

Method for Inferring AS Relationship Based on Temporal and Spatial Reliability

Lei Liu, Peidong Zhu, Zhaoming Hu

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

Received: Jan. 29th, 2016; accepted: Feb. 12th, 2016; published: Feb. 19th, 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, a new method of inferring AS relationship by analyzing the temporal and spatial reliability is proposed. Firstly, the method selects the routing tables which originate from different monitored points in different intervals. Based on the routing strategy and the hierarchical structure, the AS relationship of different monitored points is inferred in every time window. Moreover, we define and calculate the value of the temporal and spatial reliability by analyzing the spatial consistency and time stability. Furthermore, we set a threshold and regard any AS relationship with reliability value greater than the threshold as trusted relationship. In the meanwhile, we analyze the number of monitored points and the threshold to infer the influence on AS relationship's validity. Experiment result shows that the accuracy of AS relationship deduction can be largely improved by selecting appropriate number of routing monitored points and the threshold value through our proposed methods.

Keywords

AS Relationship, Reliability, Method

一种基于时空可信度推断AS商业关系的方法

刘 磊, 朱培栋, 胡照明

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1月29日; 录用日期: 2016年2月12日; 发布日期: 2016年2月19日

摘要

本文提出了一种基于时空可信度的AS商业关系推断方法，该方法选取多个时间点的不同监测点路由表，基于路由策略和层次结构初步推断每个时刻不同监测点的AS商业关系，并基于空间一致性和时间稳定性计算其可信度值，根据给定阈值选取可信的AS商业关系。同时，本文还分析了路由监测点数量与阈值对有效AS商业关系推断的影响。实验结果表明，选取恰当数量的路由监测点与阈值可以有效提高AS商业关系的精准度。

关键词

AS商业关系，可信度，方法

1. 引言

Internet 是由一系列自治系统(Autonomous System, AS)构成的集合, BGP(边界网关协议) [1]协调着 AS 间的有效运转, 并允许每个 AS 配置自己的路由策略以选择最合适的路由到达目的网络, 路由策略的配置通常由 AS 间的商业关系和采用的流量工程所决定。随着主机和网络数量的不断增加, 互联网规模正在快速膨胀, 使得 AS 的路由策略越来越商业化, 直接影响着 AS 间的商业互联关系, 使得 AS 间的物理连接并不意味着流量的传递。因此, AS 间的商业关系越发成为国家、企业、机构等不可或缺的网络资源, 并且 AS 商业关系在掌握网络结构与演化、设计和优化网络、发现网络故障点等方面也具有重要的指导意义。然而, 由于 AS 间商业关系由 ISP 之间协议决定, 属于商业机密, 且 ISP 在 IRR (Internet Routing Registry)注册留下的路由策略信息过于陈旧, 因此, 不能通过公开的数据中直接获得 AS 间商业关系, 又由于 ISP 间协议的变化会引起 AS 间商业关系的变化, 这种 AS 间商业关系的变动性以及互联网与生俱来的异构性、自组织性和非集中性, 都给 AS 商业关系的准确推断带来了技术挑战。

2. 相关工作

高立新等人最早提出 AS 间商业关系推断方法[2], 她假设一条有效的 AS 路径应该遵循 valley-free 原则, 即一条路径包含 0 个或多个客户 - 提供者(兄弟)边, 0 个或 1 个对等边, 以及 0 个或多个提供者 - 客户(兄弟)边。高立新的 valley-free 假设真实地反映了互联网中 AS 间的商业关系。高立新的算法试图找出一条路径中度数最大的 AS, 并认为具有相近度大小的两个 AS 是对等关系。

Subramanian 等人[3]正式地将 AS 关系类型(Tor)推断转化为一个组合优化问题——MaxToR 问题: 给定一个由 BGP 路由表产生的无向图, 并指定每条边为 p2c 或 p2p 类型(忽略 s2s 类型), 使得 valley-free 路径最大化。同时猜测 MaxToR 问题是 NP-完全问题, 并给出了一个解决算法 SARK: 利用不断去除 AS 拓扑图中叶子节点的方式为 AS 设定不同的等级, 并从多个采样点设定的 AS 等级进行比较来判断两个 AS 间的商业关系。

