How Malicious Tor Relays are Exploiting Users in 2020 (Part I)

>23% of the Tor network's exit capacity has been attacking Tor users

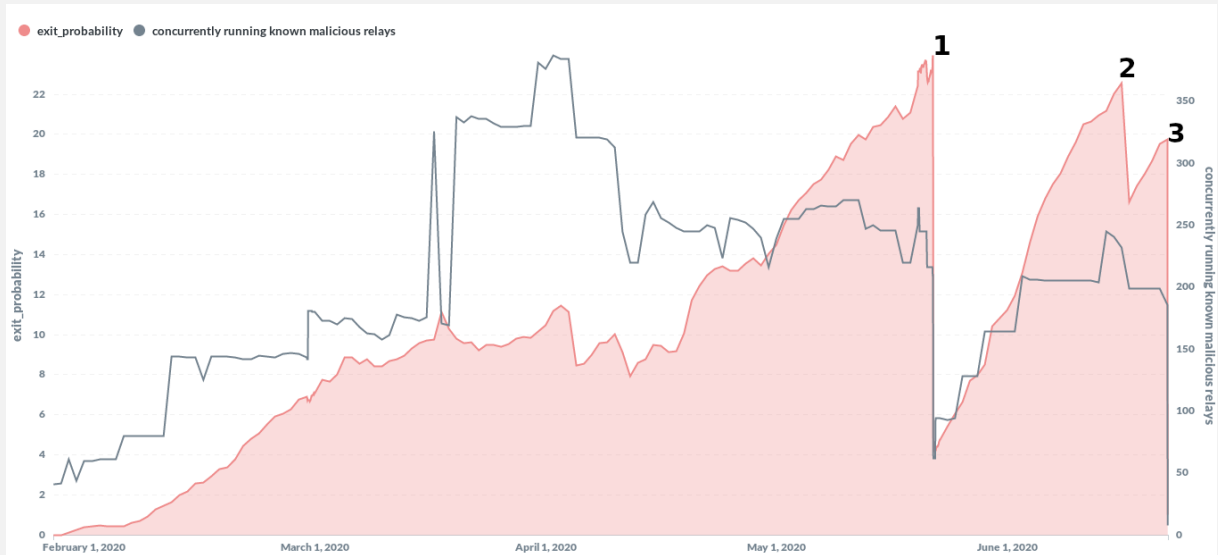

Figure 1: Confirmed malicious Tor exit capacity (measured in % of the entire available Tor exit capacity) over time (by this particular malicious entity). Graph by nusenu (raw data source: https://metrics.torproject.org/onionoo.html)

In December 2019 I wrote about The Growing Problem of Malicious Relays on the Tor Network with the motivation to rise awareness and to improve the situation over time. Unfortunately instead of improving, things have become even worse, specifically when it comes to malicious Tor exit relay activity.

Tor exit relays are the last hop in the chain of 3 relays and the only type of relay that gets to see the connection to the actual destination chosen by the Tor Browser user. The used protocol

(i.e. http vs. https) by the user decides whether a malicious exit relay can actually see and manipulate the transferred content or not.

In this post I want to give you an update on the malicious Tor relay situation for the first seven months of 2020 by looking at a single large scale malicious actor that is of ongoing concern. It demonstrates once more that current checks are insufficient to prevent such large scale attacks.

The Scale of the malicious Operator

So far 2020 is probably the worst year in terms of malicious Tor exit relay activity since I started monitoring it about 5 years ago. As far as I know this is the first time we uncovered a malicious actor running more than 23% of the entire Tor network's exit capacity. That means roughly about one out of 4 connections leaving the Tor network were going through exit relays controlled by a single attacker.

Figure 1 shows what accumulated fraction of the Tor network's exit capacity was controlled by the malicious actor and how many confirmed malicious relays were concurrently running (peak at over 380 relays). Figure 1 also tells us that we opened up Tor Browser at the peak of the attack on 2020–05–22 you had a 23.95% chance to end up choosing an attacker controlled Tor exit relay. Since Tor clients usually use many Tor exit relays over time the chance to use a malicious exit relay increases over time.

Temporary removal

The relay count line in Figure 1 shows that they added relays in big junks, which gives OrNetRadar (a relay group detector) the opportunity to detect them and it did in multiple cases (see Appendix). Most notably you can see a spike in relay count in March 2020. On 2020–03–16 OrNetRadar and the Tor Project's Sybil Attack detection reported a sudden spike of over 150 new relays. Something that basically never happens in such a short period of time. They got removed at the time, but were allowed to join the network 3 days later after the malicious operator contacted the bad-relays mailing list and configured the so called "MyFamily" setting to declare themselves as a group. Currently there are no further requirements to run such a large group of Tor relays.

