Using ML to Block BGP Hijacking

L. JEAN CAMP

L. Jean Camp is a Professor in the Luddy School of Informatics, Computing, and Engineering at Indiana University. She is currently a visiting scholar is at the University of California at Berkeley's Center for Long-Term Cybersecurity. She joined Indiana after eight years at Harvard's Kennedy School where her courses were also listed in Harvard Law, Harvard Business, and the Engineering Systems Division of MIT. After earning her doctorate from Carnegie Mellon, she spent one year as a senior member of the Technical Staff at Sandia National Laboratories. She began her career as an engineer at Catawba Nuclear Station with an MSEE from the University of North Carolina at Charlotte. She has authored more than 150 peer-reviewed publications on security and privacy, addressing trust on every layer of the OSI model. ljcamp@indiana.edu

Border Gateway Protocol (BGP) has proven to be resilient in the face of failures, attacks, and general maliciousness and incompetence. While there are no deployed mechanisms for automatically remediating BGP announcements that may be malicious, there have been many attempts at fixing this sorry state of affairs. In this article, I will describe some troublesome BGP events and how our tool, Bongo, uses machine learning (ML) and Layer 8 in the IP stack to detect malicious announcements and block traffic that would be diverted.

David Clark, in his book *Designing an Internet,* identified the Internet as a socio-technical system that could have been developed in many different ways. That the Internet is social, political, and economic is not contested in 2019. Yet the argument that BGP is social, political, and economic pushes this discussion to another level. In recent years several trends have emerged illustrating the vulnerability of the Internet's control plane. One of these trends is that, as the Internet expands, the expertise of individual operators is increasingly diverse, leading to more routing errors. The second is the increase in traffic redirection: in other words, route hijacking as an attack vector.

Consider first, misguided network configurations. China Telecom announced 15% of all IPv4 space in April of 2010, resulting in tremendous loss of traffic. This was represented as an error, and given that the traffic did not reach its intended destination, this would have been an extremely clumsy attack.

Another failure of routing was the misconfiguration of a small Australian ISP in 2012, one which took the continent "down under" down for hours. This is a most common failure of routing: a straightforward failure of human factors in configuration. These outages were immediately apparent and repaired within hours. However, the smaller customer ISP did in fact announce all routes from the two larger ISPs. The upstream ISP did not filter the route announcements appropriately and can be said to have both caused and suffered from the outage. Route leaks are a function of lack of technical competence, but this does not mean that they cannot diffuse quickly and cause harm. Previously proposed solutions have included customer route filtering, such as filtering all traffic from a client that will affect remote prefixes. The simultaneous emergence of untrustworthy behavior and the continuing need for connectivity illustrates the risk of such an approach. A different approach requires the increase of technical competence among not only ISPs but also larger end users, which appears optimistic [8].

In addition to errors and odd incidents, there has also been a wide range of political attacks. An early example originated from Pakistan in 2008, where Pakistan identified several YouTube videos as sufficiently problematic politically that the decision was made to block all of YouTube. An address internal to Pakistan for YouTube was announced within Pakistan Telecom; however, it was also broadcast across northern Africa and Europe, with a duration of hours. Arguably, while the intention to block YouTube within Pakistan was political, the leak itself was a human factors problem.

The efficacy of attacks of nation-states on the reachability of the Internet was further proven during Arab Spring, as rapid drop-offs for Egyptian and Libyan populations were easily observable but less easily repairable.

When an entity misrepresents its location in a routing path, rather than claiming to own a destination, the errors are less obvious. With this attack, traffic continues to be delivered and such a routing configuration could remain stable for long periods. Such a case occurred between China Telecom and AT&T, this time for some period of months: Facebook traffic was routed through China. Note that while the login to Facebook is protected by TLS, other uses of Facebook were not encrypted at that time. Thus a significant amount of global traffic was routed through a nation where Facebook adoption is remarkably low.

