文章编号:1001 - 2486(2006)03 - 0071 - 06

基于规则的域间路由系统异常检测*

刘 欣[1],朱培栋[1],米 强[2],杨明军[1]
(1. 国防科技大学 计算机学院,湖南 长沙 410073; 2. 国家计算机网络与信息安全管理中心,北京 100029)

摘 要:随着 Internet 的爆炸性增长,域间路由系统变得越来越复杂并难以控制,许多与域间路由安全相关的事件广泛引起了人们的关注。提出一个基于规则的域间路由监测框架,其中的规则可分为常规异常检测规则和特殊异常检测规则,它们能有效用于检测异常路由和可能的攻击行为,这两种规则的不同在于特殊异常检测规则是由大量路由得到的 Internet 模型来定义。研究了 Internet 层次模型与 ISP 商业关系模型的构造,基于这个框架实现了一个原型系统——ISP-Health,最后给出了检测能力结果。

关键词:域间路由;异常路由;路由攻击;检测规则
中图分类号:TP393.08 文献标识码:A

A Rule-based Approach to Anomaly Detection in Inter-domain Routing System

LIU Xin[1], ZHU Pei-dong[1], MI Qiang[2], YANG Ming-jun[1]
(1. College of Computer, National Univ. of Defense Technology, Changsha 410073, China;
2. National Network and Information Security Administration Center, Beijing 100029, China)

Abstract: The behaviors of the Inter-domain Routing (IDR) System are becoming rather complicated with the rapid development of the Internet. Security incidents in IDR system have attracted extensive attention among people. This paper proposes a rule-based monitoring framework to secure IDR System, in which the rules can be used to effectively detect anomalous routes and possible attacks. Unlike GADRs, SADRs were defined according to some Internet models that are behavior-models represented by large numbers of normal routes. Furthermore the construction of the Internet Hierarchy Model and ISP Commercial Relationships Model were studied, and methods based on these models were developed to detect hidden route anomalies or attacks. ISP-Health, the prototype of such a monitoring system supported by the above-mentioned framework, was implemented, and its capabilities were exhibited at last.

Key words: inter-domain routing; anomalous routes; routing attacks; detection rule

Internet 由成千上万的网络互联而成,其中边界网关协议(BGP)起到至关重要的作用。尽管从用户的角度 BGP 不为人知,但作为 Internet 的关键基础设施,BGP 的稳定、正常与否对整个 Internet 有着重要影响[1]。然而,BGP 存在许多安全缺陷[2-3]并受到多种恶意攻击的威胁,如病毒的传播等[4];更为严重的是,单台 BGP 路由器的错误配置或恶意行为能够影响其它正常的 BGP 路由器[5-6]。

由 BGP 路由而引起的全球 Internet 不稳定事件经常发生,轻则降低服务质量,重则影响网络连通性。虽然路由器提供商不断地提高路由器的路由表查找速度、容量以及包交换速度来适应 Internet 的发展,但它们在路由管理工具方面所做的努力却很少。为了观测 BGP 路由情况、寻求解决 BGP 路由问题的途径,使得整个 Internet 健康发展,迫切需要有效的 BGP 监测及异常检测工具。作为一种确保域间路由系统安全的方法,域间路由监测能被用于识别错误配置、路由不稳定、IP 地址劫持、策略违背以及伪造 AS 等问题。

传统监测系统,如基于 SNMP 的监测系统,只适用于单个 ISP 所管理的自治系统内部,对其自治系统域外的监测无能为力,因而已有监测系统能力有限,不能解决超出网络管理边界的问题。为解决现有

* 收稿日期:2005 - 12 - 01
基金项目:国家 863 高技术发展计划资助项目(2005AA121570);现代通信国家重点实验室基金资助项目(51436050605KG0102)
作者简介:刘欣(1978—),男,博士生。

监测系统不足,本文提出一个基于规则的监测框架监测整个域间路由系统,研究了两种 Internet 模型及其相关的异常路由检测规则。

1 基于规则的监测框架

本节主要讨论一个基于规则的域间路由系统监测框架,如图 1 所示。其基本思想是:若一条路由匹配了框架中的任意一条异常规则,则该路由就被认为是异常的。整个框架由多个采集器、一个规则引擎以及异常检测模块组成。这个框架不但能够对采集的路由数据进行分析,而且还能生成 Internet 模型和检测异常路由。优点在于检测能力和可扩展性,即规则引擎中的规则越多,检测能力越强。