Persistent

The 3 sharp drops in figure 1 (marked with 1, 2, 3) depict the events when some of these malicious Tor exits got detected, reported and removed from the network by the Tor directory authorities. This also shows us how fast the malicious entity recovered from a single removal event and that we didn't detect all of them at the same time. It took them less than 30 days to recover after a removal and reach 22% exit probability again (starting at 4%). It also gives us an idea that they apparently will not back-off after getting discovered once. In fact they appear to plan ahead for detection and removal and setup new relays preemptively to avoid a complete halt of their operations.

Faking multiple independent relay groups

The temporary removal event served them as a training and all relays that followed had presumably perfect MyFamily configuration, with one important caveat: Instead of declaring all of their relays in a single group they pretended to be multiple relay groups without linking them directly together. A strategy they followed from the beginning (January 2020). Figure 2 shows their exit probability by family contact information (stacked graph).

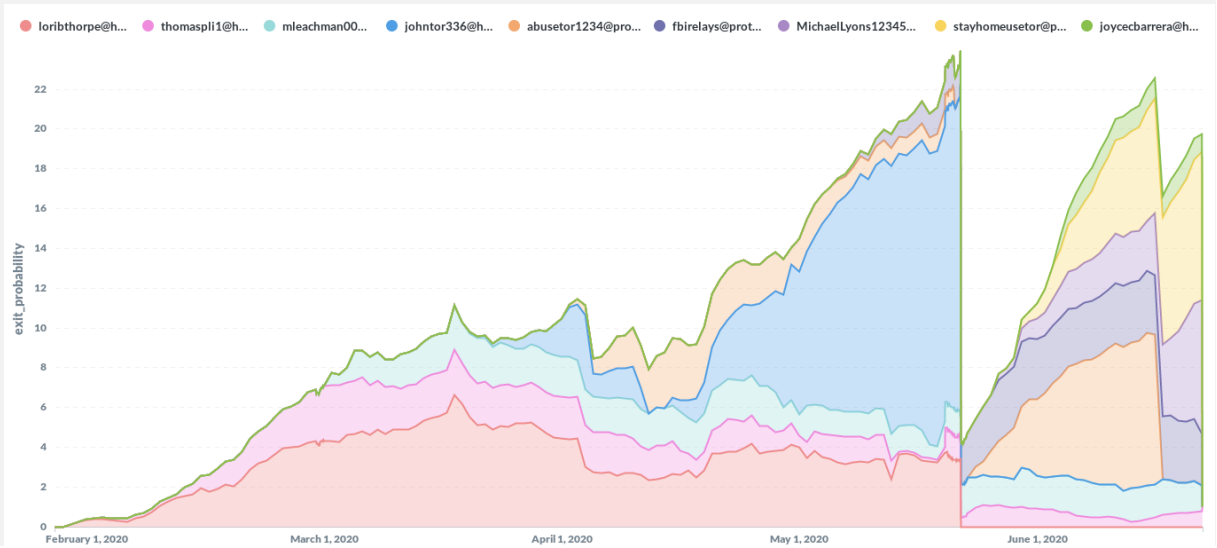

Figure 2: Confirmed malicious Tor exit fraction over time by ContactInfo (all of them are run by the same entity). Graph by nusenu (raw data source: https://metrics.torproject.org/onionoo.html)

Figure 3 shows how many malicious exit relays this particular actor operated split by given relay ContactInfo (stacked graph).