Di Battista、Erlebach 等人[4]证明了 MaxToR 问题确是 NP-完全问题且不能有效推断出 p2p 关系, 提出了 MaxToR 问题的解决算法 BPP 和 EHS。BPP 算法在结果的真实性上不如 SARK, EHS 将 MaxToR 问题转化为一个可满足问题, 利用随机游走的方法推断 AS 间商业关系。

Dimitropoulos 等人[5]基于 MAX-2-SAT 提出了一种解决办法。对于兄弟关系的确定是基于对 WHOIS 数据库的查寻, 然后最大化 valley-free 路径的数量和基于度差异推断的 c2p 关系, 通过解决这最大化 2 元可满足性问题推断 AS 关系。

加州大学的互联网研究实验室使用 Zhang 和 Oliveira 等人在文献[6]和文献[7]所述方法描绘了一张带有商业关系的 AS-level 拓扑图[8]。算法首先推断出 Tier-1 所包含的 AS，对于与 Tier-1 的 AS 相连的则推断为 p2c 关系，其余为 p2p 关系。Zhang 在文献[6]中给出了推断 Tier-1 中 AS 的方法，但 Oliveira 假定 Tier-1 中 AS 可以从维基百科获取到。由于将只有地区级提供商 AS 才能观察到 c2p 关系数量不断增多，导致此方法错误的将 p2c 关系推断为 p2p 关系。Gregori 等人[9]使用了类似的方法，对于每一条路径推断出可能的关系后，再基于路径的生命期确定 AS 关系。

上述从 AS 间商业关系推断的准确性来考虑，给出了不同的方法，但在互联网上由于商业利益或流量工程，甚至是路径伪造、路径篡改、路由泄露、路由震荡等攻击引起的路由变化导致 AS 间商业关系的变化很少被关注，因此，本文提出了一种基于时空可信度推断 AS 商业关系的方法。

3. AS 互联关系及路由策略

路由策略是网络管理员配置的一系列路由规则，在很大程度上反映了 AS 间的商业关系。主要包括提供者 - 客户关系(provider-to-customer, p2c)、客户 - 提供者关系(customer-to-provider, c2p)、对等关系(peer-to-peer, p2p)和同胞关系(sibling-to-sibling, s2s)。在 p2c 中，提供者可以将自己、它的客户、它的提供者或对等体的路由输出到客户，但不能输出它的提供者和对等体的路由；在 c2p 中，客户可以将自己和它的客户的路由输出到提供者；在 p2p 中，对等体间可以相互输出自己和客户的路由，但不能输出它们的提供者和其它对等体的路由；在 s2s 中，兄弟间可以将自己、客户、提供者和对等体的路由输出到对方。

根据以上路由策略高立新证明了[2]在一条 AS_PATH 中提供者 - 客户或兄弟边后面只能是提供者 - 客户边或兄弟边，对等边后面只能是提供者 - 客户边或兄弟边。若将只有 p2c 边或 s2s 边的 AS_PATH 路径序列称为下坡路径、将只有 c2p 或 s2s 边的 AS_PATH 路径序列称为上坡路径，那么，一条有效的 AS_PATH 路径应符合 6 种模式，即一条上坡路径、一条下坡路径、一条上坡路径加一条下坡路径、一条上坡路径加一个对等边、一条上坡路径加一个对等边加一条下坡路径、一条对等边加一条下坡路径。

4. AS 商业关系推断

AS 商业关系的推断首先要尽量整合路径集合，排除无效路由的影响，然后推断核心层 AS 并依据路由策略的六种模式对每一时刻每个监测点进行 AS 商业关系初步推断，统计形成时空上的 AS 商业关系监测表，计算每个 AS 商业关系的可信度值，根据阈值选取可信的 AS 商业关系。

