Fingerprinting Attacks on Anonymity Systems

Bc. Martin Krajči
Supervisor: Mgr. Kamil Malinka, Ph.D.

Motivation

- the existence of multiple attacks on anonymity systems, which might deanonymize the users
- fingerprinting attack is one of the most powerful, which many previous studies proved to be very dangerous, however, conditions of their experiments were unrealistic
- the necessity to test this attack under more realistic conditions to verify severity of the attack

Datasets

700 websites from Alexa top websites were used to create datasets. The first one was created with Chrome browser in default configuration, second was created by browsing the same websites from different country and third dataset was created with Adblock turned on and configured Spanish locale. Tor source codes alteration was used to bring new, more precise concept for data gathering. Gathered data was then used for training and testing of N-shot learning classification method.

Attack methodology

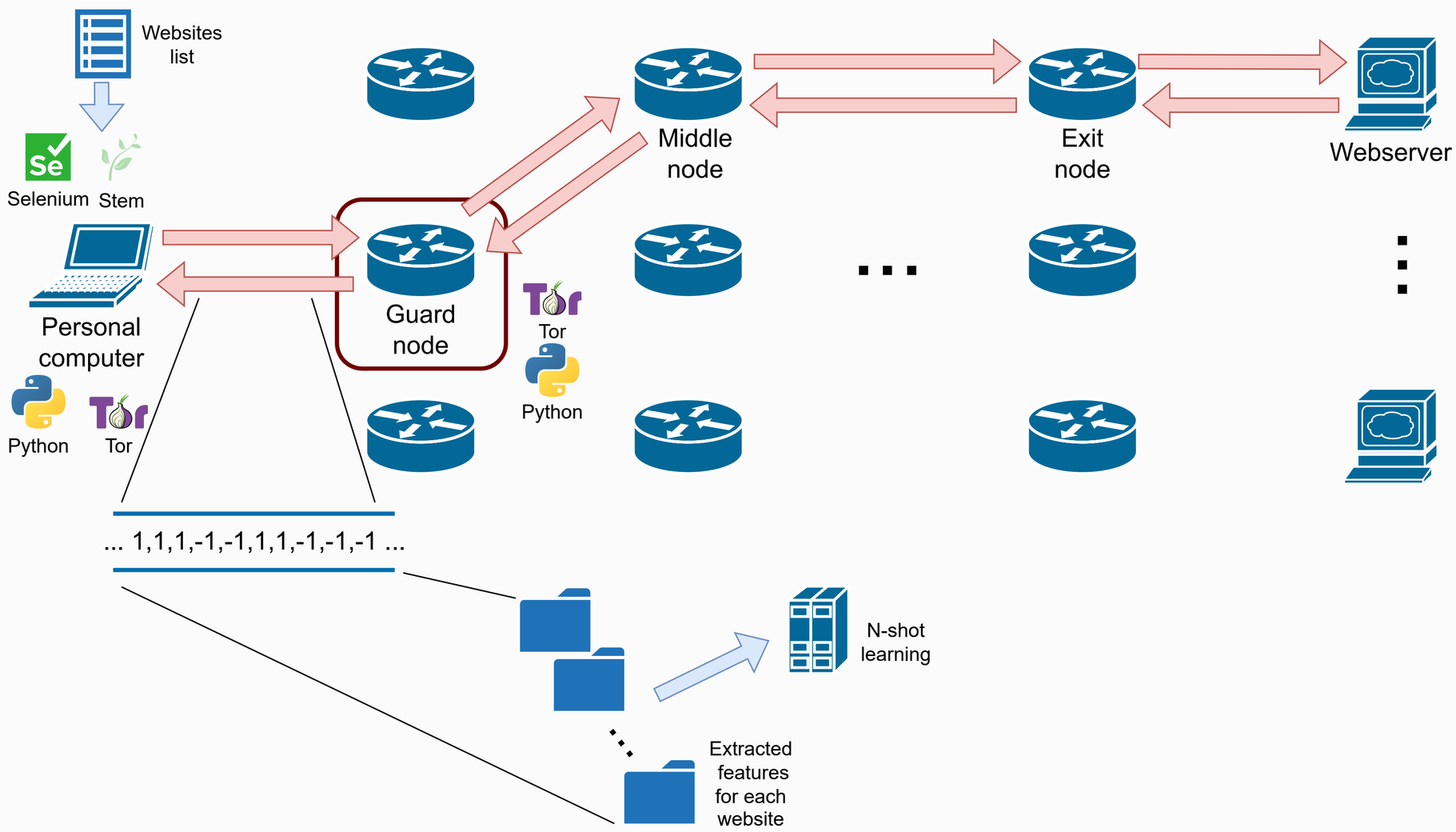

Figure 1: Diagram of entities and data flows required for succesfull attack

Results

	AWF700	AWF700_geo	AWF695_adblock
AWF700	92%	38,58%	39,65%
AWF775	71,14%	32,25%	34,19%

Table 1: Accuracy of classification for different combinations of datasets

The results show that when N-shot learning is trained and tested with very similar data, classification accuracy can be alarmingly high. However, training the classifier with data obtained from one configuration and testing it on data obtained from different configuration, can cause a reduction of the accuracy by more than half.