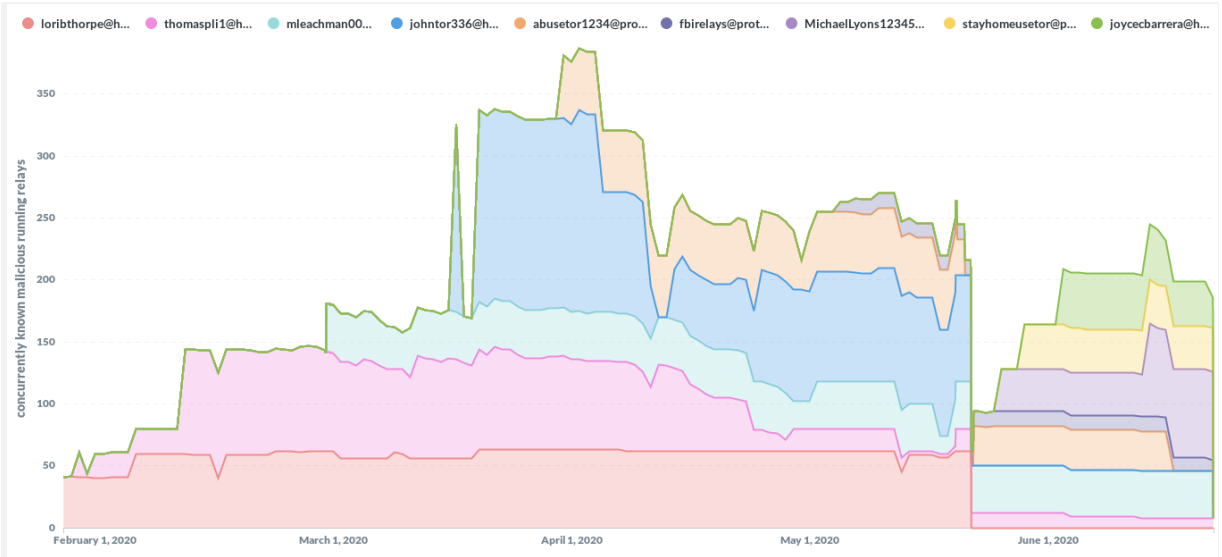

Figure 3: Confirmed malicious Tor exit relay count over time by ContactInfo. Graph by nusenu (raw data source: https://metrics.torproject.org/onionoo.html)

Contact information can be arbitrarily set by a relay operator so this information needs to be take with some caution, but since multiple of these email addresses interacted with the Tor Project's bad-relays mailing list, there is some confidence that all these addresses actually exists and are controlled by the malicious relay operator. They even impersonated the FBI by creating an address "fbirelays@..." (This email address was never used to contact the bad-relays mailing list. No, I don't believe the FBI has anything to do with these relays). We can see that the attacker likes to use common email providers (hotmail, protonmail and gmail).

Used Infrastructure

One key question of malicious relay analysis always is: What hosting companies did they use? So here is a break down by used internet service provider. It is mostly OVH (one of the — generally speaking — largest ISPs used for Tor relays). Frantech,

ServerAstra and Trabia Network are also known providers for relays. “Nice IT Services Group” looks interesting, since I’ve never seen relays on this obscure network before the attacker added some of his relays there on 2020–04–16.

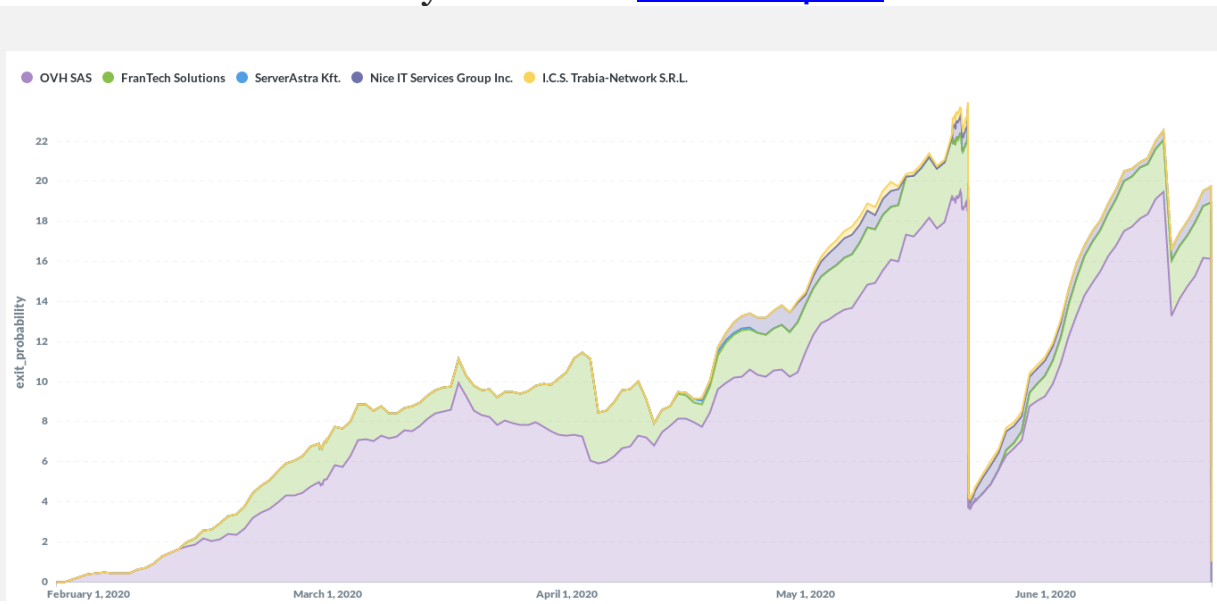

Figure 4: What ISPs did the attacker use? Mostly OVH and FranTech Solutions. Graph by nusenu (raw data source: https://metrics.torproject.org/onionoo.html)

What is this attacker actually exploiting and how does it affect Tor users?

