Software Engineering and Applications 软件工程与应用, 2016, 5(2), 93-102
Published Online April 2016 in Hans. http://www.hanspub.org/journal/sea
http://dx.doi.org/10.12677/sea.2016.52011

Hans汉斯

Clustering Analysis and Anomaly Detection Based on the Detour Path

Lei Liu, Peidong Zhu, Shuang Yan, Wei Fu

College of Computer Science, National University of Defense Technology, Changsha Hunan
Email: liulei0855@sina.com

Received: Mar. 8th, 2016; accepted: Mar. 22nd, 2016; published: Mar. 29th, 2016

Copyright © 2016 by authors and Hans Publishers Inc.
This work is licensed under the Creative Commons Attribution International License (CC BY).
http://creativecommons.org/licenses/by/4.0/
Open Access

Abstract

In this paper, the detour path is defined firstly. Through the observation of AS_PATH property of the routing table, we sum up the six forms of the detour path, *i.e.*, continuously repeated AS, loop, around the neighbor AS, around the country, around the border and around the multinational company. Moreover, we did the clustering analysis of the manifestation of detour path and put forward the routing anomaly detection method based on the detour path. The method can detect the continuous repeated AS, routing loop, domestic traffic leaked, forged path, garbled path with such anomalies. Experiments show that the proposed method can effectively detect abnormal routing behavior and also suggest that one of the main reasons for the average shortest path of Internet traffic which becomes longer is the existence of the detour paths.

Keywords

BGP, Detour Path, Clustering Analysis, Anomaly Detection

基于绕行路径的聚类分析与异常检测

刘　磊，朱培栋，闫　爽，富　威

国防科学技术大学计算机学院，湖南　长沙
Email: liulei0855@sina.com

收稿日期：2016年3月8日；录用日期：2016年3月22日；发布日期：2016年3月29日

文章引用: 刘磊, 朱培栋, 闫爽, 富威. 基于绕行路径的聚类分析与异常检测[J]. 软件工程与应用, 2016, 5(2): 93-102.
http://dx.doi.org/10.12677/sea.2016.52011

摘　要

本文对BGP路由中的绕行路径作了定义，通过观察路由表的AS_PATH属性，总结归纳了绕行路径的六种表现形式，即连续重复AS、环路、绕邻居AS、绕国、绕境、绕跨国企业。同时，对绕行路径的表现形式进行了聚类分析，并提出了基于绕行路径的连续重复AS、路由环路、国内流量外泄、路径伪造、路径篡改等异常路由检测方法。实验表明，本文所提方法能够有效检测异常路由行为，同时揭示了绕行路径的存在是使得网络流量传递平均最短路径变长的主要原因之一。

关键词

BGP，绕行路径，聚类分析，异常检测

1. 引言

在互联网关键基础应用迅速蓬勃发展的今天，互联网安全越来越发地成为一个被人们所关注的领域。BGP [1]作为域间路由的实际运行协议，协调着各 AS 间路由可达性信息的有序交换，对于互联网的正常运行起到了至关重要的作用。然而，由于 BGP 协议设计之初缺乏足够多的安全机制的考虑，无法对传播的路由信息进行完整性和真实性的有效验证[2]，使得网络的高效运转倍受域间路由的错误配置和攻击行为所扰[2] [3]。例如 2014 年以来的 Amazon 的 MITM 事件[4]、加拿大 ISP 的前缀劫持事件[5]以及印度[6]、马来西亚[7]的路由泄露事件，这些安全事件都对 BGP 路由系统的可用性和稳定性造成了极大的影响。

目前，路由异常主要包括前缀异常和路径异常两大类。前缀异常主要是前缀劫持，是指一个 AS(Autonomous System, AS)对外通告了一个未获授权的前缀，即前缀属于其他 AS 所有或者该段地址空间尚未分配，AS 违反授权对外通告非法的前缀将直接造成流量劫持的发生，甚至网络的瘫痪[8]。路径异常则主要通过对路由的 AS-PTH 属性进行攻击。而在这两类路由异常类型中，路径异常相比更加隐蔽，不容易检测。因此，针对路径的异常检测一直是路由安全系统的研究重难点之一。

