Oft Target

Tor adversary models that don't miss the mark

Aaron D. Jaggard
U.S. Naval Research Laboratory
aaron.jaggard@nrl.navy.mil

Paul Syverson
U.S. Naval Research Laboratory
paul.syverson@nrl.navy.mil

1. INTRODUCTION

The adversaries in published onion-routing and Tor research differ, to varying extents, from the adversaries against whom Tor was primarily designed to protect users.

Those published analyses have introduced and evaluated significant, often realistic, attack tactics, which have sometimes led to useful improvements to Tor's protocols and design. The adversary strategies reflected in existing work, however, generally do not serve the sensitive users that motivated Tor's design [7] and still constitute most of the examples of the types of users that Tor serves [8].

Tor is a network for traffic security of Internet communications [1] with millions of users [9]. Most Tor users are unlikely to be of specific interest to an adversary; they are primarily protected by Tor against opportunistic local eavesdroppers and local censors or against hostile destinations. Deanonymizing adversaries are generally modeled as attempting to attack as many connections as possible rather than targeting particular users or groups. But Tor is also explicitly intended to protect human rights workers, law enforcement, military, journalists, and others [8] who may face large, well-financed, and determined adversaries. More to the point, while some of these adversaries may try to hoover up whatever they can, some are interested in specific individuals or groups of Tor users, possibly based on offline or out-of-band reasons. An adversary whose interest is directed primarily or more fervently at particular users may employ different strategies. And, if Tor's design decisions are motivated by analyses ignoring such adversaries, those most in need of Tor's protections may be the least well served.

The adversaries we will describe in our talk are just such *targeting adversaries*. These need not differ at all from previously studied adversaries in terms of their capabilities or resource endowment, though they might. They differ primarily in their goals and strategies. As an example, consider a targeting adversary, Tom, who has compromised a particular user of interest, Alice, and observed her connecting to Bob, an interesting and unusual .onion website (essentially a website reachable only over Tor). Tom may wish to target other users of that site. He might also be particularly interested to learn which are the most active site users or how popular the site is in general.

Most research on security for widely-used systems follows the paradigm of assuming hoovering adversaries. Nonetheless, targeting has been shown to sometimes be much more effective than hoovering. For example, password guessing that is targeted based on knowledge about the intended victim has been shown to be more effective than hoovering in analyses of real data based on leaked datasets of passwords, and typically much more than twice as effective for security-savvy users [10]. And, NIST authentication guidelines, which had been created in consideration of hoovering strategies, were quickly modified in light of these analyses.

We will discuss targeting-strategy attacks for two example scenarios: a cabal meeting regularly on a private IRC channel, and regular visitors to a particular .onion website, such as mentioned above.

So far, we only have analytic rather than empirical results for these attacks, albeit based on empirical data concerning public Tor network composition and usage [2]. These show similar huge improvements in effectiveness versus hoovering, in fact much larger improvements than found in the password security analyses of targeting-strategy attacks. We will further argue that our adversary strategies are more realistic from a psychological and organizational resource and policy perspective than are the hoovering attack strategies typically considered.

2. TARGETING AN IRC CABAL

Suppose there is a cabal—a group of users wishing to conceal its activities, membership, and other properties—that meets regularly via a private IRC channel. Assume all of the cabal members access the IRC server exclusively via Tor.

A targeting adversary, Tom, might have seen mention of this cabal by a member, Alice, in another context, or might have targeted her for other reasons and become curious when he observed an over-Tor connection between her and an IRC server. Hoovering-strategy Tor-security analyses have focused on end-to-end correlating adversaries who try to compromise as many connections as possible. For an adversary comprising Tor relays, the best place for this is at the entry and exit relays; middles are largely useless. But a targeting adversary could be effective by starting with middle relays. Assume Tom is able to see whenever Alice is making an IRC connection to this service—e.g., if he has already compromised her guard or ISP and the IRC-server's ISP—and that he can correlate and match any, say, hour-long IRC meeting whenever its traffic goes through a relay he owns. Then, by observing from a moderate fraction of middle relays, Tom can both approximate the size of the cabal (which might be important in evaluating its importance) and identify the guards of other cabal members.

In our talk, we will compare the success of this targeting strategy to a published analysis by Johnson et al. of a hoovering adversary [3]. They set out as a behavioral user type an IRC user, who, 27 times a day, creates the same

single IRC sessions to the same server. In our specific usage scenario, the non-targeting adversary would reduce its median time to compromise by roughly 20% over their results. By contrast, the targeting adversary we consider will be one or two orders of magnitude more successful. For a 10–20 member cabal and comparable fraction of relay bandwidth (allocated to optimize each adversary model's performance), our targeting adversary will have a good idea of cabal size and will identify the guards of nearly all cabal members in under 4 days (100 meetings). The analogous adversary of Johnson et al. will require about 10 times as long to get a much rougher idea of cabal size and will have identified guards for roughly half of the cabal. To match the performance of the targeting adversary, the non-targeting adversary will take 40–50 times as long (150–200 days compared to under a week). We give details in a related preprint [2].