The full extend of their operations is unknown, but one motivation appears to be plain and simple: profit.
They perform person-in-the-middle attacks on Tor users by manipulating traffic as it flows through their exit relays. They (selectively) remove HTTP-to-HTTPS redirects to gain full access to plain unencrypted HTTP traffic without causing TLS certificate warnings. It is hard to detect for Tor Browser users that do not specifically look for the “https://” in the URL bar.

This is a well known attack called “ssl stripping” that exploits the fact that user rarely type in the full domain starting with “https://”. There are established countermeasures, namely HSTS Preloading and HTTPS Everywhere, but in practice many website operators do not implement them and leave their users vulnerable to this kind of attack. This kind of attack is not specific to Tor Browser. Malicious relays are just used to gain access to user traffic. To make detection harder, the malicious entity did not attack all websites equally. It appears that they are primarily after cryptocurrency related websites — namely multiple bitcoin mixer services. They replaced bitcoin addresses in HTTP traffic to redirect transactions to their wallets instead of the user provided bitcoin address. Bitcoin address rewriting attacks are not new, but the scale of their operations is. It is not possible to determine if they engage in other types of attacks.

I’ve reached out to some of the known affected bitcoin sites, so they can mitigate this on a technical level using HSTS preloading. Someone else submitted HTTPS-Everywhere rules for the known affected domains (HTTPS Everywhere is installed by default in Tor Browser). Unfortunately none of these sites had HSTS preloading enabled at the time. At least one affected bitcoin website deployed HSTS preloading after learning about these events.

Is the attack over?

If we look at the overall advertised exit bandwidth on the Tor network and highlight the malicious capacity, that got removed by Tor directory authorities, we can see a significant increase in advertised exit capacity after the latest removal around 2020–06–21. This part of the curve actually looks similar to the previous month when the attacker recovered it's capacity after they got removed for the first time around 2020–05–22. I added an "expected" line to the graph to show where I would roughly expect the overall capacity to be without unusual growth (approximately calculated by adding the amount of advertised bandwidth known operators did add after the removal event).

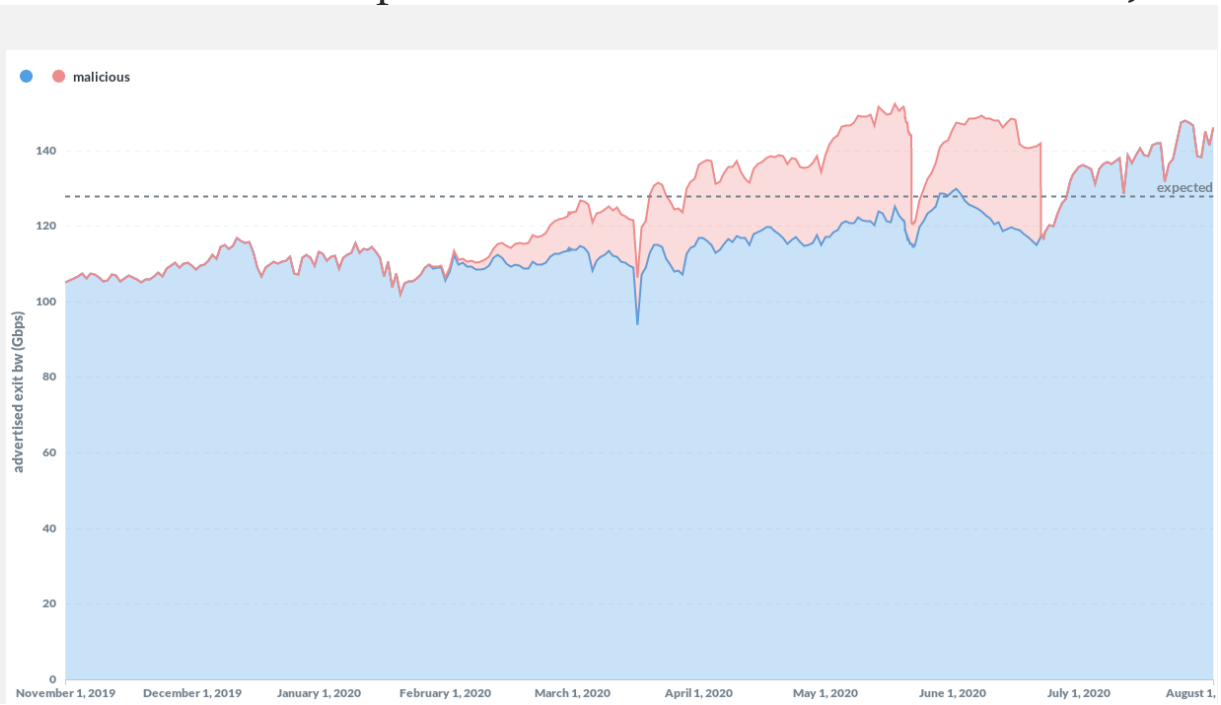

Figure 5: Overall advertised exit bandwidth in the Tor network over time shows unusual growth after removal of malicious relays. Graph by nusenu (raw data source: https://metrics.torproject.org/onionoo.html)