近年来，针对路径异常的检测已经有了很多研究。如 Kruegel，Mutz，Robertson 等人[9]提出了一种基于 AS 拓扑结构的异常检测方法。他们将边缘 AS 层划分为不同的簇，认为一条合法的 AS 路径只能包含一个核心层 AS，并且连续边缘 AS 应履同一个簇或两个非常相近的簇。实验验证了上述方法检测路径异常的有效性，但算法需要实时更新拓扑信息，边缘层 AS 的分簇也会对异常检测的结果造成很大影响。Li 等人[10]提出了基于伙伴的路径异常检测方法，认为路径遭到攻击后，异常路径与其伙伴的相似性将破坏，此方法具有良好的鲁棒性和实时性，但伙伴的判断依据的准确性有待进一步加强，并且能始终找到足够数量的异常路径的伙伴进行验证，路径异常检测的效果不佳。上述路径异常的检测方法主要针对路径异常中的伪造路径异常进行检测分析，很少对路径异常进行做进一步的聚类研究，同时，检测处理的数据量较大，方法较为复杂。Cha 等人[11]利用指纹来检测路径异常，该需要处理大量的数据信息，而且他们认为检测前出现在网络中的边均是合法的，并未考虑到边的动态性，使检测结果缺乏一定的准确性。针对上述问题，本文从路径绕行的角度出发，将绕行路径分为连续重复的 AS、环路、绕邻居 AS、绕国、绕境、绕跨国企业六种路径异常表现形式。同时，本文以 AS 邻接关系为基础，提出了基于绕行路径的连续重复 AS、路由环路、国内流量外泄、路径伪造、路径篡改等异常现象的检测方法。实验结果表明，该方法能够有效地检测路径异常行为。

2. 绕行路径聚类分析

路径属性 AS_PATH 是 BGP 协议的公认必遵属性，它用一系列有序 AS 号来描述 AS 间的路径或 NLRI(Network Layer Reachability Information)。当运行 BGP 协议的路由器发起或转发一条路由更新时，它便将自己的 AS 号附加到 AS_PATH 属性的最前面，而后再传递给自己的上游或对等 AS。AS_PATH 描述了一条从本地 AS 到宣告前缀 AS 所经过的自治系统的顺序，即去往目标前缀网络的一条路径。由于 BGP 协议设计之初并没有任何验证 AS_PATH 属性真实性的机制，致使域间路由系统及其上的 AS 节点极易受到路径伪造、路径篡改、国内流量外泄、路由黑洞、路由泄露、路由中间人等等攻击，造成网络的不可达、拥堵或数据的窃取、篡改，致使域间网络笼罩在各种安全威胁之下，而路径的绕行表征可能是造成上述威胁的潜在表现。

绕行路径是指从源 AS 到目标 AS 的一条 AS_PATH 路径本可以最短路径到达目标 AS 网络，而出于商业利益、流量工程考虑或攻击所致，导致实际到达目标 AS 网络的路径比最短路径要长的一条异常路径。通过对 routeviews 项目上的 2015 年 9 月 27 日 22 时的路由数据进行观察分析，绕行路径通常有以下六种表现形式：

1) 连续重复 AS

在 AS_PATH 路径中出现连续且相同的 AS，使得本应短的路径变长。一个影响是流量在重复 AS 内绕行，延迟流量到达时间；再一个影响是攻击者可以通过减少路径中重复 AS 的数量，增加攻击 AS，并且使得篡改后的路径长度小于篡改前的路径，达到中间人攻击的目的[12]。例如：AS_PATH 路径 13030 852 852 53359 与 AS_PATH 路径 852 7922 33287 33560 33560 33560 均在路径中的中间和末尾处出现了连续重复的 AS 852 和 AS 33560，攻击者可以伪造一条路径 852 7922 XXX 33287 33560，伪造的路径可以正常转发流量到目标网络，但其路径长度明显短于原来的路径，可被其它 AS 优先选取，而可能造成路由攻击。

2) 环路

在 AS_PATH 路径中出现重复且不连续的 AS，使得路径中存在环路和本应短的路径变长。流量在几个 AS 间绕行，甚至造成无限绕行而无法摆脱 AS 的束缚，使得流量不能够到达目标网络。例如：AS_PATH 路径 8492 6939 2711 6167 22394 6167 中，AS 6167 出现两次且中间经过 AS 22394，按常理流量第一次到达目标 AS 6167 后不应该再流出到 AS 22394 后回到 AS 6167，环路的出现，使得 AS 22394 可以截获去往 AS 6167 的流量。