4.1. 过滤无效路由

BGP 在设计应用之初，并未考虑到路由信息的安全性问题[10]，因此没有建立一种可靠的机制来证明路由信息的有效性，使得攻击者可以任意产生、宣告、伪造路由信息，进而对网络实施攻击。目前，从 Route-Views 上采集的单个文件的路由表项就达 2200 万条之多，而且从历史数据统计分析看，路由表项呈现逐年递增的趋势，这与网络规模的不断扩大密切相关，给路由分析带来了麻烦，也使得路由表处理的速度极为缓慢。通过对路由表项中的 AS-PATH 属性进行观察发现，许多 AS-PATH 是一样的，只是转发的源与目标路由器的 IP 地址不同，这给我们一个启发，因为主要的研究对象是 AS-PATH 属性，不能将 AS-PATH 数量缩小以达到快速处理的目标。通过对所有 AS-PATH 做集合操作，发现实际 AS-PATH 路径只有 300 多万条，其中含有核心层 AS 的路径就有 200 多万条，这大大降低了路由项处理的数量。同时还发现在 300 多万条路径中包含诸多无效路由，包括含私有 AS 或保留 AS、路由环、3 个及以上或 2 个不相邻核心层 AS、连续重复 AS 的路径。针对私有 AS 或保留 AS，AS-PATH 中不应出现

RFC1930 中定义的私有或保留的自治系统号的路径, 2015 年 10 月份的路由表中这样的路由平均占比 0.07%; 针对路由环, 在 AS-PATH 中若出现多个相同 AS 且至少有两个被不同的 AS 所分离, 这样的路径影响了正常路由的选择, 2015 年 10 月份的路由表中这样的路由平均占比 0.06%; 针对 3 个及以上或 2 个不相邻核心层 AS 的路径, 违反了路由选择策略, 2015 年 10 月份的路由表中这样的路由平均占比 0.06%; 针对连续重复 AS 的路径, 进行压缩处理, 例如将 A B B C 型路径压缩成 A B C, 2015 年 10 月份的路由表中这样的路由平均占比 19.64%。总的来看无效路由占比在 20% 左右, 这无疑将会造成 AS 的商业关系的错误判断[11] [12], 因此, 在推断 AS 间商业关系前要尽量排除无效路径的影响。

4.2. 推断核心层 AS

由于 AS 间存在着商业关系, 使得整个 BGP 网络具有一定的层次性。一般地, 按照 AS 在流量传递过程中所处位置一般将 Internet 分为核心层、转发层和边缘层。核心层 AS 一般被认为是顶级服务提供商的骨干网所形成的 Internet 的核心, 各个顶级服务提供商之间相互建立同级对等(p2p)商业互联关系, 且是全互联结构。由于在 5 万多个 AS 的拓扑中求出最大全连通子图是相对比较困难的, 为了降低计算的复杂度, 文献[13]给出一种核心层 AS 推断方法, 但此方法依赖于阈值 D 和经验值 P, 并未考虑阈值 D 外的其它 AS 是否与阈值 D 内的 AS 是全互连的, 可能存在误差。本文针对上述方法进行改进以提高核心层 AS 推断的准确性。具体方法为: 首先, 获取边缘层 AS 集合, 由于边缘层 AS 不为其它任何自治系统转发网络流量, 且位于网络的边缘, 只能出现在 AS_PATH 序列的末端且不会出现在其它位置, 因此将 AS-PATH 中末位 AS 组成的集合与非末位 AS 组成的集合做差集运算, 即得边缘层 AS 集合。其次, 指定 P 值为 1(即全连通), 并按照文献[13]所述方法得到部分核心层 AS。再次, 将 AS 集合中边缘层 AS 和部分核心层 AS 去除, 因为边缘层 AS 不能同时为核心层 AS, 且边缘层的数量占总 AS 数量的近 60%。最后, 逐个判断剩余 AS 与部分核心层是否为全连通, 如果为全连通则加入核心层。

以 2015 年 10 月 1 日至 30 日的路由数据进行分析, 按照上述方法提取核心层 AS 共 13 个, 分别 2914、7018、2828、174、209、1299、4436、6453、1239、3320、3257、3356、701, 且它们在 1 日至 30 日均被推断为核心层 AS, 说明核心层 AS 是相对稳定的。