Figure 6 shows the impact the malicious operator had on the probability that a Tor Browser user would choose one of the known organizations running Tor exit capacity (like those mentioned on https://torservers.net/partners.html and others

active in the tor-relay community since an extended amount of time). The attacker was able to reduce their exit probability from usually around ~60% to bellow 50%. This graph also shows that the fraction of these known organizations is decreasing despite the fact that their absolute advertised exit capacity is actually increasing.

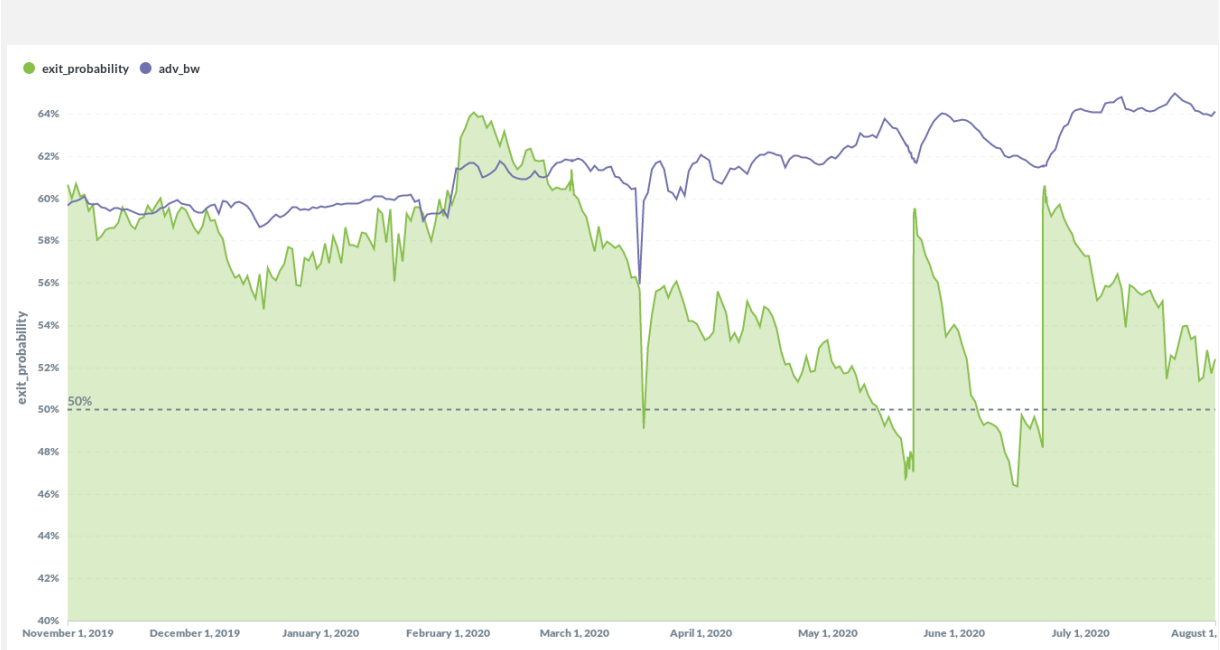

Figure 6: Exit fraction and advertised exit bandwidth by known operators/organizations. Graph by nusenu (raw data source: https://metrics.torproject.org/onionoo.html)

The fraction of known operators is decreasing, so who's share is increasing? Figure 7 and 8 show the exit fraction by unknown operators by autonomous system (figure 7) and by relay ContactInfo (figure 8). The graphs only show networks and ContactInfos with a significant fraction (>0.5% exit probability). The graphs show that the network fraction by the hosters OVH (heavily used by this attacker previously) and Liteserver Holding did significantly increase after the removal event around 2020–

06–21 and that two huge (>5% exit capacity each), new and unknown ContactInfos showed up.

