Countermeasures for RAPTOR Attacks

Yixin Sun
Princeton University

Anne Edmundson
Princeton University

1. Problem Motivation

Tor is a wildly used system for anonymous communications. However, Tor is known to be vulnerable to traffic correlation attacks in which an adversary can correlate the traffic at both ends of the communication to deanonymize the users. Recently, researchers have shown that AS-level adversaries can exploit the dynamics of BGP routing and launch new attacks on Tor [32], including both passive attacks exploiting routing asymmetry and natural churn, as well as active attacks with BGP prefix hijacking/interception. Thus, these new attacks urge the need to build countermeasures to defend Tor against such malicious AS-level adversaries. In this work, we will focus on developing countermeasures against active BGP routing attacks.

There are two potential ways to counter the attacks: (1) Mitigating traffic interception, and (2) Mitigating correlation attacks. However, mitigating correlation attacks usually involve employing extra encryption schemes and/or packet obfuscation, which either requires lots of engineering efforts, or could result in high latency of Tor. Thus, given the constraints of mitigating correlation attacks, we will build countermeasures that focus on mitigating traffic interception, more specifically, against active BGP routing attacks. Our approach includes two parts, as following.

2. Measuring Tor's Current State of Resiliency to Hijack Attacks

Because the Tor network is susceptible to network-level adversaries, we aim to quantify how much of the Tor network would be affected by a BGP prefix hijack. Proposed metrics will help enlighten the community about the state of the Tor network, in terms of how resilient the relays are to hijack attacks. Specifically, we aim to measure:

- Resiliency of the ASes that contain relays and compare the resiliency of ASes that contain more relays to those that contain few relays.
- Impact of a BGP prefix hijack on the Tor network.
- Probability of any given Tor relay being decieved by a BGP prefix hijack.

Previous work has tackled these questions using simulations of the entire Internet [22]. We will build off of this work by applying these metrics to the Tor network. There is also space for metrics regarding specific relays' resiliency to BGP prefix hijack attacks, as well as metrics regarding BGP prefix interception attacks (AS- and relay-level).

These metrics will help quantify how vulnerable the Tor network is to network-level adversaries in a novel way. We also plan to specifically look at the resiliency of guard relays (as a group) as well as exit relays (as a group).

Additionally, we will measure how the resilience of Tor relays has differed over the years. We want to answer the following questions:

- Have Tor relays become more resilient since the initial network was built?
- How fast does relay resilience change?

We plan to answer the first question by calculating the given resilience and impact metrics for each past year - similar to a longitudinal study of Tor relay resiliency. We plan to answer the second question by calculating the given resilience and impact metrics each week for the next couple of months. The results from answering the first question will also help us answer the second question.

In order to calculate these metrics, we will simulate prefix hijack attacks on an Internet derived topology. More information can be found in [22]. It's important to note that this work was done in 2007 and the number of AS links has greatly changed [3]. Additionally, we are specifically looking at Tor relays' resilience, not any given AS.

2.1 Recent Hijacks in the Wild

There have been a number of prefix hijack attacks in the past year. We plan to analyze the BGP routing announcements and withdrawals to find the prefixes that were hijacked and compare them to the list of Tor relay IP addresses at the time. This will give us information about whether or not prefix hijacks (or routing leaks) in the past year have affected Tor relays.

Some of the hijacks/leaks include:

- On November 6th, 2015, AS9498 (BHARTI Airtel Ltd.) hijacked about 16,000 prefixes [4].
- On June 12th, 2015, AS4788 Telekom Malaysia started to announce about 179,000 of prefixes to Level3 (AS3549, the Global crossing AS) [5].
- On March 27th, 2015, a BGP traffic optimizer leaked prefixes, which resulted in more than 7,000 new more-specific prefixes affecting roughly 280 Autonomous Systems, including large networks such as Rogers Cable, Telstra, Telenor, KDDI, BT-UK, Orange, Deutsche Telekom , Sprint, China Telecom, SHAW, LGI-UPC, AT&T, Comcast, Amazon, Internap, Time Warner Cable, Choopa, Syrian Telecommunications and many more [1].

3. Proactive Approaches

We take three different proactive approaches to counter RAPTOR attacks: 1) convincing relay operators to announce the relay in a /24 prefix, 2) analyzing the feasibiity of having a static route to guard relays, and 3) introducing a new path selection algorithm that minimizes the likelihood that a client sees a hijacked route in the case that her guard is hijacked.