Since 2001, route hijacking has been turned into an attack [4]. Many of these attacks appear to remain undetected and unreported, creating a call for a ubiquitous cryptographic solution, RPKI [9]. Like many previous solutions, RPKI is incentive-misaligned and requires widespread adoption. BGPSEC is another example of a solution that was operationally, and economically, misaligned to the problem it was intended to solve. BGPSEC makes both requirements and benefits for early adopters that discourage innovators. Should BGPSEC be adopted it may not be resistant to political attacks nor misconfigurations. As the experience with certificate authorities issuing X509 certificates for the Web has demonstrated, political attacks are difficult to prevent and detect with an all-or-nothing cryptographic model of trust.

Sometimes malicious authorities are generally trusted for purposes of access, interoperability, and connectivity when populations on the network experience these authorities as malicious. Other solutions call for the creation of trusted third parties [10] or other changes in the infrastructure.

Addressing Geopolitical Dimensions

Our research directly addresses the geographical and political dimensions of BGP, reaching beyond purely technical dimensions to develop operational solutions. While specific threats to the control plane have included political interference, misguided network configurations, and miscellaneous mischief, much research on hardening the control plane has tended to be carefully neutral and apolitical—which is itself an ideological choice.

We contest this viewpoint by using a variety of data, including technical, rates of change, economic, and geopolitical, as network topology changes via BGP updates can offer probabilistic (not cryptographic) trust indicators. We understand that any ML approach needs to consider that attackers can and do adjust their strategies when defenses appear, and that any mechanism needs to do more than just provide temporary benefits.

At our lab, the application of machine learning to security has three basic principles. First, do not examine something that is easy for the attacker to change. Second, focus on the minimal requirements for a successful attack. Then leverage those minimal requirements to identify features that the attacker will have difficulty altering in a successful attack. And finally, use offline information or indicators of offline information when possible to better identify those features.

Using this as a starting point, we identified the fact that a requirement for BGP routing attacks is the ability to manipulate announcements from a single autonomous system (AS). Attacks, foolishness, and manipulations have arisen from single source ASes rather than coordinating across source networks. That the geography of the network reflects political geography was an underlying assumption for our work, one that has been more recently explored and validated [7].

What are the minimal requirements for a successful attack? The attacker's routing announcements must be distributed across the Internet without arousing suspicion so that the traffic is misdirected. Detecting this behavior means identifying patterns in attacks, so we began with offline features: jurisdiction and the data arising from a nation-state approach.

Incompetence or Attack

We did wonder whether we were simply seeing the rise of less qualified operators. An IETF member (who asked to remain anonymous) once argued, "As the number of people on the internet increases, the amount of cluefulness remains constant," meaning that more incidents will occur as participation in the network increases. Could cluelessness explain BGP anomalies? Is lack of technical expertise a significant explanatory variable for why some countries initiate more BGP incidences than others? Any given anomaly could be an accident, a crime, or an attack.

To understand the nature of routing anomalies, we empirically investigated the nations of origin of the events, using multiple regression and unsupervised learning techniques to analyze anomalies over a four-year period. If BGP anomalies are a result of limited technical competence, then countries with low levels of education, few technology exports, and less expertise should be overrepresented. If BGP anomalies are crime, leveraged by criminals for profit, then economic theories and analytical approaches from criminology should show statistical significance. Using macroeconomics by leveraging three theories from criminology and global measures of technology adoption, we examined whether anomalies were likely incompetence, potential e-crime, or intelligence operations.

For the issue of technical competence we included secure Internet services, IT exports, and third-party measures of network readiness. Our variable selection was also informed by different theories of criminology. We found that exports of technology were not statistically significant, undermining the argument for incompetence. We also found support for the possibility that anomalies were driven by crime, specifically for the guardianship and relative deprivation theories of crime.

We used unsupervised learning on the sources of BGP hijacks and apparently malicious routing announcements; and the result was two categories of nations, one correlated with higher levels of corruption and one associated with conflict. This clustering indicated that civil conflict and surveillance were associated with the disproportionate origination of routing anomalies [5].