3) 绕邻居 AS

在 AS_PATH 路径中出现本应能够直接相连接的两个邻居 AS 却被其它 AS 隔断，使得路径变长。例如：AS_PATH 路径 11537 10764 513 12654 中，AS 11537 与 AS 513 是相互连接的邻居，可以直接从 AS 11537 到达 AS 513，但此路径中却被 AS 10764 的隔开，使得从 AS 11537 去往 AS 513 的流量流经 AS 10764。

4) 绕国

在 AS_PATH 路径中，同一国家的 AS 中间出现其它国家的 AS，按常理同一个国家的 AS 应是相互连接的，且流量应首选本国的 AS 进行传递后再经其它国家传递，而此类型的路径绕行可能造成国家流量外泄，致命国家层面上面临着重大威胁。例如：AS_PATH 路径 23673 38726 7497 11537 22388 7660 24287 24489 23911 9401 中，AS 7497、AS 24489、AS 23911、AS 9401 所属国家为中国，而其中间的 AS 11537、AS 22388 所属国家分别为美国，AS 7660、AS 24287 所属国家为日本。路径片段 7497 11537 22388 7660 24287 24489 23911 9401 中，流量从中国(AS 7497)经由美国和日本后，又回到中国，这种行为显然是不合理的，因为从 AS 7497 出去的流量本应直接到达国内的其它 AS 而不需要经过其它国外 AS。

5) 绕境

在 AS_PATH 路径中，在同一国家的不同境内外的 AS 间来回进行流量的传递。例如：AS_PATH 路径 3277 3267 20388 7497 4641 4641 24151 中，AS 7497 和 AS 24151 属中国大陆拥有 AS，AS 4641 属中国香港 AS。流量从中国内陆 AS 7497 本应直接传递到内陆 AS 24151，而不应经由境外香港 AS 4641 后回到内陆 AS 24151。

6) 绕跨国企业

在 AS_PATH 路径中，在本国 AS 与本国跨国企业的它国的 AS 间进行的流量传递。例如：AS_PATH 路径 852 2914 4134 36678 55992 中，AS 4134 和 AS 55992 均属中国 AS，AS 36678 属美国所有，但 AS 36678 注册组织为中国电信美国分公司。

通过对上述绕行路径的表现形式的分析可以看出，路径绕行的共同特征是增加了路径的长度，使得本来短的路径变长，这其中的原因可能是基于流量工程或利益的考量，异或是攻击行为所致，因此，对于路径的绕行行为应着重对待。

3. 基于绕行路径的异常检测

绕行路径的异常检测基于 AS 基本信息表、AS 邻接点对集、AS 间商业关系等基本知识库，主要检测路径绕行中可能存在的路由环路、国内流量外泄、路径伪造(也可能是路由黑洞)、路径篡改(甚至可能是中间人攻击)等异常现象。这些异常现象的产生极有可能导致流量导向非正确网络或数据内容遭到窃取篡改。如果由于路由泄露将大量流量导向特定网络将可能会造成大规模的拒绝服务式攻击，如果由于路径伪造将流量引向伪造 AS 内且丢弃将可能会造成路由黑洞攻击，如果由于路径篡改将流量引入篡改 AS 内后再将流量转发至目标网络将造成数据被窃取或篡改，这些都将给网络和用户造成严重的影响。因此，需要一种路径异常检测的方法来有效防范或预警异常现象的发生，尽量减少路径异常所带来的损失。

3.1. 基本知识库的构建

AS 基本信息主要从网络上公开的信息中获取，包括注册 ISP 名称、国家及城市等基本信息；AS 间的邻居信息主要从 routeviews.org 公开的路由表中获得，将出现在 AS_PATH 属性中的相邻 AS 记作邻居关系，其具体算法如算法 1 所示；由于 AS 商业关系是 ISP 的商业机密，非公开的，无法从公共网络上直接获得，本文采用文献[13]所提的方法进行 AS 商业关系的推断。

3.2. 异常检测

主要对路径绕行中可能存在的路由环路、国内流量外泄、路径伪造(也可能是路由黑洞)、路径篡改(甚至可能是中间人攻击)等异常进行检测。

3.2.1. 重复 AS 异常的检测

重复 AS 的异常包括连续重复 AS 的异常和路由环路异常两种情况。给定路径 $P=\left(p_1 p_2 \cdots p_i \cdots p_j \cdots p_n\right)$，其中，$p_1$ 至 p_n 代表 AS 号。如果在 P 中出现两个及以上相同 AS 号 $p_i=p_j=\cdots=p_k$，则为重复 AS 异常，进一步考虑，令 $\left(a_1,a_2,\cdots,a_n\right)=\left(i,j,\cdots,k\right)$，如果 $=\ n-1$，则为连续重复 AS 异常，反之，则为路由环路异常。