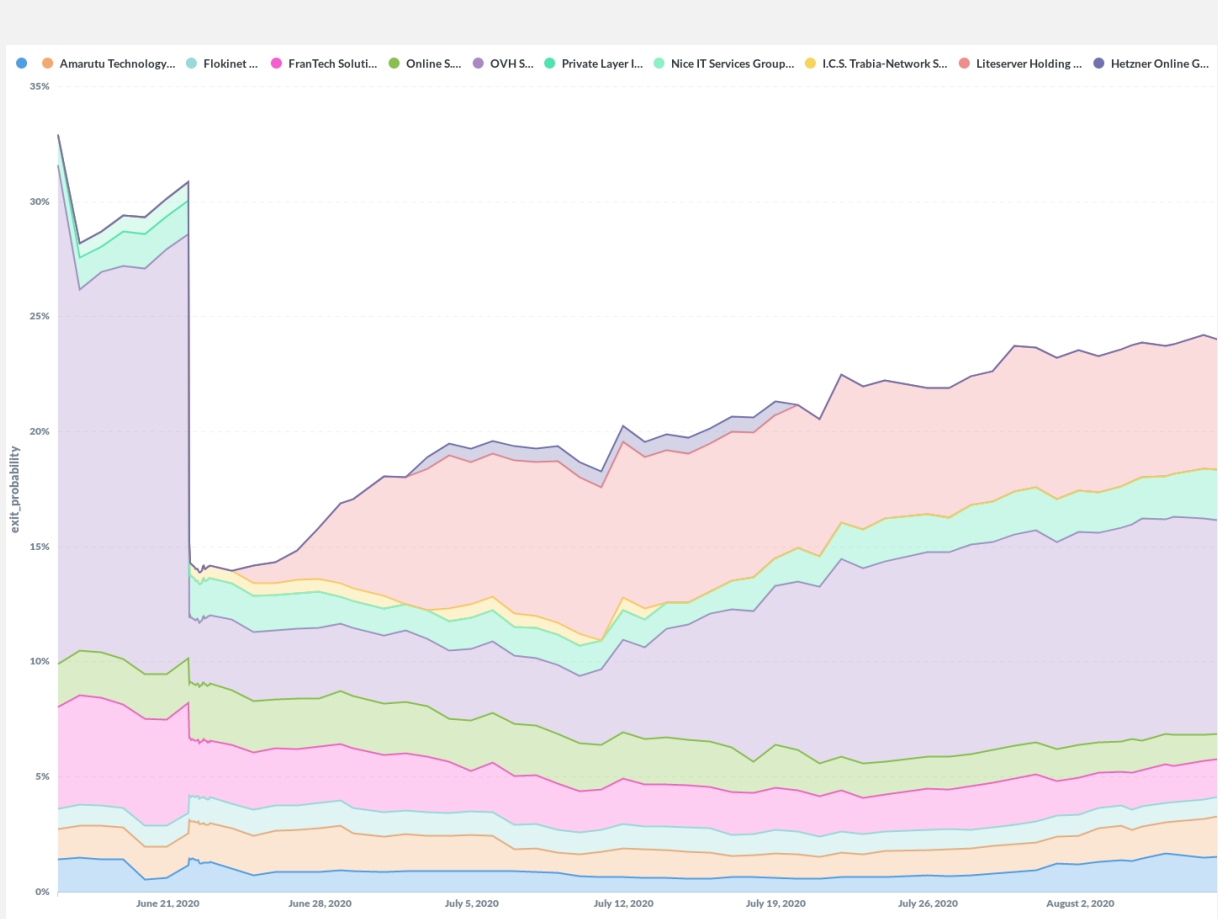

Figure 7: Exit fraction from unknown operators since the last removal of malicious exits (2020–06–21) by Autonomous System (showing ASNs >0.5% exit probability only). Two networks are significantly growing: OVH (again) and Liteserver Holding. Graph by nusenu (raw data source: https://metrics.torproject.org/onionoo.html)

