A Novel Unsupervised Method for Securing BGP Against Routing Hijacks

Georgios Theodoridis, Orestis Tsigkas and Dimitrios Tzovaras

Abstract In this paper, a BGP hijack detection mechanism is presented. The proposed methodology is utterly unsupervised and no assumptions are made whatsoever, but it is developed upon the extraction of two novel features related to the frequency of appearance and the geographic deviation of each intermediate AS towards a given destination country. The technique is tested under a real-world case of BGP hijack and the efficiency of the features and the corresponding proximity measures is assessed. It is proven that the proposed approach is capable of decisively capturing such events of malicious routing path anomalies.

Keywords BGP · Hijack · Security · Detection · Routing · Anomalies

1 Introduction

According to its fully decentralized structure, Internet is established as the sum of numerous administrative regions called Autonomous Systems (ASes), which are interconnected through the existing backbone on the basis of the interdomain routing protocol, i.e. Border Gateway Protocol (BGP) [1]. BGP is responsible for maintaining and communicating the routing directives among the Internet comprising entities and hence defining the actual operational network topology.

This work has been partially supported by the European Commission through project FP7-ICT-257495-VIS-SENSE funded by the seventh framework program. The opinions expressed in this paper are those of the authors and do not necessarily reflect the views of the European Commission.

G. Theodoridis (✉) · O. Tsigkas · D. Tzovaras
Centre for Research and Technology Hellas, Information Technologies Institute,
6th km Charilaou-Thermi Road, P.O. Box 60361, 57001 Thessaloniki, Greece
e-mail: gtheo@iti.gr

E. Gelenbe and R. Lent (eds.), *Computer and Information Sciences III*,
DOI: 10.1007/978-1-4471-4594-3_3,

Routing paths undergo continuous alterations as a result of hardware failures and the varying inter-AS relationships. In consequence, the volume of BGP activity is significantly increased, while phenomena of intense BGP disturbances also appear at high frequency. Additionally, due to the fact that it is developed upon the concept of mutual trust, BGP suffers from inherent security vulnerabilities, which can severely compromise its functionality and the integrity of the routing paths. In this context, cyber attacks against BGP have nowadays emerged as one of the most prominent Internet threats. Hence, given the key role of BGP in conjunction with the utmost importance of Internet, it becomes an urging necessity to develop the adequate mechanism that would allow for the effective detection and root cause analysis of potentially malicious BGP anomalies.

In this context, significant research activity has been drawn on the detection of BGP anomalies [2]. Ballani et al. perform a thorough study of the prefix hijacking mechanisms with special emphasis on the interception cases [3], while Gao aims at inferring the inter-AS relationships by solely utilizing the raw BGP data [4]. One of the most common approaches is based on the extraction and unsupervised statistical analysis of the most appropriate features [5, 6]. Alternatively, in a semi-supervised learning manner, data mining upon well known BGP behavior is performed, so as to calibrate the detection algorithms [7, 8]. Several monitoring systems have also been developed that build a view of the AS/prefix topology and report on any alterations against this state of Refs. [9, 10]. Furthermore, the *Argus* system focuses specifically on cases of traffic blackholing, by combining info from both the control and the data-plane [11].

Towards this ultimate goal of decisively capturing and attributing any abnormal routing alterations, a completely novel methodology is hereby presented, which introduces two significant novelties. First, it is completely unsupervised, requiring no a-priori knowledge of the BGP dynamics. Second, it manages to efficiently make virtue of the underlying geo-spatial coherence of the Internet routing information by grouping the BGP activity on a per destination-country basis; to this aim two new feature related to each AS's frequency of appearance and geographic divergence are extracted.

The paper is structured as follows. Section 2 describes the proposed methodology in detail, while in Sect. 3 the technique is implemented and evaluated in a real-world scenario. Finally, Sect. 4 summarizes the paper.