3.2.2. 国内流量外泄异常的检测

在给的路径 $P=\left(p_1 p_2 \cdots p_i \cdots p_j \cdots p_n\right)$ 中，如果存在 p_i 与 p_j 同属一个国家 C，且 $\left|i-j\right|>1$，令 p_i 与 p_j 之间的 AS 序列为 $Q=\left(q_1 q_2 \cdots q_n\right)$，扫描新的序列 Q，如果发现其中存在不属于国家 C 的 AS，则路径判断为国内流量外泄异常。

Algorithm 1. AS adjacent relation extraction
算法 1. AS 邻居关系提取

输入：BGP 路由表的 AS_PATH 集合
输出：AS 邻接关系集

```
1) AS_PATH_List = AS_PATH.strip(r' ')
2) AS_PATH_Length = len(AS_PATH_List)
3) for i in range(AS_PATH_Length):
4)       AdjacencyAS_Set = set()
5)       if i == 0:
6)             AdjacencyAS_Set.add(AS_PATH_List [i+1])
7)             d[AS_PATH_List [i]] = AdjacencyAS_Set
8)       elif i == AS_PATH_Length:
9)             AdjacencyAS_Set.add(AS_PATH_List [i-1])
10)            d[AS_PATH_List [i]] = AdjacencyAS_Set
11)      else:
12)            AdjacencyAS_Set.add(AS_PATH_List [i-1])
13)            AdjacencyAS_Set.add(AS_PATH_List [i+1])
14)            d[AS_PATH_List [i]] = AdjacencyAS_Set
15) return AdjacencyASList
```

3.2.3. 邻居绕行异常的检测

在给的路径 $P=\left(p_1 p_2 \cdots p_i \cdots p_j \cdots p_n\right)$ 中，如果存在 P 的子集序列 $Q=\left(q_1 q_2 \cdots q_n\right)$，如果存在 q_1 与 q_n 具有邻居关系，即在其它的路径中存在 q_1 与 q_n 直接相连的情况，则有理由相信去往 q_n 的流量可以直接从 q_1 到达 q_n 而不需要经过其中间的若干个 AS。若出现此种情况，则判断为邻居绕行异常。

3.2.4. 路径伪造异常的检测

通常情况下，路径伪造使数据流向伪造 AS 内，造成数据被窃听或被丢弃，在邻居绕行异常的路径片段序列 $Q=\left(q_1 q_2 \cdots q_n\right)$ 中，q_1 与 q_n 是邻居关系，那么，如果序列 $\left(q_2 \cdots q_{n-1}\right)$ 中存在 q_i，且 q_i 与 q_{i-1} 是邻居关系，而 q_i 与 q_{i+1} 为非邻居关系，如果，流量传递到 q_i 不再继续延着 q_{i+1} 进行传递，可能造成的威胁行为是窃听数据或丢弃数据，如果大量的数据流量都转向了这类伪造的路径，使得流量不能正确的进行传递，很可能形成路由黑洞。

3.2.5. 路径篡改异常的检测

路径篡改较路径伪造而言，它不会造成数据流量的丢弃，而是通过改变路径传递方向，使流量传递到攻击者的 AS 内，进行流量的截获或修改，而后再将流量转发到正确的目标网络，对流量的发起方和接收方并没有产生可观察到的影响，近年来，比较凸显的路由中间人攻击就是以这种形式进行攻击的，这使得网络管理人员很难发现此类异常行为，从绕行路径的角度来看，路径的篡改或中间人攻击行为的表现形式就是在原本可直接到达的两个 AS 间插入了若干 AS，使流量经同插入的 AS 后传递到目标网络，显然路径绕行具有以上行为的基本特征，如果在邻居绕行异常的路径片段序列 $Q=\left(q_1 q_2 \cdots q_n\right)$ 中，q_1 与 q_n 是邻居关系，那么，如果序列 $\left(q_2 \cdots q_{n-1}\right)$ 中存在 q_i 与 $q_{i+1}\left(1<i<n-2\right)$ 均是邻居关系，即序列 $\left(q_2 \cdots q_{n-1}\right)$ 是一条通路，且没有违反 AS 间的商业关系，则判断此为路径篡改异常。

