Walkie Talkie: An efficient defence against passive website fingerprinting attacks - Write-up*

Avishek Mondal

I. How does the defence work?

Walkie Talkie (WT) is a defence against passive website fingerprinting (WF) attacks that has 2 main thrusts.

A. Half-duplex communication

This is how the name "Walkie-Talkie" came about, because it is similar to how people using a pair of walkie-talkies would communicate. Client queues up all request messages and only sends them once the server has completed all previous requests. Effectively there's only ever one-way communication with no overlap, and has to use the optimistic data method to achieve the required functionality. The advantage of using half-duplex communication is to reduce the information available to an adversary about a cell sequence to just the information about the burst sequence. This makes the next step, burst molding, easier. It also has the advantage that bursts are easier to make identical and thus beat ML attacks, which is what the next step does.

B. Burst molding

The main purpose of burst molding is to simulate the simultaneous loading of multiple pages without having to load multiple pages. This is achieved by splitting long bursts into separate sequences, and adding fake cells to short bursts to maintain a constant "supersequence" length. Choosing an appropriate "decoy" pages is up to the client, who can choose popular websites to cover for their access to the actual site. Choosing a more popular site can reduce the effectiveness of the attack, because probabilisticaly, an attacker trying to claim the client visited a non-popular sensitive page is likely to be wrong.

II. What are the strengths and weaknesses of the paper?

A. Strengths

1) Deployability: The authors have implemented their solution in an already deployed system and demonstrated its effectiveness. The changes to existing infrastructure is minimal - only 143 lines of code to the existing Tor browser. The bandwidth and time overhead is also minimal. This is good from a security point of view, because if user experience is not eroded too much, more users are likely to use the system which leads to greater security through a larger anonymity set as compared to a more theoretically secure system with a worse user experience that would lead to fewer users.

2) Performance against attacks: As the authors themselves have demonstrated, WT holds up to attacks, showing the proof of their concept. A later paper, [2] has also validated WT's ability to stand up to an attack that uses a deep-learning classifier to deploy WF attacks on Tor. It is also superior to other defences that have been proposed

B. Weaknesses

1) Threat model: The threat model assumed for this defence is that of an adversary with visibility over the client-guard connection. This is a weaker threat model than that envisioned by the original proposers of the Tor network, where an adversary has visibility over both client-guard and exit-destination. If a stronger adversary is considered, it is not difficult to see that WF can be carried out by correlating the metadata of the super-sequences and deanonymising clients and receipients.

2) Nature of sites: WT only defends for access to static sites and not dynamic ones. In a real-world setting, this is not a realistic assumption.

3) Choice of decoy sites: The client is given the freedom to choose which decoy site to choose. The authors use the example of a German speaking client using a popular German page to use as a decoy. This could be a weakness. As the authors themselves note, choosing sensitive pages as decoys is necessary for this defence to work; otherwise an adversary can simply infer the non-sensitive page is always a decoy. If left to lay users who may be journalists, whistleblowers etc. who are not trained in the nuances of the defence, it is likely that only non-sensitive pages are always chosen as decoys, and the adversary has an advantage.

4) Scaling of burst modelling: This criticism is somewhat covered in [2] as well. To accurately build the "super-sequence" to beat the adversary, clients need to know the up-to-date padding patterns of what the burst sequence would be of decoy sites. This has to be distributed to all clients, and be specific to their local network conditions. This means a central directory that maintains and updates a list of plausible decoy websites and their corresponding burst sequences would be required and has to distribute this to all the clients. This could hamper the scalability of this defence.

*Paper is [1]

References

[1] Tao Wang and Ian Goldberg. "Walkie-talkie: An efficient defense against passive website fingerprinting attacks". In: *26th USENIX Security Symposium (USENIX Security 17)*. 2017, pp. 1375–1390.

[2] Payap Sirinam et al. "Deep fingerprinting: Undermining website fingerprinting defenses with deep learning". In: *arXiv preprint arXiv:1801.02265* (2018).