4.3. 基于时空的 AS 商业关系可信度计算

由于网络的动态变化, 诸如 ISP 间的合作关系的变化、针对 BGP 路由的路径伪造、路径篡改、路由泄露等等都会影响 AS 间的商业关系, 需要一种有效的方法来辨别 AS 间商业关系的可靠性。为此, 本文提出一种基于时空的 AS 商业关系可信度的计算方法, 在基于时间的统计分析上加以空间上的统计分析, 即在某一时刻将路由表按监测点进行分离后单独对 AS 间商业关系进行推断与监测, 形成空间与时间的立体式统计分析计算。其具体计算方法如公式(1)和(2)。

$$Z_i = \begin{cases} \sum_1^{n1} V_{MPi} / m, & \sum_1^{n1} V_{MPi} < m \\ 1.0, & \sum_1^{n1} V_{MPi} \geq m \end{cases} \quad (1)$$

$$S = \sum_1^{n2} Z_i * 2^{\frac{i}{3}} * i / \sum_1^{n2} 2^{\frac{i}{3}} * i \quad (2)$$

空间可信度由公式(1)给出, 其中, $n1$ 为监测点的数量, V_{MPi} 的值为 0 或 1, $V_{MPi} = 1$ 表示在监测点 MPi 推断出了 AS 商业关系 $\langle ASN_x, ASN_y, R \rangle$, 反之, $V_{MPi} = 0$ 表示在监测点 MPi 没有推断出 AS 商业关系 $\langle ASN_x, ASN_y, R \rangle$, 考虑到监测点的数量较多, 这里指定在有 m 个及以上监测点监测到了某一 AS 间商业关系对的存在, 即认为此 AS 商业关系对在空间上是真实可信的。

基于时空的 AS 商业关系可信度 S 由公式(2)给出, 主要利用对空间可信度加权平均来实现, 使得在最近频繁出现的 AS 商业关系的权重变大, 说明在时间维度上越稳定可信。给定 AS 商业关系可信度阈值 β , 若 $S > \beta$, 则认为对应的 AS 商业关系在时空上是可信的, 反之亦然。

4.4. 基于可信度的 AS 商业关系推断算法

依据路由策略模式, 基于时空可信度的 AS 间商业关系算法将监测点的数量及路由表采集的时间综合考虑进去进行 AS 商业关系判断。首先, 获取单个监测点的路由表并进行 AS 商业关系初步推断, 具体算法如算法 1, 其基本过程如下:

(1) 去除 as-path 中的含私有 AS 或保留 AS、路由环、3 个及以上或 2 个不相邻核心层 AS、含环、即在 as-path 中出现至少 2 个相同 AS 的路径, 压缩含连续重复 AS 的 as-path。

(2) 得到核心层 AS 集合, 若两个不同 AS 都是核心层 AS, 则推断它们为对等关系。

(3) 若 AS-PATH 中含有核心层 AS, 则推断第一个核心层 AS 及其左侧为上坡路径, 且按路径顺序从左至右每相邻两个 AS 的商业关系为客户 - 提供者关系; 最后一个核心层 AS 及其右侧为下坡路径, 且按路径顺序从左至右每相邻两个 AS 的商业关系为提供者 - 客户关系。

(4) 若 AS-PATH 中不含核心层 AS, 且已知两个不相邻 AS 的商业关系为提供者-客户(或客户-提供者)关系, 则推断二者之间 AS 的商业关系为提供者 - 客户(或客户 - 提供者)关系。

(5) 若在不合核心层 AS 的 AS-PATH 中存在提供者 - 客户关系, 且其右侧 AS 关系还没有被标记, 则推断其右侧为提供者 - 客户关系。

(6) 若在不合核心层 AS 的 AS-PATH 中存在客户 - 提供者关系, 且其左侧 AS 关系还没有被标记, 则推断其左侧为客户 - 提供者关系。

(7) 重复步骤(3)、(4)和(5), 直到没有新的 AS 商业关系产生。

(8) 遍历上述产生的 AS 商业关系集, 若存在两个 AS1 与 AS2 的关系既为提供者 - 客户关系也为客户 - 提供者关系, 则推断其为兄弟关系。

然后, 在单个监测点 AS 商业关系初步推断的基础上, 进行空间和时间上的统计, 并计算每一对 AS 商业的可信度值 S , 若 S 大于给定的 β , 则认为此 AS 商业关系在时空上是稳定的、可信的。具体算法如算法 2, 其基本过程如下:

(1) 将路由表按监测点进行分离, 对于每个监测点的路由表进行 AS 商业关系初步推断, 得到空间上 AS 商业关系的统计表如表 1。

(2) 按时间获取路由表, 针对每张路由表执行(1)操作, 得到时空上监测到的 AS 商业关系统计表如表 2。

(3) 按照公式(1)和(2)计算每个 AS 商业关系的可信度值, 舍弃 S 值小于 β 的 AS 商业关系。

5. 实验结果及分析

本实验路由数据来源于开源公共项目 Route-Views [14], 因其路由数据更新的连续性与完整性, 且单个路由表快照包含多个地区 AS 监测点。以 2015 年 10 月 1 日路由数据为例共观察到 41 个监测点, 这 41 个监测点遍布世界各地, 可以较为全面地反映互联网的路由情况, 是实验基础数据的最佳选择。Route-Views 上提供两种格式[15]的路由表供下载, 一种是从 Cisco 路由器以“sh ip bgp”命令产生的路由表数据的形式提供, 每 2 个小时提供一次路由数据镜像, 但这个过程相对比较慢, 而且路由的时间戳可能会受到影响, 由于路由器的缓存以及本地网络的维护问题, 可能导致路由数据部分丢失; 另一种是从软件路 Zebra 上采集到的路由数据, 数据以 MRT 格式存储, 这种格式方便被计算机处理, 路由数据每

Algorithm 1. AS relationship inference of single monitored point**算法 1.** 单监测点 AS 商业关系推断输入：单监测点路径集合 AS_PATH_SET 输出：AS 对的关系集 $Relationship_SET$ ，在关系集中每对 AS 商业关系用三元组 $\langle ASN_x, ASN_y, R \rangle$ 表示，其中 ASN_x, ASN_y 是 AS_PATH_SET 中出现的任意两个互异的自治系统号， R 是 ASN_x 与 ASN_y 的商业关系 ($R \in \{p2p, p2c, c2p, s2s\}$)

- 1) 设 P 为任意一条路径，记为 $P = a_1 a_2 a_3 \dots a_i \dots a_n (1 \leq i \leq n)$
- 2) for ASN_x, ASN_y in $CoreAS_SET$
- 3) if $ASN_x, ASN_y \in CoreAS_SET$ and $ASN_x \neq ASN_y$
- 4) $Relationship_SET.add(\langle ASN_x, ASN_y, p2p \rangle)$
- 5) 从 AS_PATH_SET 中提取出含有核心层 AS 的路径，构成集合 $Core_AS_PATH_SET$ ，其余路径构成集合 $None_Core_AS_PATH_SET$
- 6) for $path$ in $Core_AS_PATH_SET$
- 7) if $path$ 中只有一个核心层 AS 且序列号为 m
- 8) for i in $rang(0, m)$
- 9) $Relationship_SET.add(\langle ASN_i, ASN_{i+1}, c2p \rangle)$
- 10) for j in $rang(m, len(path))$
- 11) $Relationship_SET.add(\langle ASN_j, ASN_{j+1}, p2c \rangle)$
- 12) elif $path$ 中只有两个核心层 AS 且序列号为 m, n
- 13) for i in $rang(0, m)$
- 14) $Relationship_SET.add(\langle ASN_i, ASN_{i+1}, c2p \rangle)$
- 15) for j in $rang(n, len(path))$
- 16) $Relationship_SET.add(\langle ASN_j, ASN_{j+1}, p2c \rangle)$
- 17) while 有新的 AS 商业关系产生
- 18) for $path$ in $None_Core_AS_PATH_SET$
- 19) if $path$ 中 $\langle ASN_i, ASN_{i+1} \rangle$ 和 $\langle ASN_j, ASN_{j+1} \rangle (i < j)$ 为 *customer-provider* 关系
- 20) for k in $rang(i, j)$
- 21) $Relationship_SET.add(\langle ASN_k, ASN_{k+1}, c2p \rangle)$
- 22) if $path$ 中 $\langle ASN_i, ASN_{i+1} \rangle$ 和 $\langle ASN_j, ASN_{j+1} \rangle (i < j)$ 为 *provider-customer* 关系
- 23) for k in $rang(i, j)$
- 24) $Relationship_SET.add(\langle ASN_k, ASN_{k+1}, p2c \rangle)$
- 25) if $path$ 中存在 *customer-provider(provider-customer)* 关系，且其左(右)侧 AS 关系未确定
- 26) 推断其左(右)侧 AS 为 *customer-provider(provider-customer)* 关系
- 27) for x in $Relationship_SET$
- 28) if $\langle ASN_x, ASN_y, p2c \rangle$ 和 $\langle ASN_y, ASN_x, c2p \rangle$
- 29) $Relationship_SET.remove(\langle ASN_x, ASN_y, p2c \rangle)$
- 30) $Relationship_SET.remove(\langle ASN_y, ASN_x, c2p \rangle)$
- 31) return $Relationship_SET$