3.1 Using /24 Prefixes

Sun *et al.* [32] recently found that >90% of BGP prefixes hosting relays are shorter than /24, making them vulnerable to a more-specific BGP prefix attack. Thus, one quick way to make Tor relays more resilient to such active routing attacks is to announce /24 prefix covering Tor relays. In order to make real world impact of this approach, we plan to start a campaign by contacting network operators whose prefixes contain Tor relays, and asking them to announce a more specific /24 prefix covering the relay.

3.2 Static Routing (or path protection mechanisms?)

There has been progress in protecting certain BGP paths by using static routes, or other protection mechanisms. We plan to explore how difficult it is to create static routes to guard relays in order to help previent prefix hijacks that aim to hijack a guard relay.

In the past, there has been work on protecting routes to top-level DNS servers. This was done by a fairly aggressive set of filters to be a little more more conservative in accepting alternate routes to the DNS servers [6].

3.3 Path Selection of Guard Relay

Guard relay is at an important position that it has direct connection with the Tor client. Thus, securing the guard relay would be our first step. It has been shown that AS-level adversaries can launch a more-specific prefix attack to intercept the Tor traffic from the guard relay to the malicious AS [32], and this can be potentially prevented by advertising /24 prefixes. However, even if the guard relay belongs to a /24 prefix, it is still subject to an equally-specific prefix attack. Unlike more-specific attacks which spread through the whole internet, equally-specific attacks can only affect connections within a small range - i.e., ASes that are within a certain number of hops away, depending on the influence of the announcing AS. Thus, picking a guard relay that is relatively close to the client AS could make it more resilient to such equally-specific attack.

On the other hand, we also don't want to pick a guard relay that is too close - in an extreme case, picking a guard within the same AS as the client will reveal client location to the adversary.

Therefore, we plan to develop new path selection algorithm that incorporates this aspect, picking a guard relay that satisfies the optimal balance between resilience to routing attacks and privacy protection.

4. Reactive Approaches

In addition to proactive approaches to helping prevent RAPTOR attacks, we have also taken a reactive approach. In the case that an attack is happening, we can detect it using a live monitoring framework.

4.1 A Live Monitoring Framework for Tor

The live monitoring framework aims at detecting suspicious routing attacks that affect Tor relays, and then react correspondingly to alert Tor clients of the scenario. The monitoring framework consists of two parts: BGP monitoring and Traceroute monitoring.

Relay Info from Tor Consensus Tor consensus releases up-to-date information for current running relays every hour. Our system automatically grabs this consensus data once it's updated. We focus on guard relays and exit relays, which reside at the two ends of the communication path and can easily be the target of an adversary. Further more, since we focus on AS-level adversaries, so it is unnecessary to monitor each individual relay by its IP address. But instead, we monitor the /24 prefixes which contain Tor guard and exit relays. Note that, there is no need to monitor a more specific prefix than /24, since generally /24 is the longest prefix accepted in BGP announcement.
Thus, we construct a live monitoring database which is being updated every hour with latest Tor relay data. The table contains the following fields:

/24 Prefix	Total Bandwidth	Number of Guards	Number of Exits	Timestamp

Each /24 prefix that contains any guard/exit relays will have one entry in the table, and the list of /24 prefixes will be used for our BGP and Traceroute monitoring frameworks, which we describe in the following.

BGP Monitoring Framework monitors the control plane of internet routing. We collect live BGP announcements data from BGP Stream, in combination with the latest Tor relay data. We monitor all the Tor-related /24 prefixes we obtain, as described

in the table above. We check if any activity related with the prefixes exhibit any anomaly. These anomalies can be detected by developing certain heuristics, such as the amount of time that a BGP path is used or the frequency that a path us announced; if certain anomalies fall under a threshold for a given heuristic, they should be flagged as potential attacks. This analysis will require saving offline relay bgp info for some period of time. This framework should help:

- Differentiate between hijack attacks and interception attacks [9].
- Differentiate between attacks and "normal" behavior, such as multiple origin AS conflicts, backup paths, etc [41]. We use Team Cymru to obtain AS ownership of prefixes. Some prefixes are owned by organization with multiple AS numbers, so we take this aspect into consideration and store the AS origins of these prefixes. If we observe any change in AS paths, we will first check if the prefix has multi-AS origins, and if so, as long as the new on-path AS also owns the prefix, then it would not be seen as an attack.

