Raptor - Write-up*

Avishek Mondal

I. What has the paper contributed?

The paper shows 3 main ways user anonymity is decreased on the Tor network -

1) A new, asymmetric traffic analysis method - TCP headers and ACKs are monitored by a compromising AS (i.e. an AS that is both in the path of client-guard and exit-destination) and correlation is computed (i.e. Spearman's rank correlation coefficient is computed) and the server trace with the highest correlation is matched to the client traffic.
2) Natural churn - Analysis of how path changes due to the BGP protocol increases the chances of an AS becoming a compromising AS
3) Active attacks - Both hijacks, where packets are dropped, and interception, where path vector is modified so that traffic passes through an adversarial AS before going through the guard relay so that attack 1 can be carried out, are shown.

II. What are the assumptions, and their strengths and weaknesses?

A. Threat model

The bar is assumed to be quite high - AS level adversary, which has visibility of traffic both from the client to the guard, as well as from the exit to the destination. But this is a realistic assumption, given recent disclosures.
The experimental set-up assumes the client and the legitimate transit portal (TP), i.e. the one at Georgia Tech, has no BGP defences available. The attack lasts for only 5 mins (300 seconds). When compared to recent BGP hijacks, such as the April 26th 2017 attack by AS12389 [2] where the attack on major financial institutions of the world by a Russian AS lasted only 7 minutes, this time frame of 5 minutes is a realistic time frame. Thus, although in a real world setting detection is usually available that can allow the true owner of the hijacked prefix to become aware of the hijack and thus take counter measures (such as by broadcasting longer prefixes), during the short frames of these attacks these defences cannot be deployed, so the assumption of the client etc. having no defences is valid.

B. Experimental set-up

During the experiment, clients are only running Tor client software, and thus, all the network traffic generated from the 50 client machines can be assumed to be Tor traffic. However, in a real world setting, it can be expected that clients will be running non-Tor processes as well, that could generate non-Tor traffic. Thus, an additional step should be added to the attack, where network data is initially parsed to extract the Tor traffic.
One potential area for improvement could be scaling both the number of servers, as well as the complexity of the operation, which is currently a simple wget for a 100MB image to see how the co-relation attack scales. Intuition suggests that with more traffic to analyse (there will be more "noise", the accuracy could potentially reduce, but will still likely be significant.

III. How effective is the attack?

Given the current Tor topology and the BGP framework, it is a very efficient attack. Both user anonymity and receiver anonymity are compromised. It could however, be interesting to see the efficiency of this attack if more secure BGP protocols are adopted. These include cryptographic (for example source authentication etc.) and non-crypto methods (lifesaver ASes), but often involve changes to the entire BGP protocol and would be unlikely to be rolled out quickly.
It must be noted that in the original Tor paper [3] explicitly mentions that the Tor network is not able to defend against an adversary that has visibility over both client-guard traffic and exit-destination traffic. What this attack shows is that a weaker adversary than that originally envisioned by the Tor designers can *become* the stronger adversary that can deanonymise Tor users through active attacks, as well as the natural churn that exists in the BGP protocol.

IV. Comparison to the WF attack by DL models [4]

Both attacks show how a weaker adversary than that envisioned by the original designers of Tor can de-anonymise certain parts of the Tor network ([4] deanonymises recipient if the recipient is part of a close-world that the adversary has seen before). Perhaps it is wise to come up with a metric that characterizes the "effectiveness" of an attack in a more concrete way - comparing whether or not the attack requires the adversary to be active/passive, how quickly clients and destinations can be deanonymised etc. so that there is a basis for comparison between different attacks.

Effectiveness of proposed counter-measures are discussed in the write-up for Counter-RAPTOR

*Paper is [1]

REFERENCES

[1] Yixin Sun et al. "RAPTOR: Routing Attacks on Privacy in Tor". In: *24th USENIX Security Symposium (USENIX Security 15)*. Washington, D.C.: USENIX Association, 2015, pp. 271–286. ISBN: 978-1-931971-232. URL: https://www.usenix.org/conference/usenixsecurity15/technical-sessions/presentation/sun.

[2] Noticon. *BGP Hijacking overview. Routing incidents prevention and defense mechanisms*. URL: https://www.noction.com/blog/bgp-hijacking.

[3] Paul Syverson, R Dingledine, and N Mathewson. "Tor: The second generation onion router". In: *Usenix Security*. 2004.

[4] Vera Rimmer et al. "Automated website fingerprinting through deep learning". In: *Network & Distributed System Security Symposium (NDSS)*. 2018.