Algorithm 2. AS business relationship inference based on temporal and spatial reliability**算法 2.** 基于时空可信度的 AS 商业关系推断输入：时间 T 内多监测点的路由表输出：AS 商业关系集 $Relationship_SET$

- 1) for t in T
- 2) 选取 t 时刻的路由表
- 3) 得到有效路径集 AS_PATH_SET
- 4) 按监测点进行路由分离，生成包含每个监测点路径集合的列表 $mpASPathSetList$
- 5) for $mpASPathSet$ in $mpASPathSetList$
- 6) 按照算法 1 进行 AS 商业关系初步推断
- 7) 得到空间上 AS 商业关系统计表
- 8) 得到时空上 AS 商业关系统计表
- 9) 按照公式(1)和(2)计算 AS 间商业关系的可信度，并去除 S 值小于 β 的 AS 商业关系
- 10) return $Relationship_SET$

Table 1. AS relationship monitoring statistics in space**表 1.** 空间上 AS 商业关系监测统计表

AS 商业关系对	监测点	MP1	MP2	MPi	MPn1
$\langle ASN_x, ASN_y, R \rangle$		V_{MP1}	V_{MP2}	V_{MPi}	V_{MPn1}
.....	

15 分钟更新一次, 相比第一种路由信息更加丰富、准确, 因此实验数据选用 MRT 格式。采集 Route-Views 上 2015 年 10 月 1 日至 30 日每日零点的路由表快照作为实验数据进行 AS 商业关系推断, 并统计 m , β 取不同值时产生的结果, 如图 1 所示。

从图 1 可以看出, 当 m 值逐渐增大时推断的 AS 商业关系数量呈现递减的趋势, 且随着 β 取值的增大, 这种递减的趋势也不断增大。这是因为随着 m , β 取值的增大, 对于基于时空可信度的 AS 商业关系推断越趋严格, 即可信度值越大, 符合可信度值的 AS 商业关系数量就越少, 这在一定程度上可以减少由于商业利益、流量工程、路径攻击等对 AS 商业关系引起的变化的影响, 防止非正常 AS 商业关系被推断出来而掩盖非法攻击行为。

为了评估本算法推断的 AS 商业关系的有效性并分析 m 、 β 值的选取对于有效性的影响, 本实验产生的结果将与 10 月 1 日 CAIDA 提供的 AS 商业关系进行核实验证。不同 m 、 β 值对应的有效 AS 商业关系占比如图 2 所示。

由图 2 可以看出 m 、 β 值的递增对有效 AS 商业关系占比的影响不是线性变化的, 在 m 值小于 10 时的曲线中, 都有陡增的部分, 说明某些 AS 商业关系只会有限的监测点被观察到, 这样的变化有可能是商业利益、流量工程、路径攻击等引起的, 需要其他佐证才有进一步确定。此外, 当 $m = 1$, $\beta \geq 0.0$ 时, 即 41 个监测点中只要有一个监测点观察到 AS 商业关系且可信度为 1 时所推断出来的 AS 商业关系, 从图 1 和图 2 中可以看出, 虽然此时推断出来的 AS 商业关系数量最多, 但其可信 AS 商业关系占比却是最低的, 说明其中有很多不稳定或错误的 AS 商业关系; 当 $m = 14$, $\beta \geq 0.9$ 时, 有效 AS 商业关系占比值最大, 由此可确定 m 、 β 值可以最大化 AS 商业关系推断的有效性。

Table 2. AS relationship monitoring statistics in time

表 2. 时间上 AS 商业关系监测统计表

AS 商业关系对	监测点	Time_1	Time_2	Time_i	Time_n2
		<ASN _x , ASN _y , R>	Z ₁	Z ₂	Z _i
.....