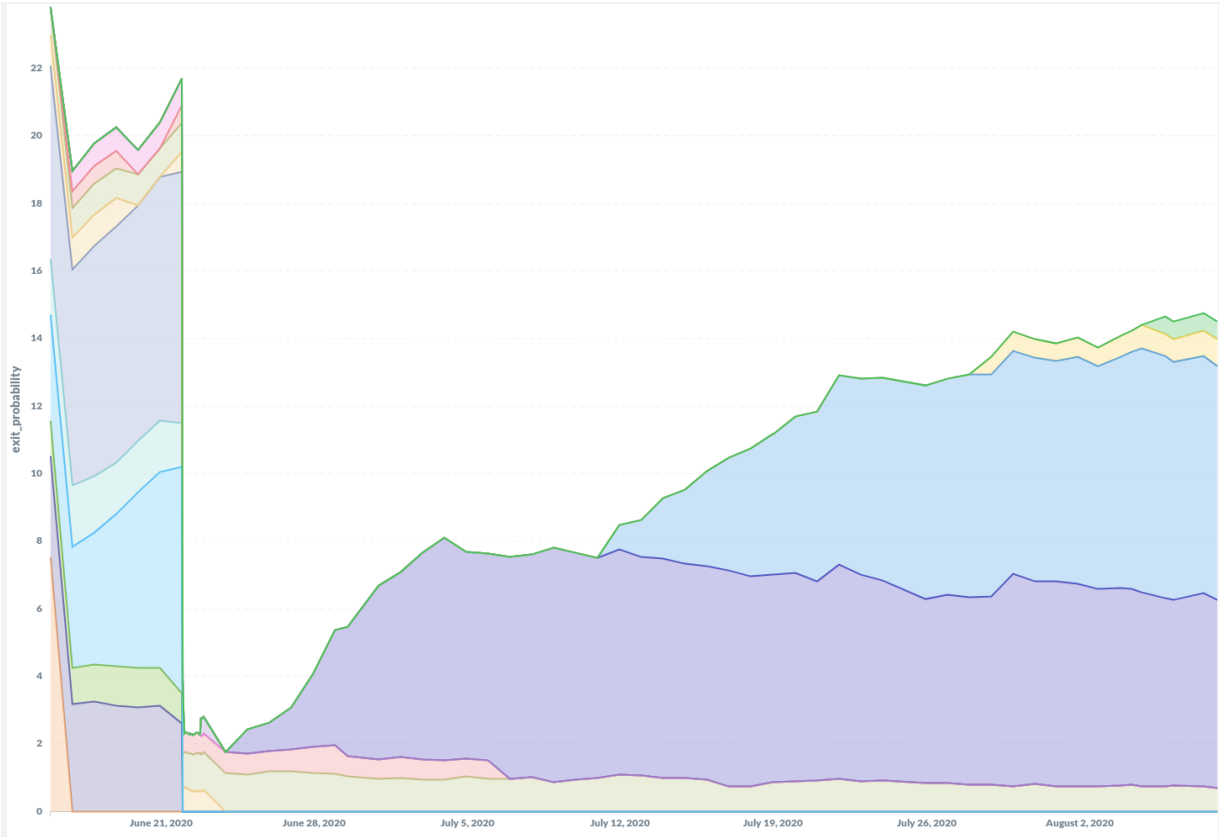

Figure 8: Exit fraction by unknown operators since the last removal of malicious exits (2020–06–21) grouped by exit relay contact information (stacked). Showing ContactInfos with >0.5% exit probability only. Exits with no ContactInfo are not included. Graph by nusenu (raw data source: https://metrics.torproject.org/onionoo.html)

These and some additional indicators, which I'm not publishing to avoid burning them, suggest that the attacker is not gone, but since exploitation of known victims became harder, the attacker might has chosen new victims or other types of attacks. This is an ongoing analysis and details might get published in a follow-up blog post.

Countermeasures

Bad-Relay Handling Situation

After the blog post from December 2019 the Tor Project had some promising plans for 2020 with a dedicated person to drive improvements in this area, but due to the recent COVID19 related layoffs that person got assigned to an other area. In addition to that, Tor directory authorities apparently are no longer removing relays they used to remove since 2020–06–26. It is unclear what triggered this policy change but someone apparently likes it and is adding undeclared relay groups (that used to get removed due to missing ContactInfo and MyFamily). This means the discovery of the attacker that did run over 10% of the Tor network's guard capacity in December 2019 apparently did not lead to any improvements. Since previous reports have not been acted or responded upon yet since over a month I've stopped reporting relays for removal, but let's leave that issue aside for the moment (topic for a separate blog post) and let's just keep in mind that the Tor Project has no dedicated resources to tackle this (relevant information when trying to make some progress).

Better visualizations for "known" vs. "unknown" network fractions

> "we lack the tools for tracking and visualizing which relays we trust" — Roger Dingledine

It is crucial to notice when the Tor network's known vs. unknown fractions significantly change. I'm aiming to tackle this by incorporating graphs like figure 6 and 7 into the daily

generated OrNetStats, but I'm only able to accomplish that when verification of some static operator identifier can be automated since manual verification is too time consuming on the long run. I'm working on version 2 of the ContactInfo Information Sharing Specification to provide Tor relay operators with two easy to implement options to allow for automated verification of the "operatorurl" field. Verified fields can then be used as an input for the manual assignment (done once) of a "known" label which is then used for the graphs. Once version 2 of the specification is released (that should happen before the end of August 2020) and deployed by enough relay operators, such graphs can be added to OrNetStats and other tools. This is also aiming to help with Roger Dingledine's plans on this matter. One crucial factor will be the adoption of the version 2 specification by relay operators.

Short term harm reduction

How do we make it more expensive and time consuming for malicious actors
to run such a big chunk of the the Tor network on an ongoing basis? Currently there are no requirements for (large scale) Tor relay operators. So currently there is nothing that stops a malicious actor from adding 150 malicious relays, as has been demonstrated by this attack in March 2020.