Choosing Safety over Availability

The lack of support for the hypothesis that BGP anomalies are generated by incompetents, the support for the role of crime, and the indication that political stability is an indicator provided a design guideline for our next step. If there is a threat from a particular jurisdiction, then announcements that change stable routing patterns to include these jurisdictions are suspicious.

Beyond the refusal to treat the Internet as if it were some third space, unmoored to political nor geography reality, our next decision made explicit the tradeoff between availability and confidentiality. Is there data where it is better not to send it, rather than sending it and risking exposure? We developed a proof of concept that generates firewall rules based on BGP changes according to the jurisdiction of the nation of the announcing AS.

We examined the feasibility of changing the calculus of the confidentiality, integrity, and availability model by providing the choice to select confidentiality and integrity over availability [3]. We named this package Bongo to fit with the ungulate method of naming in the open source route reflector tradition. We distinguish this from Quagga and Zebra in that bongos are the only spiral-horned antelopes (tragelaphid) where both male and females have horns. Our goal was to offer a pointed defense. Bongo uses the Routing Information Base (RIB) to identify changes not in routes but in AS, and then assigns each AS a risk parameter. Based on this parameter, Bongo can create a firewall rule. Alternatively, Bongo may simply issue an alert to the operator, if the change is worthy of note but not high risk.

The approaches for route filtering also would work with the next generation Internet architecture, software-defined networking (SDN). SDN is an established alternative architecture, where flows (rather than routes) define the paths of data using information from multiple layers (rather than IP and NAT information only). To say that SDN was not built with security as a design goal is a bit of an understatement. Yet the nature of flows inherently offers potential against BGP attacks. Changes in flows and delaying flows should be less prone to overlap and confusion than potentially quickly changing firewall rules. If this holds under further study, then SDN allows the promise of rejecting changes in the BGP topology.

Bongo also implements the creation of OpenFlow rules based on risk estimates, including jurisdiction. The key difference between Bongo and other works is the assumption that hijacks and leaks will happen upstream beyond the control of the end of the network. The goal is to enable what a BGP end user—that is, companies operating networks receiving BGP updates from their upstream ISPs—can do to react to these events.

Additionally, even if prefix hijacking and path-length attacks were to be completely eliminated, an organization may want to cease sending data in reaction to a topology change that would send sensitive data to an undesirable geographical region. Bongo enables prohibition of traffic from traversing certain jurisdictions. At some level, this seems as simple as the promulgation of simple yes or no rules. However, full path verification is clearly not feasible (upstream routers may lie about AS paths), so this detection inherently includes uncertainty.

A Pragmatic Approach

Bongo is designed to defend the network as it is currently operated, without assumptions about adoption of PKI or any other technology in the future. Bongo approaches route updates as risk decisions and can estimate the risk of adopting a route based on any filter the programmer deems appropriate. Bongo allows an organization to decide exactly how much and what kind of risk it will accept from the control plane.

In addition to publishing this in a traditional technical domain—an ACM workshop—we also sought feedback from experts in communications and Internet policy. An earlier version, before the ACM publication, was presented at the Telecommunications Policy Research Conference. This venue has a rich history of publishing visions before there is a reality. Software-defined radio was included in the call for papers before it was a respectable technical idea; and Internet commerce was a session at TPRC before the First Workshop on Electronic Commerce was hosted by one of the big three (ACM, IEEE, and, of course, USENIX) [1].

The feedback at that venue was that the reality of the politics of delay, misdirection, and black holing of traffic was already known in policy circles. The roles of criminals and intelligence agencies was much discussed at this Telecommunications Policy Research Conference, but rarely with engineers in the room.

To determine how disruptive such an approach might be, we examined the paths to the top global financial institutions. We found only the Bank of Tehran would have suffered any denial of service connecting from California while using Bongo, and that may be a result of route flapping. Other routes were extremely consistent in terms of jurisdictions traversed. We identify these as political and geographical decisions [2].