3.3. 异常检测流程

基于路径绕行的异常检测首先根据路径本身的表现形式特点检测重复 AS 的异常，而后依据 AS 所属国家信息检测国内流量外泄异常，同时，根据路由表推断出各 AS 的邻居关系，进而检测邻居绕行的异常，并在邻居绕行路径的基础上进行路径伪造和路径篡改异常的检测。其流程图如图 1 所示。

判定规则：

1) AS-PATH 中出现重复的 AS；

2) 非连续重复 AS；

3) 邻居绕行：形如 A-C-D-B 路径中，存在 A-B 的路径且 A!=B!=C；

4) 国内流量外泄：形如 A-C-D-B 路径中，存在 A、B 属于同一国家，而路径序列 C-D 中存在 AS 属于另一国家；

5) 绕行 AS 在 AS-PATH 中与其相连的 AS 均是邻居关系；

6) 绕行 AS 在 AS-PATH 中与其相连的上游 AS 是邻居关系，而与其下游 AS 甚少有一个是非邻居关系；

7) 违反 AS 间商业关系。

注：所谓绕行 AS 是指两个互为邻居的 AS 中间的 AS。

4. 实验结果与分析

实验数据来自开源项目 routeviews 上 2015 年 11 月 1 日至 30 日，每日 0 点的路由更新，通过对路由数据进行基于绕行路径的部分异常检测结果如表 1 至表 4 所示。

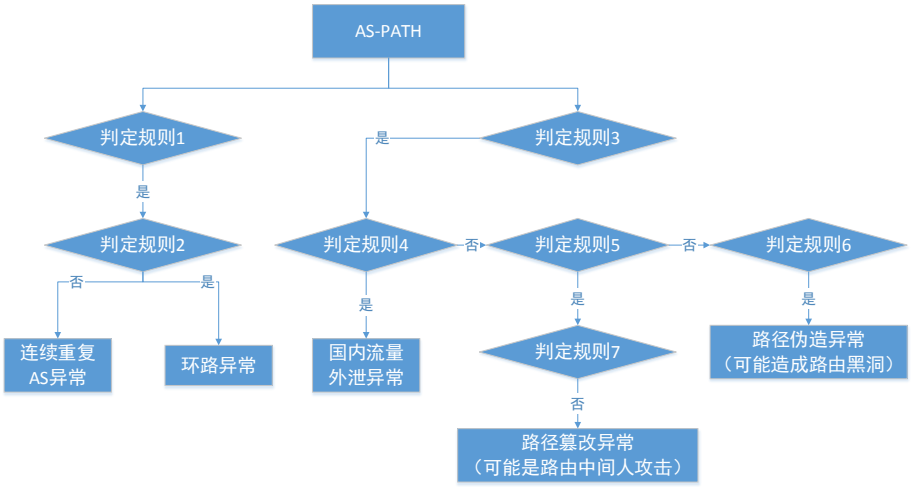

Figure 1. Flow chart of anomaly detection
图 1. 异常检测流程图

Table 1. The abnormality of continuous repeated AS
表 1. 连续重复 AS 异常

1	11537 2603 2603 2603 2603 1103 12654|1--4|
2	11686 12989 22381 22381 22381 22381 22381 22381 262429 52770 28605 19990|2--7|
3	11686 12989 28640 262401 262401 262401 262401 262401 262401 262401 262401 262949|3--10|
4	11686 19151 15399 37084 37545 37586 37586 37586|5--7|
5	11686 19151 174 12714 8905 8905 8905 39792 51230 56636|4--6|
6	11686 19151 174 20764 21453 21453 47321 47321 47321 48098 48098 48098 48098 48098 48098 51230 56636|4--14|
7	11686 19151 174 3741 3741 3741 3741 3741 3741 30988 37125|3--8|
8	11686 19151 20485 5563 56947 56947 56947|4--6|
9	11686 19151 3356 12956 60725 52253 52253|5--6|
10	11686 19151 3356 1299 1299 7029 6316|3--4|
11	11686 19151 3356 1299 60205 60205 60205 60205 60205|4--8|
	……