Note that even if Tom has already compromised a session from Alice and been able to observe all content and usernames for that session, the targeting strategy will remain quite useful. Though he will already know cabal size automatically, after about a week of IRC usage he will also have a good sense of cabal guards, and client send-receive activity per cabal guard (which may indicate cabal leaders and *will* indicate which members send the most).

Once guards are identified, a moderately-resourced targeting adversary can bring additional capabilities to bear in bridging guard to locate clients, but need do so *only for guards of those clients discovered to be worth locating.* He might, e.g., compromise a guard or its ISP, or coerce or extort operators or owners of either (legally, physically, etc.). Or he might use network-level attacks—e.g., over 90% of Tor relays are vulnerable to easy-to-mount BGP prefix hijacks that would reveal client IPs, given identified guards [6].

3. PICKING RIPE ONIONS

The set of users of a particular site may be similar to a cabal communicating via IRC. While they may not hold simultaneous meetings or even see themselves as a group, an adversary may target them because they are users of that site. A site may be interesting for exogenous reasons, e.g., because of a public mention of it or its presence in the browser bookmarks of a previously compromised target.

Our analysis [2] essentially applies to Tor users visiting many ordinary Internet sites, but we focus on onionsites, particularly hidden web services. These were designed to hide many features typically visible for ordinary websites. They have also had recent design changes specifically intended to make it harder for an adversary to discover a site's .onion address, popularity [5], or network location [4].

As observed, discovering site popularity may be an adversary goal or may be a criterion for deciding to target a site. We will note variants of the techniques for IRC cabals that can be used to measure onionsite popularity *at the client end, regardless of whether onionsite guards can be compromised or even identified.* These also show which users are the most active and show distribution of site activity in general. We will also note from our analysis capture-recapture techniques we introduce from population ecology to estimate the number of clients visiting a site n or more times.

4. SUMMARY

We will describe targeting adversaries in general and analyze some examples: targeting an IRC cabal and targeting frequent users of a particular onionsite.

We will also discuss psychological and organizational incentives and justifications for deployment of attacks: Short-duration and focused attacks that also give both intermediate feedback on success and further-action decision points are more psychologically and organizationally justifiable than are unfocused strategies with a much longer timeline for both similar-probability success and feedback about it. Even an adversary with generally hoovering goals would be foolish not to add targeting strategies when hoovering uncovers targets. While both have a place and may be in operation together, targeting adversaries have been overlooked despite being much more quickly and fully successful against vulnerable Tor users than are attacks by comparably resourced hoovering-strategy adversaries that have driven analysis and design to date.

Finally, we will discuss some countermeasures to targeting attacks: layered guards, randomized guard-set size and duration, leveraging trust diversity, and standardized templates for onion-service traffic.

5. REFERENCES

[1] R. Dingledine, N. Mathewson, and P. Syverson. Tor: The Second-Generation Onion Router. In *Proc. 13th USENIX Security Symposium*, 2004.

[2] A. D. Jaggard and P. Syverson. Onions in the crosshairs: When the man really *is* out to get you. `https://arxiv.org/abs/1706.10292`, June 2017.

[3] A. Johnson, C. Wacek, R. Jansen, M. Sherr, and P. Syverson. Users get routed: Traffic correlation on Tor by realistic adversaries. In *Proc. ACM Conf. on Computer & Communications Security*, CCS '13, pages 337–348, 2013.

[4] G. Kadianakis and M. Perry. Defending against guard discovery attacks using vanguards (Tor proposal #247). `https://gitweb.torproject.org/torspec.git/tree/proposals/247-hs-guard-discovery.txt`, 2015.

[5] N. Mathewson. Next-generation hidden services in Tor, (Tor proposal #224). `https://gitweb.torproject.org/torspec.git/tree/proposals/224-rend-spec-ng.txt`, 2013.

[6] Y. Sun, A. Edmundson, L. Vanbever, O. Li, J. Rexford, M. Chiang, and P. Mittal. RAPTOR: Routing attacks on privacy in Tor. In *Proc. 24th USENIX Security Symposium*. USENIX Association, 2015.

[7] P. Syverson. A peel of onion. In *Proc. 2011 Annual Computer Security Applications Conference (ACSAC'11), Orlando, FL, USA*, December 2011.

[8] Who uses Tor. `https://www.torproject.org/about/torusers.html`. Accessed February 2017.

[9] Tor metrics portal. `https://metrics.torproject.org/`.

[10] D. Wang, Z. Zhang, P. Wang, J. Yan, and X. Huang. Targeted online password guessing: An underestimated threat. In *Proc. 2016 ACM Conf. on Computer and Communications Security*, CCS '16, pages 1242–1254, 2016.