2 Proposed Methodology for AS-Path Hijack Detection

A primary threat against BGP is related to the case that a malicious AS announces itself as an intermediate hop towards an already occupied prefix. As a result, all the IP traffic of the victim prefix is compelled to traverse the attacking AS, in order to be either blackholed or intercepted.

More precisely, let $\mathbf{W^{M,p}}$ be the initial AS-Path connecting the monitoring AS (M) with the destination AS (O^p) that owns prefix p and let $\mathbf{W'^{M,p}}$ be an alternative AS-Path announced at a later time instant.

$$\mathbf{W^{M,p}} = \{M, A_1^{M,p}, \ldots, A_K^{M,p}, O^p\} \tag{1}$$

$$\mathbf{W'^{M,p}} = \{M, A_1'^{M,p}, \ldots, A_{K'}'^{M,p}, O^p\} \tag{2}$$

Then, $\mathbf{F^M}(\mathbf{W}, \mathbf{W}')$ denotes the set of non-common ASes for $\mathbf{W^{M,p}}$ and $\mathbf{W'^{M,p}}$:

$$\mathbf{F^M}(\mathbf{W}, \mathbf{W}') = \mathbf{W^{M,p}} \cup \mathbf{W'^{M,p}} - \mathbf{W^{M,p}} \cap \mathbf{W'^{M,p}} \tag{3}$$

An AS-Path anomaly is identified if and only if $\mathbf{F^M}(\mathbf{W}, \mathbf{W}') \neq \emptyset$, i.e. there is at least one non-common intermediate AS between the competing routes. Each non-common AS, $A_h \in \mathbf{F^M}(\mathbf{W}, \mathbf{W}')$, is suspicious of performing BGP hijacking. If $A_h \in \mathbf{AP^{M,p'}}$, the later announcement is the root cause of the hijacking. On the contrary, if $A_h \in \mathbf{AP^{M,p}}$, then the BGP anomaly event is related to an attempt of the prefix's owner (O^p) to restore the normal path.

Hence, in order to be able to evaluate the normality of an AS-Path alteration, it is necessary to assess the legitimacy of each AS's appearance within the announced AS-Paths. To this end, two primary features are extracted, while, different similarity measures are utilized for each one of them, in order to capture the root causes of the BGP activity. In more detail, let $\mathbf{C} = \{C_1, \ldots, C_D\}$ be the set of all the country-entities identified in Internet. Then, for each country $C_d \in \mathbf{C}$, $\mathbf{I^d}$ is introduced as the set of all the intermediate ASes that are traversed in order to reach C_d from M for any prefix $p \in \mathbf{P^d}$, where $\mathbf{P^d}$ is the set of all the prefixes hosted by ASes located in C_d and $C(X)$ denotes the country of origin of AS X.

$$\mathbf{I^d} = \bigcup_{k=1}^{K} \left\{A_k^{M,p}\right\}, \quad \forall p \in \mathbf{P^d}, \quad C(A_k^{M,p}) \neq C(M), \quad C(A_k^{M,p}) \neq C^d \tag{4}$$

Namely, $\mathbf{I^d} = \{I_1^d, \ldots, I_Q^d\}$ comprises the Q different ASes that have ever appeared as intermediate hops between $C(M)$ and C^d. Although, the sequence of intermediate hops towards each prefix's owner AS is dependent on the monitoring point, the notation M will be hereafter omitted for ease of reference.

Furthermore, $\mathbf{U^d}$ is introduced as the set of all the countries that are recorded as intermediate hops towards hosts of C_d, i.e. $\mathbf{U^d}$ is the set of all the countries of origin of the ASes comprising $\mathbf{I^d}$.

$$\mathbf{U^d} = \{U_1^d, \ldots, U_Z^d\} = \bigcup_{q=1}^{Q} \left\{C(I_q^d)\right\}, \quad \mathbf{U^d} \subseteq \mathbf{C} \tag{5}$$