Table 2. The abnormality of forged path
表 2. 路径伪造异常

1	11537 10764 6509 376|0--2|
2	11537 2603 2603 2603 2603 1103 12654|0--5|
3	11686 19151 174 12714 8905 8905 8905 39792 51230 56636|1--7|
4	11686 19151 174 3741 3741 3741 3741 3741 3741 30988 37125|1--8|
5	11686 19151 8468 31708 31708 31708 31708 31708 31708 50882 50882 50882 50882 50882 50882 201902|1--8|
6	11686 3356 1299 1299 7029 6316|1--5|
7	11686 3356 1299 198781 36776 36776 36776 36776 36776 36776|1--9|
8	11686 3356 1299 6939 6939 3212 44647|0--4|
9	11686 3356 1299 6939 6939 44356 44356 37076 37125 37125 37125 37125 37125|0--4|
10	11686 3356 1299 8262 48452 48452|2--5|
11	……

Table 3. The abnormality of garbled path
表 3. 路径篡改异常

1	11537 10764 513 12654|0--2|
2	11537 20965 1103 12654|0--2|
3	11537 20965 2603 1880|0--2|
4	11537 20965 27750|0--2|
5	11537 2153 2152 567 226 47065|1--3|
6	11537 2603 1653 1880|1--3|
7	11537 2603 2603 2603 2603 1103 12654|0--5|
8	11537 27750 1916 12654|0--2|
9	11686 19151 10026 9498 9829|1--3|
10	11686 19151 1273 3216 8905 51230 56636|1--3|
11	……

Table 4. The abnormality of domestic traffic leaked
表 4. 流量外泄异常

1	11686 19151 3356 10794|0--2|
2	11686 19151 3356 11911|0--2|
3	11686 19151 3356 12956 60725 52253 52253|0--2|
4	11686 19151 3356 1299 1299 7029 6316|0--2|
5	11686 19151 3356 1299 18351 59149|0--2|
6	11686 19151 3356 1299 60205 60205 60205 60205 60205|0--2|
7	11686 19151 3356 15510|0--2|
8	11686 19151 3356 174 12389 8382|0--2|
9	11686 19151 3356 201287|0--2|
10	11686 19151 3356 201369|0--2|
	……

表 1 展示了部分连续重复 AS 异常的路径。如路径 11537 2603 2603 2603 2603 1103 12654 的第 2 至第 5 个 AS 均为 2603，使得流量在 AS2603 内连续跳转 4 次才流出 AS2603，这样的行为让始于 AS11537 通向 AS12654 的流量在网络上行走时间延长。这样的路径，多半是 ISP 为了本流量工程所采用的一种常见的配置策略。然而，这样的配置亦可成为路由攻击的一种手段，攻击者可能为了攻击而找到这样一条路径将其修改为 11537 2603 2603XXX 1103 12654，其中，“XXX”为攻击者所在 AS，这样修改后，攻击者添加了自己的 AS 而又使路径长度由原来 7 变为现在的 6，根据路径选择的原则，变短后的路径将后取代原来的路径成为路由，这样攻击者便能够达到攻击的目的。

表 2 展示了部分路径伪造的异常路径。如路径 11537 10764 6509 376，AS11537 和 AS6509 在路由表中存在一条互为邻居的路径 11537 6509 376，而中此异常路径中，二者被 AS10764 所分开，致使流量本应从 AS11537 直接流向 AS6509 而不需要经过 AS10764，这样的异常路径称作绕行路径。此外，路径的伪造还体现在 AS11537 和 AS10764 具有邻居关系，而 AS10764 和 AS6509 并非邻居关系，也就是说 AS10764 没有到达 AS6509 的路径，因此，这样的路径是不正常的，从表象上来看，符合路径伪造的特征。

表 3 展示了部分的路径篡改路径。可知路径 11537 10764 513 12654 中，AS11537 和 AS513、AS10746 均是邻居关系，且 AS10746 和 AS513 亦是邻居关系，这样，本可以从 AS11537 直接流向 AS513 的流量却经过了 AS10746，使得流量的流转绕道了，正常情况下，这种路径是不应该出现的。可能的原因是错误的路由配置所导致，或是路径篡改的攻击的行为所致，不管是哪种都是一种路径遭到篡改的异常表现，都值得关注。

表 4 展示了部分的流量外泄异常的路径。如其中一条路径 11537 22388 7660 9264 7497 4635 38345，AS7497 和 AS38345 均属于中国内陆所有，而 AS4635 属于中国香港所有，按理说，由中国大陆 AS 出发的流量应该有直接到达中国大陆内目标 AS 的路径，而不需要经过境外到达香港后再流回境内。又如路径 11686 3356 4134 36678 55992 55992，中国 AS4134、AS55992 中间流经美国 AS36678，此种情况属流经外国 AS，从国家安全的角度来考虑，这种流量外泄的路径异常行为可能造成国内流量受到侦听或篡改，亦是倍受关注的重点异常行为之一。

