffic from the exit Tor router to a destination client to de-anonymize the connection
- Examples: Tor router owner, ISP, or local network administrator



Data & Features

- 10 onion services and 10 general websites
- Due to the simplicity of the design of the onion services, 10 general websites were also chosen to be made with relatively simple designs
- Each service has 150 instances

Studied	Feature	No.
Oh *et al.* [5]	Packet general information (44) Cell sequence length (4) Packet inter arrival time (27) Burst information (24) Cell ordering (18) Concentration (8)	125

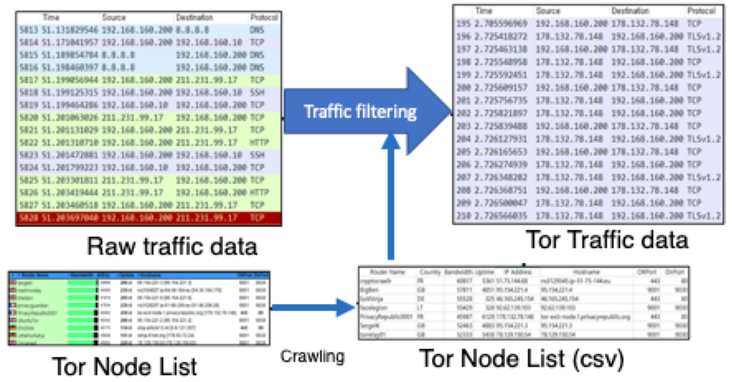
Table 1: Feature Vectors

Approach

To find out the vulnerability of Tor onion services and genera websites, we implemented a framework.

❑The framework can collect network traffic

❑The framework can filter the collected traffic to consist only of Tor-related traffic



Analysis

- The classification was conducted with Decision Tree, Random Forest, and XGBoost.
- Table 2 shows the results with several classification experiments with 125 features

			non-cutting	50 cutting
Decision Tree	10 general	Accuracy	0.7393	0.7171
		Time	0.072	0.066
	10 onion	Accuracy	0.7454	0.7272
		Time	0.062	0.066
Random Forest	10 general	Accuracy	0.8080	0.7616
		Time	0.050	0.050
	10 onion	Accuracy	0.8101	0.7797
		Time	0.046	0.047
XGBoost	10 general	Accuracy	0.8262	0.8424
		Time	0.960	1.373
	10 onion	Accuracy	0.8202	0.8202
		Time	0.885	1.088

Table 2: Website fingerprinting comparison (Training Time (sec))

- When using XGboost, 10 onion services and 10 general websites have the highest accuracy of 82.02% and 82.62%, respectively.
- Even with decision trees and random forest, we found that for 10 onion services and 10 general websites, the difference in accuracy was not significant.
- The initial 30 packets are known as important features because they contain important information about protocols and websites. However, this is a feature common to websites, especially for onion services so it may not that important for onion services

Conclusion

- With the same features, there is no significant difference in accuracy.
- The initial packet does not significantly affect.
- Tor network is vulnerable not only to general websites and but also to onion services if important features are extracted

HOTSOS 2021
8TH ANNUAL
HOT TOPICS *in the* SCIENCE OF SECURITY
APRIL 13-15, 2021 | VIRTUAL