Specifically, we selected the top 50 banking websites from Alexa as a rough approximation for sensitive addresses to which an organization may want to apply transit restrictions. We examined the BGP state from a single ISP based in San Francisco over the course of April 2016 and identified every time the country associations changed in the AS path to each banking site. Out of the 50 financial services IP addresses analyzed over the one-month period, only four of them experienced country changes in the path to their destination and one experienced an outage. That means that for 46 of the IP addresses, an extremely strict exfiltration policy, restricting to no AS path country changes, would not have caused any outages during this time period from our vantage point in San Francisco.

Only four showed a change in jurisdiction along their paths. Online.citibank.com, during a period of 24 hours on April 5 to April 6, 2016, had routes to their network completely withdrawn from the routing table. The path for www.nbg.gr, the website for the National Bank of Greece, changed on April 9 to a direct peering between a US provider and a Greek (GR) provider, eliminating an intermediary peer in the EU. HSBC.com alternated between being advertised directly in the US and being advertised by an ISP in Great Britain (GB). For US-only paths, this would have caused an eight-day outage and then a three-day outage, but for US and GB paths, there would have been no impact.

The one likely impacted domain would have been Bsi.ir, Bank Saderat Iran, headquartered in Tehran. From the start of the period until April 17, the path traversed the US, Russia (RU), Azerbaijan (AZ), and Iran (IR). For the following eight days, it traversed Oman (OM), the US, and Iran. Then for four hours on April 25, the path traversed Germany (DE), the US, and Iran, after which it switched back to the Oman route. Thus, the use of Bongo for preventing connections through specified jurisdictions would have affected only Bank Saderat in its connections to the US were Germany, Russia, or Azerbaijan disallowed.

Another Approach

Given that it is feasible to defend against possible targeted events, can we detect larger-scale assaults on BGP? Again we began by asking the three core questions for applying machine learning to security: what is required, what is the goal, and what features can be identified with this understanding? Building on the requirements to defend against attacks that require large-scale redirections, we defined the goal of the attack as fast, undetected changes in topology.

Our analysis showed that large-scale attacks could be identified at an earlier stage than reported in the literature, again by focusing on the single-AS source and the necessary results for a successful attack. For a successful attack, attackers must issue a large number of route changes, leading to detectable bursts in their announcements. So we examined each AS as a data source and compared the interarrival times of announcements not only from the (malicious) AS but also to the neighboring ASes. BGP announcements that are associated with disruptive updates occurred in groups of relatively high frequency, often multiple standard deviations from normal rates, followed by periods of infrequent activity.

Conclusion

Together this set of analysis and open code illustrates that it is possible to quickly identify some BGP attacks and then mitigate the risks of hijacks when confidentiality is more valuable than availability [6].

Since a core problem with BGP resiliency is the concept of trust, then trust and risk must be a core of the solution. Understanding routing updates as a function of trust and risk enables approaching such updates as partially trusted. Cryptographic solutions attempt to provide perfectly trustworthy sources and paths. Yet certificate authority subversion in the TLS realm have shown that today's certificates are not themselves trustworthy; nor does this proposed solution address misconfiguration or malicious configurations.

Trust is not globally encompassing. The Internet is no less affected by the transit across geographical boundaries than were telephone calls, letters, or other communications infrastructures. Recognizing the connection between the political, the physical network, and the requirements for security in the sociopolitical reality of the Internet offers the potential for more robust defense and earlier indicators. Not taking advantage of this would be an ideological, not technical, choice.

References

[1] K. Benton and L. J. Camp, "Examining the Jurisdictions of Internet Routes to Prevent Data Exfiltration," 44th Research Conference on Communications, Information and Internet Policy (TPRC44), 2016.

[2] K. Benton and L. J. Camp, "Firewalling Scenic Routes: Preventing Data Exfiltration via Political and Geographic Routing Policies," in *Proceedings of the 2016 ACM Workshop on Automated Decision Making for Active Cyber Defense* (*SafeConfig '16),* ACM, 2016, pp. 31–36.

[3] K. Benton, L. J. Camp, and M. Swany, "Bongo: A BGP Speaker Built for Defending against Bad Routes," in *MILCOM 2016 IEEE Military Communications Conference*, IEEE, 2016, pp. 735–739.