The driving notion behind the definition of $\mathbf{I^d}$ and $\mathbf{U^d}$ resides in the necessity to quantitatively define the legitimacy of an AS's occurrence as an intermediate hop. More concisely, in order to draw safe conclusions concerning the legitimacy of an announced AS-Path, it would optimally be required to retain the complete history of the specific prefix and all the involved ASes. Nevertheless, the frequency of a route's announcement cannot be regarded as a safe criterion for judging its normality, since any alternative route announced for the first time would be always condemned to be labelled as suspicious. Moreover, in the case that the competing paths are interchangeably announced, the respective routes will be found to present comparably high probability values and thus the event would be disregarded, despite the fact that such phenomena may correspond to repeating hijacks-responses. Additionally, an analysis performed at per prefix/AS level would impose substantial memory and processing overhead.

On the contrary, by aggregating the BGP activity on per country level, the proposed methodology exploits the inherent geo-spatial coherence of the Internet infrastructure and routing policies. Specifically, following the inter-AS agreements, for every destination-country there is a finite, semi-constant set of Intermediate-Countries (ICs) that provide its connectivity with the rest of the world. Hence, any path alterations involving ICs that are not common for the specific destination-country are bound to raise significant suspicions of interception. Additionally, special emphasis must be laid on the fact that such malicious Internet activity is usually carried out by (or through) remote hosts, in order to decrease the probability of being tracked down as well as to escape any legal actions and countermeasures. In this context, the two features that form the cornerstone of the proposed technique are presented below.

2.1 Probability of an IC's Appearance

For every country $C_d \in \mathbf{C}$, the vector $\mathbf{V_C^d}$ is estimated, which contains the number of appearances of each IC across the path towards hosts of the destination-country C^d.

$$\mathbf{V_C^d} = \begin{bmatrix} N(U_1^d) \\ \vdots \\ N(U_Z^d) \end{bmatrix} \tag{6}$$

Utilizing $\mathbf{V_C^d}$, *the probability* $(B(U_z^d))$ *of a country's appearance in a path towards destination-country* C_d is introduced, in order to allow for the quantitative assessment of the legitimacy of an AS's appearance.

$$B(U_z^d) = \frac{N(U_z^d)}{\sum_{z=1}^{Z}\{N(U_z^d)\}}, \quad \forall U_z^d \in \mathbf{U^d} \tag{7}$$

The calculated probability is equal to the conditional probability that an AS X from country $C(X)$ appears within a routing path for a prefix hosted in C_d $\left(B(U_z^d) = PR\left[X \in \mathbf{AP^p} | (C(O^p) = C_d)\right]\right)$ and hence it is equal to the fraction of the $C(X)$'s appearances towards country C_d against the aggregate appearances of all the ICs for C_d. $B(U_z^d)$ is a measure of how frequently U_z^d serves as an intermediate hop for C_d and therefore how commonly expected is for U_z^d to appear in a BGP announcement that refers to a C_d host.

2.2 Geographic Disparity of a Route's ASes

Routing algorithms generally opt for the path with the lowest latency and thus the minimization of the end-to-end geographic distance is a primary routing objective. Hence, apart from specific agreements/policies or infrastructural malfunctions, there is no operational reason for selecting routes that significantly diverge from the direct route. Furthermore, cyber criminal activity is anticipated to originate from remote countries with favourable legal system and/or international relationships. Consequently, profound geographic divergence along the routing path can be safely regarded as enough evidence for raising an alarm. In this context, in order to numerically approach the geographic anomaly imposed by the appearance of an AS along a BGP route, two measures are defined.

Geographic length introduced by each IC. $\forall U_z^d \in \mathbf{U^d}$, the geographic length $(L(U_z^d))$ introduced by U_z^d in reference to the ideal direct path is defined as:

$$L(U_z^d) = \frac{L(C(M), U_z^d) + L(U_z^d, C_d)}{L(C(M), C_d)} \tag{8}$$

where $L(C_x, C_y)$ is the geographic distance between countries C_x and C_y. $L(U_z^d)$ is the geographic length of the $C(M) \rightarrow U_z^d \rightarrow C_d$ path, normalized against the length of the direct source-destination link. The estimation of $L(U_z^d)$ allows for tracing the countries and thus the corresponding ASes that prominently diverge from the expected path.