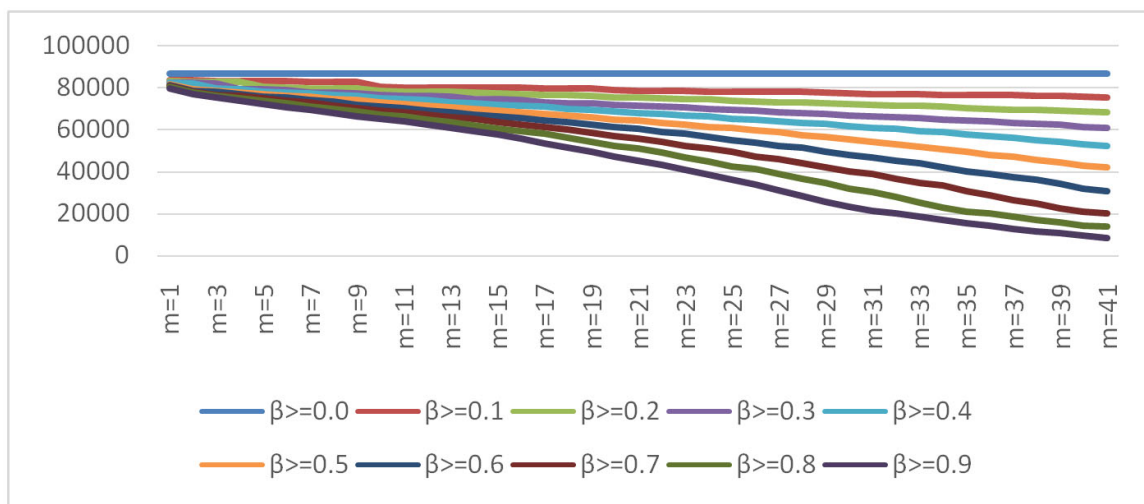


Figure 1. m , β value of impact on the number of AS relationship

图 1. m 、 β 值对 AS 商业关系数量的影响

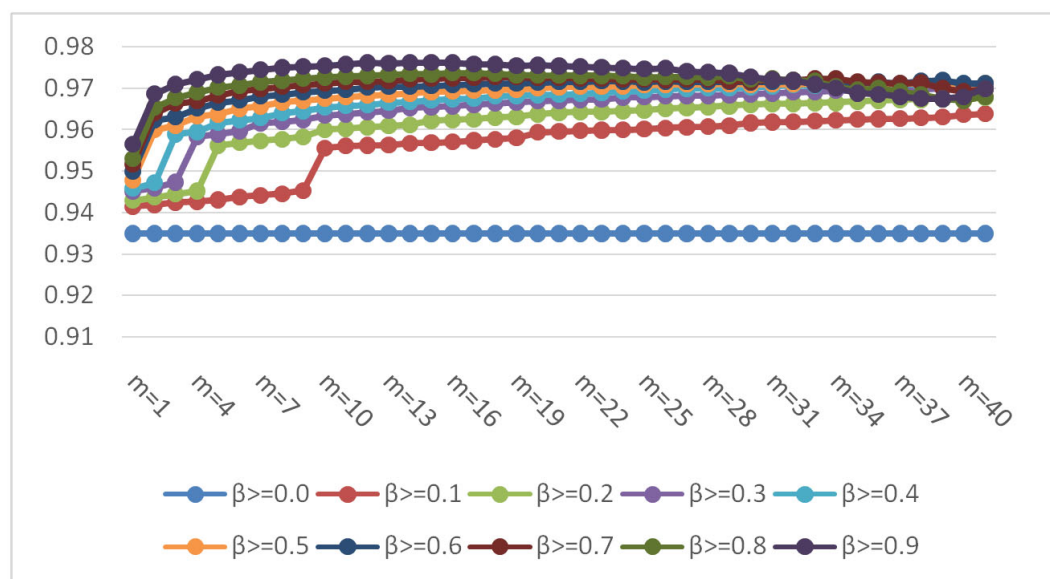


Figure 2. m , β value of impact on the ratio of effective AS relationship

图 2. m 、 β 值对应的有效 AS 商业关系占比的影响

6. 结束语

本文通过对路由表中的无效路由进行简单分析与统计, 说明无效路由在网络分析与 AS 商业关系推断中具有一定的影响, 应在 AS 商业关系推断前尽量排除或重整无效路由。同时本文给出一种基于时空可信度的 AS 商业关系推断算法, 并分析了 m 、 β 取值对 AS 商业关系推断数量及有效 AS 商业关系占比的影响。实验结果表明, 选取适当的 m 、 β 值可以有效提高 AS 商业关系推断的精准度。

基金项目

新一代互联网社会网络感知模型与能力设计自然科学基金(编号: 61170285); 面向国家关键基础设施的大规模人机物融合网络安全可控性模型与机制自然科学基金(编号: 61572514)。

参考文献 (References)

- [1] Chen, E. (2004) BGP Support for Four-Octet AS Number Space. IETF Internet-Draft-4893, December 2004.
- [2] Gao, L. (2001) On Inferring Autonomous System Relationships in the Internet. *IEEE/ACM Transactions on Networking*, **9**, 733-745. <http://dx.doi.org/10.1109/90.974527>
- [3] Subramanian, L., Agarwal, S., Rexford, J. and Katz, R.H. (2002) Characterizing the Internet Hierarchy from Multiple Vantage Points. *Proceedings of IEEE 21st Annual Joint Conference of the IEEE Computer and Communications Societies*, **2**, 618-627. <http://dx.doi.org/10.1109/infcom.2002.1019307>
- [4] Di Battista, G., Erlebach, T., Hall, A., et al. (2007) Computing the Types of the Relationships between Autonomous Systems. *IEEE/ACM Transactions on Networking*, **15**, 267-280. <http://dx.doi.org/10.1109/TNET.2007.892878>