[4] T. Mizuguchi and T. Yoshida, "Inter-Domain Routing Security ˜BGP Route Hijacking˜," Asia Pacific Regional Internet Conference on Operational Technologies (APRICOT '07), 2007.

[5] P. Moriano, S. Achar, and L. J. Camp, "Incompetents, Criminals, or Spies: Macroeconomic Analysis of Routing Anomalies," *Computers & Security*, vol. 70, 2017, pp. 319–334.

[6] P. Moriano, R. Hill, and L. J. Camp, "Using Bursty Announcements for Early Detection of BGP Routing Anomalies": https://arxiv.org/abs/1905.05835, 2019.

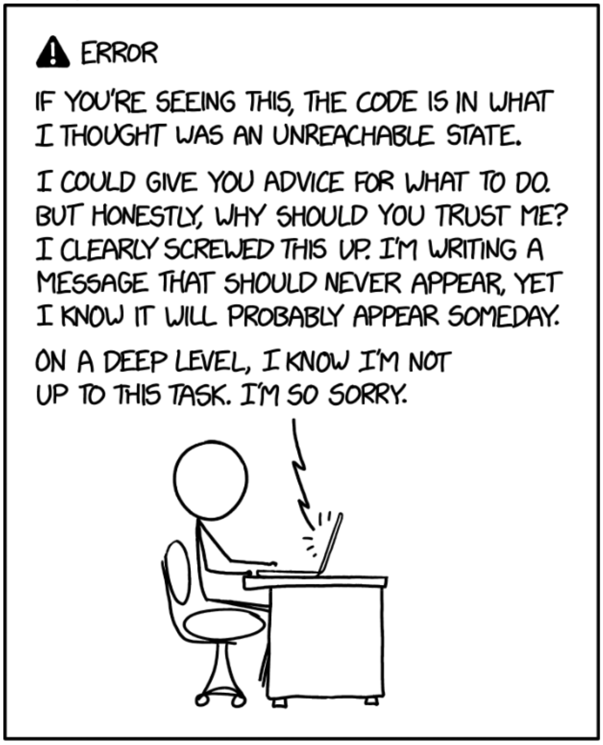
[7] L. Petiniaud and L. Salamatian, "Geopolitics of Routing," RIPE Network Coordination Center: https://labs.ripe.net/Members/louis_petiniaud/geopolitics-of-routing, 2019.

[8] V. Valancius, N. Feamster, J. Rexford, and A. Nakao, "Wide-Area Route Control for Distributed Services," in *Proceedings of the USENIX Annual Technical Conference (ATC '10)*, pp. 17–30: https://www.usenix.org/legacy/events/atc10/tech/full_papers/Valancius.pdf.

[9] M. Wählisch, O. Maennel, and T. C. Schmidt, "Towards Detecting BGP Route Hijacking Using the RPKI," *ACM SIGCOMM Computer Communication Review*, vol. 42, no. 4, 2012, pp. 103–104.

[10] Z. Zhang, Y. Zhang, Y. C. Hu, and Z. M. Mao, "Practical Defenses against BGP Prefix Hijacking," in Proceedings of the 2007 ACM Conference on Emerging Network Experiment and Technology (CoNEXT 2007), article 3.

XKCD **xkcd.com**


ERROR
IF YOU'RE SEEING THIS, THE CODE IS IN WHAT I THOUGHT WAS AN UNREACHABLE STATE.
I COULD GIVE YOU ADVICE FOR WHAT TO DO. BUT HONESTLY, WHY SHOULD YOU TRUST ME? I CLEARLY SCREWED THIS UP. I'M WRITING A MESSAGE THAT SHOULD NEVER APPEAR, YET I KNOW IT WILL PROBABLY APPEAR SOMEDAY.
ON A DEEP LEVEL, I KNOW I'M NOT UP TO THIS TASK. I'M SO SORRY.

NEVER WRITE ERROR MESSAGES TIRED.