The implementation of the BGP Monitoring framework is based on BGP Stream [2]. We analyze the live stream of BGP updates and withdrawals, focusing just on the prefixes that contain a Tor relay. We monitor the prefixes that are reported through Team Cymru, as well as the /24 that contains each relay; we do this because we would like to be able to detect sub-prefix hijack attacks, so we must monitor longer prefixes in addition to the reported prefix.

As of now, our analysis involves checking the origin AS in the live BGP update and comparing it to the owner AS reported by Team Cymru. If these don't match up, then we flag the update (and prefix) as a potential hijack, otherwise we ignore it.

In the future, we may use heuristics (such as frequency, timing, and the number of collectors that saw the specific update) and store past updates/withdrawals, to increase the accuracy of our framework.

Traceroute Monitoring Framework monitors the data plane of internet routing, i.e., how packets travel through the Internet in reality. The traceroute monitoring framework is used as a verification mechanism if the BGP monitoring framework flags certain behavior. There may be many false positives in detecting hijack/interception attacks due to the nature of BGP. With many false positives, the traceroute monitoring framework will be used often for verification - this raises a question of optimization, which we will also address.

Our Traceroute monitoring framework retrieves updated Tor relay data hourly from the database described above. Since running large number of continuous traceroutes to relay IP addresses may create unnecessary extra traffic to the Tor network, so we selectively monitor a subset of prefixes at certain frequency rates when there is no anomaly from BGP monitoring data, and when there is suspicious activity report by BGP monitoring, the traceroute monitoring can be "triggered" to target the suspicious prefix announcement to verify the anomaly.

- Selectively monitor relays of interest.
  A /24 prefix can be evaluated based on several factors: (1) total combined bandwidth of Tor relays it covers, denoted as b_i for prefix i; (2) total number of guard relays it covers, denoted as g_i; (3) total number of exit relays it covers, denoted as e_i; and (4) resilience of the prefix to BGP hijack/interception attacks. These factors can make the prefix an attractive target to adversaries. Using these factors, we want to formulate the overall security of the system as following, and the goal is to find the monitoring frequency f_i for each relay i that maximizes the overall security of the network.
  (not including the resilience factor for now. can add it after we figure out the resilience stuff.)

$$\max \sum_{i=1}^{N} \frac{\log(f_i+1)}{b_i+g_i+e_i} \quad (1)$$

$$\text{s.t.} \sum_{i=1}^{N} f_i \leq F \quad (2)$$

$$0 < f_i \leq M, \forall i \quad (3)$$

  N is the total number of prefixes we want to monitor, F is the constraint on total number of traceroutes we can send from each Planetlab node per day, and M is the constraint on total number of traceroutes needed for each prefix. Given the solution, we use a collection of Planetlab nodes located in different ASes to send traceroutes to the prefix at its frequency rate.

- Monitoring target prefixes triggered by BGP.
  If we detect any anomaly from the BGP monitoring framework, we will immediately send traceroutes to the suspicious prefixes to verify whether there is truly a path change happening on the data plane to the Tor relays.

- Detecting anomaly from Traceroutes
  Even when there is no suspicious activity reported by BGP monitoring, it is also possible we detect anomaly from our selective traceroute monitoring. We keep track of the past traceroute monitoring results in a database table, as following:

Source Prefix	Dest Prefix	AS Path	Time Created	Time Last Updated	Current

  With this table, we will be able to compare the current AS path with past AS paths to detect any path changes, which may indicate a hijack event happening.

4.2 Framework Evaluation

The live monitoring framework will be evaluated on a number of characteristics, including false positive rate, false negative rate, as well as performance and overhead.

4.3 Deployment Experience

5. Related Work

BGP Attacks and Security. BGP attacks are well-studied, particularly prefix hijack and interception attacks [9, 25, 35]. Arnbak, et al. showed that prefix interceptions could be used by nation-states as a way to conduct surveillance on their citizens [8]. It's also known that routing anomalies can lead to network snapshots that look similar to attack scenarios. These are due to a range of routing policies, misconfigurations, and multiple origin AS conflicts [11, 24, 41].

