# Deep Fingerprinting: UnderminingWebsite Fingerprinting Defenses with Deep Learning

Paper Summary

Ruturaj Kiran Vaidya
University of Kansas

## 1 Summary

Fingerprinting attacks are shown effective against the tor browser, which has been considered as one of the safest browsers. Tor is vulnerable to traffic analysis attacks and website fingerprinting is particularly a traffic analysis attacks, which breaks the tor security. WTF-PAD and Walkie-talkie are existing defenses (the primary strategy is to add dummy packages) against these types of attacks. Paper proposes "Deep fingerprinting", a website fingerprinting attack against tor, which is based on convolutional neural networks (CNN). This attack is designed using deep learning methods, which as authors claimed outperform traditional machine learning techniques. Authors experimented in close world (95 sites), as well as open world (20000 sites) settings. Authors studied SDAE, AWF, which helped them to build a sophisticated architecture and tune the model's hyper-parameters that fit well. They evaluated this attack against WTF-PAD and Walkie-talkie, which shows effectiveness against WTF-PAD with over 90% accuracy and against Walkie-talkie with 49.7% accuracy, highlighting the need of effective defenses is tor against this attack.

## 2 Contribution

The paper presents "Deep fingerprinting", a website fingerprinting attack against tor, attaining over 98% accuracy on Tor traffic without defenses. The attack uses deep learning which is considered more sophisticated than machine learning methods. In the open world settings, the attack showed 0.99 precision and 0.94 recall on undefended traffic. They also provided some alternative defenses.

## 3 Strengths

I liked the description provided by the authors about CNN with easy to understand architecture diagram. I also liked that they used deep learning techniques to build the attack. DF model has varying values of hyper-parameters for different layers, it shows no evidence of over-fitting in their experiments and performs better that AWF model. I also liked that they tested it in close world, as well as open world settings.

## 4 Limitations

DF showed significantly more training cost. Also the gap between close world and open world performance is larger. Also, in general paper introduces an attack and provide solid defenses against it (I may be wrong, but this is my general observation), but this paper doesn't do so.

## 5 Future Work

Authors mentioned that because of the large amount of training data and large number of hyper-parameters, model has an exhaustive search prohibitive of computational resources. Thus for attacks, they aim at good-enough classifier and claim that their model can be optimized further. And then future work can be done on building defenses against this type of attack.