通过对实验的路由异常数据的统计可得，连续重复 AS 异常数量平均占比 29.6%，路由环路异常数量平均占比 0%，邻居绕行异常数量平均占比 35.3%，路径伪造异常数量平均占比 5.3%，路径篡改异常数量平均占比 31.6%，国内流量外泄数量平均占比 0.8%。上述异常数量统计情况如表 1 所示，从表 1 中可以看出，路由环路异常出现次数为 0，是因为路由环路容易被发现，且出现环路的危害比较大，这在路由配置策略中做了很好的预防。

从图 2 中异常路径和路径数量的变化情况还可以发现，在 2 日至 7 日这段期间内，虽然路径的数量变化曲线明显，但其异常路径数量却没有太明显的变化，说明此段时间内的路由更新量比较大，而 18 至 22 日这段时间内，异常路径数量较前后时间急剧增多，这样情况可能是路由攻击的征兆，因为在正常情况下，路径异常的数量会维持在一个相对稳定的状态。当有新的路径异常路由的出现后会将此异常路由传播给更多的 AS 进行路径选择，而出现这种异常路径急剧增加的情况，说明受到影响的路由也急剧增加，推测可能受到范围比较广的路由攻击。

此外，如图 3 所示，提取出绕行路径并计算出其平均最短路径的值为 4.02，同时，去除绕行路径中绕行的部分并重新计算，其最短路径的值为 2.06，而路由更新表的平均最短路径为 4.02，可见由于绕行路径的存在使得流量传播的最短平均路径增加了近 2 倍，无论此种现象是基于商业利益或流量工程的考虑，还是人为的恶意攻击行为所致，都不仅给网络的承载能力带来了极大的负担，也让网络中受到攻击的可能性极大增加。