Z-Score of the an IC's geographic length. According to the definition of the Z-Score,

$$S^L(U_z^d) = \frac{L(U_z^d) - E[L(U_z^d)]}{\sigma[L(U_z^d)]} \tag{9}$$

where $E[L(U_z^d)]$ and $\sigma[L(U_z^d)]$ are respectively the mean value of all $L(U_z^d), \forall U_z^d \in \mathbf{U^d}$. The statistical analysis of potential geographic anomalies on the basis of Z-Score, allows to assess each IC's role in comparison with the general deviations of the routing path for the specific destination-country under investigation. In particular,

- $S^L(U_z^d) < (>)\ 0$: The geographic divergence introduced by U_z^d is lower (higher) than the average distance of all the perceived routing paths between the $C(M)$ and C_d.
- $|S^L(U_z^d)| \uparrow$ AND $S^L(U_z^d) > 0$: U_z^d's geographic location significantly distances from the average route, while there are only few alternative countries towards C_d that are located far away from the direct route.

2.3 Implementation Issues

Despite its solid theoretical background, the aforementioned methodology is associated with a real-world implementation complication. A substantial fraction of the intermediate hops along an announced AS-Path are higher-tier ASes that provide connectivity across multiple countries/continents. However, despite their global presence, higher-tier ASes are uniquely identified by a sole country of origin, which usually coincides with the location of the enterprise's headquarters. Therefore, including the higher-tier ASes in the calculation of $\mathbf{U^d}$ shall utterly distort the overall statistical analysis.

Let us assume the example of AS-Path {AS15469, AS174, AS6762, AS8966, AS13224}. By including AS174 (tier-1 AS) in $\mathbf{I^d}$ (4), it will be erroneously taken for granted that it is usual for IP traffic originating from Switzerland (CH) to traverse USA (US) so as to reach Kenya (KE). As a result, the efficiency of the BGP hijack detection mechanism will be severely degraded, since: (i) announcements that include higher-tier ASes established in US would be erroneously considered as suspicious due to the high $L(US)$ value and (ii) BGP hijacks executed by ASes located in countries hosting higher-tier ASes would remain unobserved.

In this respect, a subset $\mathbf{I'^d}$ of $\mathbf{I^d}$ is defined that comprises all the ASes that belong to $\mathbf{I^d}$ and which are not classified as higher-tier ASes. Correspondingly, $\mathbf{U'^d}$ is also defined on the basis of $\mathbf{I'^d}$ (5). Eventually, $B(U_z^d)$, $L(U_z^d)$ and $S^L(U_z^d)$ are calculated for the set $\mathbf{U'^d}$. However, for simplicity, the initial notation will be kept. Finally, it must be underlined that this exclusion of the higher-tier ASes does not compromise the reliability of the BGP hijack mechanism, since higher-tier ASes are not expected to deploy criminal activity.

2.4 Transition to AS-Path Anomaly Level

The metrics $B(U_z^d)$, $L(U_z^d)$ and $S^L(U_z^d)$ are defined on a per IC level. Thus, in order to study BGP hijacks, it is necessary to implement the proposed methodology at AS-Path anomaly level. According to (3), each AS-Path anomaly comprises two competing AS-Paths, while in turn, each AS-Path is identified by a sequence of ICs. In this context, let $\mathbf{F_j}$ be the set of non-common ASes $\mathbf{F}(\mathbf{W}, \mathbf{W'})$ involved in the

AS-Path anomaly j. Then, considering that suspicions of malicious BGP activity are raised for ASes with low values of $B(U_z^d)$, and high values of $L(U_z^d)$ and $S^L(U_z^d)$, the corresponding scores at AS-Path level are defined:

- Country Appearance Probability per AS-Path anomaly (CAP)

$$CAP_j = min\{B(C(A_f))\}, \quad \forall A_f \in \mathbf{F_j} \tag{10}$$

- Country geographic length per AS-Path anomaly (CGL)

$$CGL_j = max\{L(C(A_f))\}, \quad \forall A_f \in \mathbf{F_j} \tag{11}$$

- Z-Score of country geographic length per AS-Path anomaly ($CGLZ$)

$$CGLZ_j = max\{S^L(C(A_f))\}, \quad \forall A_f \in \mathbf{F_j} \tag{12}$$

where $A_f \in \mathbf{I^d}$ and $C(A_f) \in \mathbf{U^d}$ and $\mathbf{J}$ are all the monitored AS-Path anomalies.