The research community has contributed a number of protocols to help secure interdomain routing [10, 12, 15, 17, 36, 34]. Unfortunately, it has also been shown that partial deployment of secure interdomain routing protocols does not provide much security [23].

There is also a large body of research with the goal of defending against and detecting prefix hijacks and interceptions. These include defensive and detection tools [21, 16, 29, 38, 42, 30, 37], as well as mechanisms such as PGBGP, which allow network administrators more time to determine if an attack is happening before using new routes [20]. There has also been research not only on detecting attacks, but on determining the location of the attacker [28]. Qui, et al. detected any bogus routes, not just hijacks or interceptions [27]. In addition to detection algorithms, there has been research in visualization of real-time detection algorithms [33]. Our work does not aim to contribute a new hijack detection tool, but rather compliments existing tools by applying a monitoring framework to the Tor network.

BGP Attack Resiliency. Prior research on prefix hijack attack resilience has been simulated on the Internet for equal-length prefix hijacks [22]. They find that customers of Tier-1 ASes are the most resilient and also create the most impact (if they were to hijack a prefix). There has been some related work in relating hijack attacks to the Internet hierarchy [40, 39]. This differs from our work; we focus on the resilience of ASes that contain Tor relays, as well as measure the resilience of guard relays and exit relays as groups.

Network Adversaries on Tor. Network-level adversaries are known in anonymity networks. Feamster and Dingledine [14] first investigated AS-level path in anonymity networks, which showed that some AS could appear on nearly 30% of entry-exit pairs. Murdoch and Zielinski [26] later demonstrated the threat posed by network-level adversaries who can deanonymize users by performing traffic analysis. Furthermore, Edman and Syverson [13] demonstrated that even given the explosive growth of Tor during the past years, still about 18% of Tor circuits result in a single AS being able to observe both ends. In 2013, Johnson *et al.* [18] evaluated the security of Tor users over a period of time, and the result indicated that a network-level adversary with just low bandwidth cost can deanonymize any users within three months with over 50% probability and within six months with over 80% probability.

AS-level Tor Path Selection. The existence of network-level adversaires urges the need to incorporate AS-awareness path selection in Tor. In 2012, Akhoondi *et al.* [7] proposed LASTor, a Tor client which takes into account AS-level path and relay locations in path selection, although LASTor neglected relay capacity and its AS resilience to active attacks. Recently, Nithyanand *et al.* [31] constructed a new Tor client, Astoria, which adopted a new path selection algorithm which considered more aspects - relay capacity, asymmetric routing, colluding ASes, etc.. However, Astoria only considers a passive AS-level attacker, while does not evaluate the AS resilience to an active routing attack.

Towards this goal, it is important to understand AS-level internet topology and network path predictions. Lad *et al.* [22] investigated the relation between internet topology and prefix hijacking, and provided a metric for evaluating AS resilience to active prefix hijack attacks. Although, the study was conducted in 2007 when there were far less ASes than now. Recently, Juen *et al.* [19] performed a measurement study using Tracecroutes on network-level paths that Tor traffic actually get routed through.

6. References

[1] Bgp optimizer causes thousands of fake routes. `http://www.bgpmon.net/bgp-optimizer-causes-thousands-of-fake-routes/`.

[2] Bgp stream. `http://bgpstream.caida.org/`.

[3] Internet as-level topology archive. `http://irl.cs.ucla.edu/topology/`.

[4] Large scale bgp hijack out of india. `http://www.bgpmon.net/large-scale-bgp-hijack-out-of-india/`.

[5] Massive route leak causes internet slowdown. `http://www.bgpmon.net/massive-route-leak-cause-internet-slowdown/`.

[6] Protecting bgp routes to top-level dns servers. `http://web.cs.ucla.edu/~lixia/papers/03TPDS.pdf`.

[7] AKHOONDI, M., YU, C., AND MADHYASTHA, H. V. Lastor: A low-latency as-aware tor client. In *Security and Privacy (SP), 2012 IEEE Symposium on* (2012), IEEE, pp. 476–490.

[8] ARNBAK, A., AND GOLDBERG, S. Loopholes for circumventing the constitution: Warrantless bulk surveillance on americans by collecting network traffic abroad, 2014.