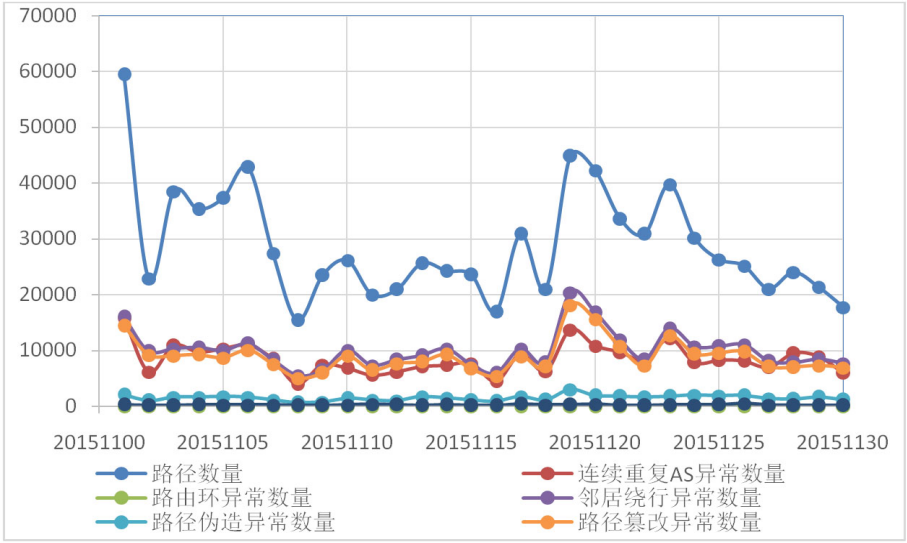

Figure 2. Statistical chart of anomaly detection
图 2. 异常统计图

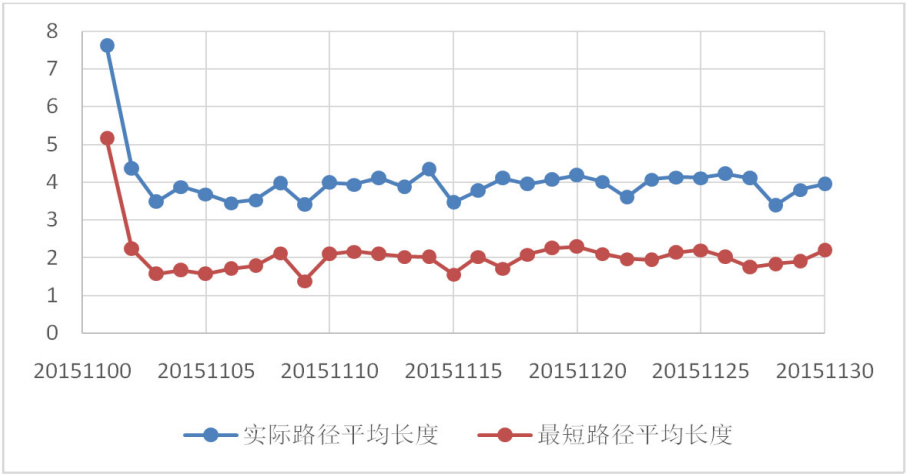

Figure 3. Statistical chart of average path length
图 3. 平均路径长度统计图

5. 结束语

本文首先对绕行路径作了定义，表明实际路由表中的路由本可以更短的路径到达目标网络进行流量的传递。针对路径异常进行了基于绕行路径的聚类分析，以 AS 邻接关系为先验知识，根据路径绕行的行为方式将绕行路径的表现形式分为连续重复 AS、环路、绕邻居 AS、绕国、绕境、绕跨国企业。同时提出了基于绕行路径的异常检测方法及流程。实验表明，本文所提方法能够有效检测出连续重复 AS、路由环路、国内流量外泄、路径伪造、路径篡改等路由异常行为，能够对域间路由安全威胁做出预警，提高网络安全性，同时，实验还表明，由于绕行路径的存在，使得路由平均最短路径长度变长，使得网络的安全性受到威胁。

基金项目

国家自然科学基金(编号：61572514)。

参考文献 (References)

[1] Vohra, Q. and Chen, E. (2007) RFC 4893: BGP Support for Four-octet AS Number Space. Internet Engineering Task Force (IETF).

[2] Bono, J.V. (1997) 7007 Explanation and Apology. NANOG.

[3] Lad, M., Oliveira, R., Zhang, B. and Zhang, L. (2007) Understanding Resiliency of Internet Topology against Prefix Hijack Attacks. *Proceedings of* 37*th Annual IEEE/IFIP International Conference on Dependable Systems and Networks*, 2007, *DSN*'07, Edinburgh, 25-28 June 2007, 368-377.

[4] Toonk, A. BGP Optimizer Causes Thousands of Fake Routes. http://www.bgpmon.net/bgp-optimizer-causes-thousands-of-fake-routes/

[5] Toonk, A. The Canadian Bitcoin Hijack. http://www.bgpmon.net/the-canadian-bitcoin-hijack/

[6] Toonk, A. Large Scale BGP Hijack Out of India. http://www.bgpmon.net/large-scale-bgp-hijack-out-of-india/

[7] Toonk, A. Massive Route Leak Causes Internet Slowdown. http://www.bgpmon.net/massive-route-leak-cause-internet-slowdown/

[8] 黎松, 诸葛建伟, 李星. BGP 安全研究[J]. 软件学报, 2013, 24(1): 121-138.

[9] Kruegel, C., Mutz, D., Robertson, W. and Valeur, F. (2010) Topology-Based Detection of Anomalous BGP Messages. In: *Recent Advances in Intrusion Detection*, Springer, Berlin Heidelberg, 17-35. http://dx.doi.org/10.1007/978-3-540-45248-5_2

[10] Li, J., Ehrenkranz, T. and Elliott, P. (2012) Buddyguard: A Buddy System for Fast and Reliable Detection of IP Prefix Anomalies. 2012 20*th IEEE International Conference on Network Protocols* (*ICNP*), Austin, 30 October-2 November 2012, 1-10.

[11] Hong, S.C., Hong, J.W.K. and Ju, H. (2011) IP Preifx Hijacking Detection Using the Collection of AS Characteristics. 2011 13*th Asia-Pacific Network Operations and Management Symposium* (*APNOMS*), Taipei, 21-23 September 2011 1-7.

[12] Zhang, Y. and Pourzandi, M. (2012) Studying Impacts of Prefix Interception Attack by Exploring BGP AS-PATH Prepending. *Proceedings of* 2012 *IEEE* 32*nd International Conference on Distributed Computing Systems* (*ICDCS*), Macau, 18-21 June 2012, 667-677.

[13] 刘磊, 朱培栋, 胡照明. 一种基于时空可信度推断 AS 商业关系的方法[J]. 软件工程与应用, 2016, 5(1): 38-46.