3 Evaluation Under Real-World BGP Hijack Events

On August 20, 2011, a Russian telecommunication company, hereafter referred to as victim-AS (AS_V), reported to the North American Network Operators Group (NANOG) that five of its prefixes had been hijacked. The prefixes' ownership was not affected, but forged routes were announced that dictated any traffic to traverse through the hijacking AS (AS_H), which is located in US, for interception purposes. As a countermeasure, AS_V responded on August 24, by announcing longer subprefixes with the correct paths. All AS-Path anomalies taking place on that date are recorded and, for each one, the values of CAP, CGL and $CGLZ$ are calculated. The response of the AS_V triggers an AS-Path anomaly as a result of the juxtaposition of the new legitimate route (rather straight path) against the existing path (detour through US) previously forged by AS_H. The raw BGP data are obtained through the RIPE repository [12] and AS15469, which is situated in CH (Vantage Point *rrc00*), is chosen as the monitoring point (M).

The distributions of CAP and CGL are presented in Fig. 1, so as to assess their capability to efficiently discriminate the bulk of AS-Path anomalies. From Fig. 1a it becomes apparent that the vast majority of the AS-Path anomalies involve rather common ICs, while, according to Fig. 1b, almost 98 % of the AS-Path anomalies involve ICs that do not introduce additional geographic path length higher than the direct source-destination distance ($CGL < 2$). Figure 2a, b presents the scatter plot of CAP against CGL and $CGLZ$ respectively. With the solid black triangle it is marked the AS-Path anomaly that corresponds to the incident under investigation. As it becomes apparent, taking into account the CAP in conjunction with $CGLZ$ of all the monitored BGP path alterations, the proposed methodology is capable of