- [5] Dimitropoulos, X., Krioukov, D., Fomenkov, M., et al. (2007) AS Relationships: Inference and Validation. *ACM SIGCOMM Computer Communication Review*, **37**, 29-40. <http://dx.doi.org/10.1145/1198255.1198259>
- [6] Zhang, B., Liu, R., Massey, D., et al. (2005) Collecting the Internet AS-Level Topology. *ACM SIGCOMM Computer Communication Review*, **35**, 53-61. <http://dx.doi.org/10.1145/1052812.1052825>
- [7] Oliveria, R., Pei, D., Willinger, W., Zhang, B. and Zhang, L. (2008) Quantifying the Completeness of the Observed Internet AS-Level Structure. Technical Report TR-080026-2008, UCLA CS Dept.
- [8] Bron, C. and Kerbosch, J. (1973) Finding All Cliques of an Undirected Graph. *Communications of the ACM*, **16**, 575-576. <http://dx.doi.org/10.1145/362342.362367>

- [9] Gregori, E., Improta, A., Lenzini, L., *et al.* (2011) BGP and Inter-AS Economic Relationships. NETWORKING 2011. Springer Berlin Heidelberg, 54-67.
- [10] Rekhter, Y. and Li, T. (1994) A Border Gateway Protocol 4 (BGP-4). RFC. <http://dx.doi.org/10.17487/rfc1654>
- [11] 王洪君, 于晓鹏. 一种 BGP 无效路由检测方法[J]. 吉林师范大学学报: 自然科学版, 2008(3) :54-57.
- [12] Luckie, M. (2014) Spurious Routes in Public BGP Data. *ACM SIGCOMM Computer Communication Review*, **44**, 14-21. <http://dx.doi.org/10.1145/2656877.2656880>
- [13] 邓文平, 郭敏, 胡晓峰, 等. 互联网 AS 拓扑的结构与连通性研究[J]. 计算机工程与科学, 2012, 34(6): 1-6.
- [14] Mayer, D. (2015) University of Oregon Route Views Project. <http://www.routeviews.org>
- [15] Route Views Project (2015) Route-Views Data. <http://www.routeviews.org/data.html>