[9] BALLANI, H., FRANCIS, P., AND ZHANG, X. A study of prefix hijacking and interception in the internet. In *ACM SIGCOMM Computer Communication Review* (2007), vol. 37, ACM, pp. 265–276.

[10] BOLDYREVA, A., AND LYCHEV, R. Provable security of s-bgp and other path vector protocols: model, analysis and extensions. In *Proceedings of the 2012 ACM conference on Computer and communications security* (2012), ACM, pp. 541–552.

[11] CAESAR, M., AND REXFORD, J. Bgp routing policies in isp networks. *Network, IEEE 19*, 6 (2005), 5–11.

[12] CHAN, H., DASH, D., PERRIG, A., AND ZHANG, H. *Modeling adoptability of secure BGP protocol*, vol. 36. ACM, 2006.

[13] EDMAN, M., AND SYVERSON, P. As-awareness in tor path selection. In *Proceedings of the 16th ACM conference on Computer and communications security* (2009), ACM, pp. 380–389.

[14] FEAMSTER, N., AND DINGLEDINE, R. Location diversity in anonymity networks. In *Proceedings of the 2004 ACM workshop on Privacy in the electronic society* (2004), ACM, pp. 66–76.

[15] GILL, P., SCHAPIRA, M., AND GOLDBERG, S. Let the market drive deployment: A strategy for transitioning to bgp security. In *ACM SIGCOMM Computer Communication Review* (2011), vol. 41, ACM, pp. 14–25.

[16] HU, X., AND MAO, Z. M. Accurate real-time identification of ip prefix hijacking. In *Security and Privacy, 2007. SP'07. IEEE Symposium on* (2007), IEEE, pp. 3–17.

[17] HU, Y.-C., PERRIG, A., AND SIRBU, M. Spv: Secure path vector routing for securing bgp. In *ACM SIGCOMM Computer Communication Review* (2004), vol. 34, ACM, pp. 179–192.

[18] JOHNSON, A., WACEK, C., JANSEN, R., SHERR, M., AND SYVERSON, P. Users get routed: Traffic correlation on tor by realistic adversaries. In *Proceedings of the 2013 ACM SIGSAC conference on Computer & communications security* (2013), ACM, pp. 337–348.

[19] JUEN, J., DAS, A., JOHNSON, A., BORISOV, N., AND CAESAR, M. Defending tor from network adversaries: A case study of network path prediction. *arXiv preprint arXiv:1410.1823* (2014).

[20] KARLIN, J., FORREST, S., AND REXFORD, J. Pretty good bgp: Improving bgp by cautiously adopting routes. In *Network Protocols, 2006. ICNP'06. Proceedings of the 2006 14th IEEE International Conference on* (2006), IEEE, pp. 290–299.

[21] LAD, M., MASSEY, D., PEI, D., WU, Y., ZHANG, B., AND ZHANG, L. Phas: A prefix hijack alert system. In *Usenix Security* (2006).

[22] LAD, M., OLIVEIRA, R., ZHANG, B., AND ZHANG, L. Understanding resiliency of internet topology against prefix hijack attacks. In *Dependable Systems and Networks, 2007. DSN'07. 37th Annual IEEE/IFIP International Conference on* (2007), IEEE, pp. 368–377.

[23] LYCHEV, R., GOLDBERG, S., AND SCHAPIRA, M. Bgp security in partial deployment: is the juice worth the squeeze? *ACM SIGCOMM Computer Communication Review 43*, 4 (2013), 171–182.

[24] MAHAJAN, R., WETHERALL, D., AND ANDERSON, T. Understanding bgp misconfiguration. In *ACM SIGCOMM Computer Communication Review* (2002), vol. 32, ACM, pp. 3–16.

[25] MCARTHUR, C., AND GUIRGUIS, M. Stealthy ip prefix hijacking: don't bite off more than you can chew. In *Global Telecommunications Conference, 2009. GLOBECOM 2009. IEEE* (2009), IEEE, pp. 1–6.

[26] MURDOCH, S. J., AND ZIELIŃSKI, P. Sampled traffic analysis by internet-exchange-level adversaries. In *Privacy Enhancing Technologies* (2007), Springer, pp. 167–183.