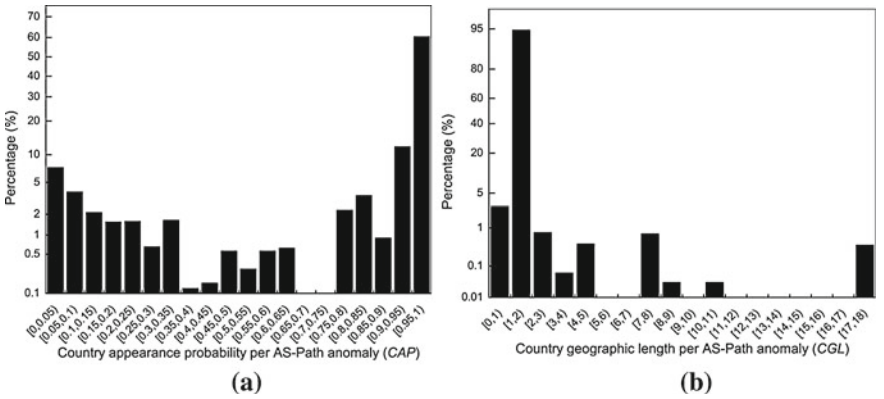

Fig. 1 Distribution of (**a**) *CAP* and (**b**) *CGL*

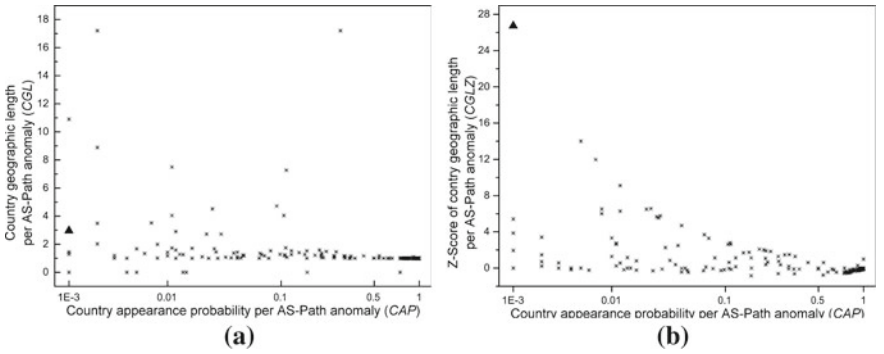

Fig. 2 Scatter plots between proximity measures of different features. **a** Scatter plot of *CAP* against *CGL*. **b** Scatter plot of *CAP* against *CGLZ*

decisively pinpointing the BGP hijack. $CGLZ$ is chosen instead of CGL since it incorporates the overall distribution of the routes' geographic length.

4 Conclusions

In this paper, a fully-unsupervised methodology for the detection of malicious AS-Path anomalies has been described. The novelty of the proposed technique lies within the basic notion of statistically analysing the BGP activity on a per hosting-country basis, so as firstly to make virtue of the inherent geo-spatial coherence of the Internet infrastructural functionality and secondly to formulate a solid background regarding the normal routing status. In this context, two completely novel features are extracted and corresponding proximity measures are defined, while no assump-

tions whatsoever are being introduced. The presented mechanism is implemented and evaluated against a positively known event of AS-Path hijack that has been publicly reported. Through this real-world case study, the behavior of each feature and metric are studied and assessed thoroughly and eventually the efficiency of the proposed methodology in capturing BGP attacks is proven.

References

1. Rekhter, Y., Li, T.: A border gateway protocol 4 (bgp-4). IETF RFC 1771. http://www.ietf.org/rfc/rfc1771.txt
2. Sriram, K., Borchert, O., Kim, O., Gleichmann, P., Montgomery, D.: A comparative analysis of bgp anomaly detection and robustness algorithms. In: Proceedings of the CATCH '09, (March 2009)
3. Ballani, H., Francis, P., Zhang, X.: A study of prefix hijacking and interception in the internet. In: Proceedings of the SIGCOMM '07, Kyoto (2007)
4. Gao, L.: On inferring autonomous system relationships in the internet. IEEE/ACM Trans. Netw. **9**(6), 733–745 (2001)
5. Deshpande, S., Thottan, M., Ho, T.K., Sikdar, B.: An online mechanism for bgp instability detection and analysis. IEEE Trans. Comput. **58**(11), 3296–3304 (2009)
6. Zhang, K., Yen, A., Zhao, X., Massey, D., Felix Wu, S., Zhang, L.: On detection of anomalous routing dynamics in bgp. Lecture Notes in Computer Science. Springer, Berlin (2004)
7. Li, J., Dou, D., Wu, Z., Kim, S., Agarwal, V.: An internet routing forensics framework for discovering rules of abnormal bgp events. In: Proceedings of the SIGCOMM '05, (Oct. 2005)
8. de Urbina Cazenave, I., Kosluk, E., Ganiz, M.: An anomaly detection framework for bgp. In: Proceedings of the INISTA '11, (June 2011)
9. Lad, M., Massey, D., Pei, D., Wu, Y., Zhang, B., Zhang, L.: PHAS: A prefix hijack alert system. USENIX Security Symposium, In (2006)
10. Qiu, J., Gao, L., Ranjan, S., Nucci, A.: Detecting bogus bgp route information: going beyond prefix hijacking. In: Proceedings of the SecureComm '07, (Sept. 2007)
11. Xiang, Y., Wang, Z., Yin, X., Wu, J.: Argus: An accurate and agile system to detecting ip prefix hijacking. In: Proceedings of the IEEE ICNP '11, (Oct. 2011)
12. Ripe, NCC. http://www.ripe.net/datatools/stats/ris/ris-raw-data