The following suggestions are taking into account that the Tor Project has no dedicated resources to address this issue.

As an immediate countermeasure against this ongoing issue the Tor Project could require physical address verification for all new (joined in 2020) Tor relay operators that run more than 0.5% of the Tor network's exit or guard capacity. Why 0.5%? It is a balance between the risk of malicious Tor relay capacity and the required effort for verification. Using 0.5% as a threshold is a realistically low number of operators to verify. As of 2020–08–08 there are just five exit and one guard operator that match these criteria (new and big). Some of them have similarities to previously detected malicious groups. Others are somewhat known with a good reputation already. So the amount for this initial verification is limited to sending 6 letters to a provided physical address (more likely actually 3 since some might not request the physical address verification).

This is also about empowering those that have to make difficult decisions when handling suspicious relays.

Long term: Limiting attackers by allocating a minimal network fraction to known operators

Roger Dingledine's plan is to allocate a fixed lower boundary to the "known" operators pool. This limits the damage a malicious operator can do, no matter how good they are at hiding their individual relays / relay groups. This approach is strong but should be combined with the following to further increase the efforts required by an attacker:

1. require a verified email address to gain the exit or guard relay flag.
Email verification can be fully automated. Since the malicious relay operators can easily register email addresses, this is combined with the second point.

2. require a verified physical address for large operators (>=0.5% exit or guard probability)

Summary

- Since the disclosure of large scale attacks on the Tor network (malicious operator did run >10% of Tor's guard capacity) in December 2019, no improvements with regards to malicious Tor relays have been implemented.
- The malicious Tor relay operator discussed in this blog post controlled over 23% of the entire Tor network exit capacity (as of 2020–05–22)
- The malicious operator demonstrated to recover their capacity after initial removal attempts by Tor directory authorities.
- There are multiple indicators that suggest that the attacker still runs >10% of the Tor network exit capacity (as of 2020–08–08)
- The reoccurring events of large scale malicious Tor relay operations make it clear that current checks and

approaches for bad-relays detection are insufficient to prevent such events from reoccurring and that the threat landscape for Tor users has changed.

- Multiple specific countermeasures have been proposed to tackle the ongoing issue of malicious relay capacity.
- It is up to the Tor Project and the Tor directory authorities to act to prevent further harm to Tor users.

Acknowledgements

I'd like to thank the person who initially reported some of the malicious relays, which allowed for the broader discovery of this huge malicious Tor exit relay fraction. (The reporting person asked to remain anonymous.)

Supporting this Research (section added on 2020-08-12)

In case you find investigating malicious Tor relays as important as I do: I'd like to further (and continuously) investigate this (and other) suspicious Tor network activities that might pose a risk to it's users. To allow me to further dig into this I'm looking for a Maltego Classic license since I'm running into some limits when using the free Maltego Community edition (an open source intelligence and graphical link analysis tool). In case you happen to work at Maltego or wanted to sponsor this type of research with such tooling support.

Appendix

OrNetRadar references to known malicious Tor exit relays by this actor:

https://nusenu.github.io/OrNetRadar/2020/01/29/a5
https://nusenu.github.io/OrNetRadar/2020/02/05/a3
https://nusenu.github.io/OrNetRadar/2020/02/11/a4
https://nusenu.github.io/OrNetRadar/2020/02/16/a2
https://nusenu.github.io/OrNetRadar/2020/02/29/a4
https://nusenu.github.io/OrNetRadar/2020/03/16/a5
https://nusenu.github.io/OrNetRadar/2020/03/30/a6
https://nusenu.github.io/OrNetRadar/2020/03/30/a7
https://nusenu.github.io/OrNetRadar/2020/04/11/a4
https://nusenu.github.io/OrNetRadar/2020/04/13/a8
https://nusenu.github.io/OrNetRadar/2020/04/14/a3
https://nusenu.github.io/OrNetRadar/2020/04/16/a7
https://nusenu.github.io/OrNetRadar/2020/04/21/a4
https://nusenu.github.io/OrNetRadar/2020/05/04/a9
https://nusenu.github.io/OrNetRadar/2020/05/06/a2
https://nusenu.github.io/OrNetRadar/2020/05/09/a2
https://nusenu.github.io/OrNetRadar/2020/05/22/a7
https://nusenu.github.io/OrNetRadar/2020/05/25/a4
https://nusenu.github.io/OrNetRadar/2020/05/28/a1
https://nusenu.github.io/OrNetRadar/2020/06/02/a1
https://nusenu.github.io/OrNetRadar/2020/06/13/a2