[27] QIU, J., GAO, L., RANJAN, S., AND NUCCI, A. Detecting bogus bgp route information: Going beyond prefix hijacking. In *Security and Privacy in Communications Networks and the Workshops, 2007. SecureComm 2007. Third International Conference on* (2007), IEEE, pp. 381–390.

[28] QIU, T., JI, L., PEI, D., WANG, J., XU, J. J., AND BALLANI, H. Locating prefix hijackers using lock. In *USENIX Security Symposium* (2009), pp. 135–150.

[29] SHI, X., XIANG, Y., WANG, Z., YIN, X., AND WU, J. Detecting prefix hijackings in the internet with argus. In *Proceedings of the 2012 ACM conference on Internet measurement conference* (2012), ACM, pp. 15–28.

[30] SRIRAM, K., BORCHERT, O., KIM, O., GLEICHMANN, P., AND MONTGOMERY, D. A comparative analysis of bgp anomaly detection and robustness algorithms. In *Conference For Homeland Security, 2009. CATCH'09. Cybersecurity Applications & Technology* (2009), IEEE, pp. 25–38.

[31] STAROV, O., NITHYANAND, R., ZAIR, A., GILL, P., AND SCHAPIRA, M. Measuring and mitigating as-level adversaries against tor. *arXiv preprint arXiv:1505.05173* (2015).

[32] SUN, Y., EDMUNDSON, A., VANBEVER, L., LI, O., REXFORD, J., CHIANG, M., AND MITTAL, P. Raptor: routing attacks on privacy in tor. *arXiv preprint arXiv:1503.03940* (2015).

[33] TEOH, S. T., RANJAN, S., NUCCI, A., AND CHUAH, C.-N. Bgp eye: a new visualization tool for real-time detection and analysis of bgp anomalies. In *Proceedings of the 3rd international workshop on Visualization for computer security* (2006), ACM, pp. 81–90.

[34] VAN OORSCHOT, P. C., WAN, T., AND KRANAKIS, E. On interdomain routing security and pretty secure bgp (psbgp). *ACM Transactions on Information and System Security (TISSEC) 10*, 3 (2007), 11.

[35] ZHANG, Y., AND POURZANDI, M. Studying impacts of prefix interception attack by exploring bgp as-path prepending. In *Distributed Computing Systems (ICDCS), 2012 IEEE 32nd International Conference on* (2012), IEEE, pp. 667–677.

[36] ZHANG, Y., ZHANG, Z., MAO, Z. M., AND HU, Y. C. Hc-bgp: A light-weight and flexible scheme for securing prefix ownership. In *Dependable Systems & Networks, 2009. DSN'09. IEEE/IFIP International Conference on* (2009), IEEE, pp. 23–32.

[37] ZHANG, Z., ZHANG, Y., HU, Y. C., AND MAO, Z. M. Practical defenses against bgp prefix hijacking. In *Proceedings of the 2007 ACM CoNEXT conference* (2007), ACM, p. 3.

[38] ZHANG, Z., ZHANG, Y., HU, Y. C., MAO, Z. M., AND BUSH, R. Ispy: detecting ip prefix hijacking on my own. In *ACM SIGCOMM Computer Communication Review* (2008), vol. 38, ACM, pp. 327–338.

[39] ZHAO, J., AND WEN, Y. Analysis on the effect of prefix hijacking attack and internet hierarchy. In *Computer and Information Technology (CIT), 2012 IEEE 12th International Conference on* (2012), IEEE, pp. 375–382.

[40] ZHAO, J., WEN, Y., LI, X., PENG, W., AND ZHAO, F. The relation on prefix hijacking and the internet hierarchy. In *Innovative Mobile and Internet Services in Ubiquitous Computing (IMIS), 2012 Sixth International Conference on* (2012), IEEE, pp. 415–420.

[41] ZHAO, X., PEI, D., WANG, L., MASSEY, D., MANKIN, A., WU, S. F., AND ZHANG, L. An analysis of bgp multiple origin as (moas) conflicts. In *Proceedings of the 1st ACM SIGCOMM Workshop on Internet Measurement* (2001), ACM, pp. 31–35.

[42] ZHENG, C., JI, L., PEI, D., WANG, J., AND FRANCIS, P. A light-weight distributed scheme for detecting ip prefix hijacks in real-time. In *ACM SIGCOMM Computer Communication Review* (2007), vol. 37, ACM, pp. 277–288.