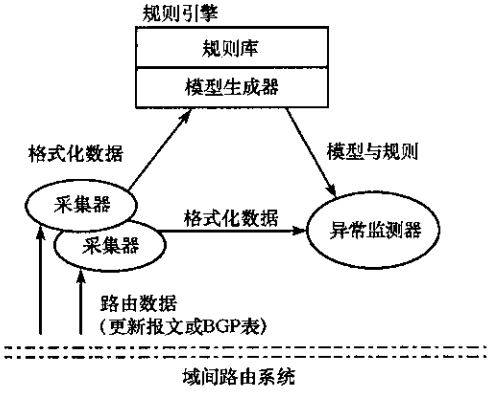

图 1 基于规则的域间路由系统监测框架

Fig.1 The rule-based framework for monitoring IDR

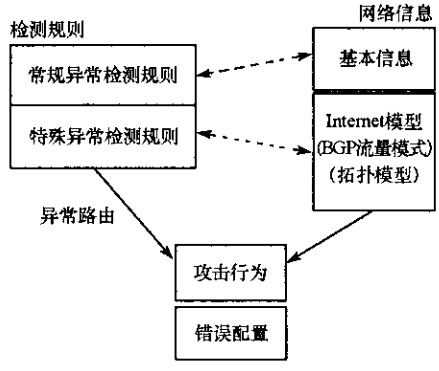

图 2 规则与网络基本信息的关系

Fig.2 Relations of the rules and the network information

采集器从监测点得到原始数据(BGP 更新报文或 BGP 表数据),然后把这些数据整理为统一的格式以便于其它模块的处理。为了生成路由失效警报并定位问题来源,需要对多监测点进行监测。由图 1 可以看到,整个框架的核心就是规则引擎,它由一个规则库和一个模型生成器组成,模型生成器负责构造 Internet 的各种模型(Internet 模型被认为是大量 BGP 路由所表现出来的特性或正常行为)。规则库中的规则被分为两大类:常规异常检测规则和特殊异常检测规则。这些规则将在后面的内容中详细讨论。异常检测器处理由采集器送来的路由数据,使用规则引擎中的模型与规则来评估路由数据,并判断是否受到攻击。

在图 2 中展示了框架中规则与网络基本信息的关系。网络信息由基本网络信息和网络的模型信息两部分组成。根据不同的网络模型,配合基本网络信息就可以定义不同的异常检测规则。基本网络信息包括已分配的自治系统号与 IP 前缀的信息以及它们之间的映射关系。网络信息中的 Internet 模型有 ISP 的商业关系、Internet 层次模型、地理关系模型、小世界特性和密率特性等。

2 常规异常检测规则

通常,一条路由如果不遵守常规异常检测规则,那么它不允许在 Internet 上传播。与特殊异常检测规则相比,常规异常检测规则相对明显并已被用于检测错误路由,如含有私有前缀或 AS 号的路由。许多常规异常检测规则可直接从 RFC 文档中推导出来[9]。以往,人们并没有把这些规则加以提炼并给出定义,本文总结了 10 条这种路由规则。尽可能多地定义这些路由规则有许多作用,如可把这些规则用于对路由数据进行常规检测,消除在推断路由策略时问题路由的干扰等等。

规则 a1:被监测网络的 BGP 路由表项中包含的 IP 前缀,若属于 RFC1918 中定义的私有地址块,则判定该路由为‘含私有地址’异常;

规则 a2:被监测网络的 BGP 路由表项中包含的 IP 前缀,若不属于基本信息库中定义的已分配 IP 地址块,则判定该路由为‘含未分配地址’异常;

规则 a3:被监测网络 BGP 路由表项中包含的 IP 前缀,若不符合基本信息库定义的 AS-IP 地址映射关系,则判定该路由为‘未授权使用地址块’异常,可能存在伪造(或劫持)路由攻击;

规则 b1:被监测网络的 BGP 路由表项中 AS-PATH 包含的自治系统号,属于 RFC1930 中定义的私有自治系统号,则判定该路由为"含私有 AS 号"异常;

规则 b2:被监测网络 BGP 路由表项中 AS-PATH 包含的自治系统号,不属于基本信息库中定义的已分配自治系统号,则判定该路由为"含未分配 AS 号"异常;

规则 b3:被监测网络的 BGP 路由表项中 AS-PATH 包含的自治系统号,出现重复且重复的自治系统号在 AS-PATH 中不连续,则判定该路由为"含 AS 环"异常;

规则 c1:若同一地址前缀有多个发起者,则判定该路由有潜在的 MOAS(多来源自治系统)冲突;

规则 c2:具有潜在 MOAS 冲突的网络前缀,若基本信息库中的自治系统号与 IP 地址块的映射信息表明,发生冲突的自治系统之间没有存在隶属关系或未对该网络前缀进行授权,则判定该路由为路由劫持异常;

规则 c3:具有潜在 MOAS 冲突的网络前缀,若冲突的两个自治系统都部署了监测点(例如监测点 A 和 B)并已经获得该监测点的 BGP 路由表数据,监测点 A 的 BGP 路由表指示监测点 B 为发起者,但该前缀在监测点 B 路由表中不是源发路由,则判定发生针对监测点 B 所在网络的伪造异常;

规则 d:若某个监测点(例如监测点 A)的路由表指示某路由的发起者或路由上一跳为另一个监测点(例如监测点 B),但是监测点 B 并未发起该路由或传播该路由,则判定发生针对监测点 B 所在网络的路由伪造异常。

3　特殊异常检测规则

与常规异常检测规则不同,特殊异常检测规则不能由 RFC 文档直接推导出来,它们通常来自于各个 ISP 的运营经验[7,9-10]以及大量 BGP 路由表现出来的行为模式。大量的 BGP 路由应该表现出某些共同的特征,本文把这种共同的特性或行为称为 Internet 模型。由此,利用特殊异常检测规则来进行检测异常路由的基本思想就是找出那些与某一模型不匹配的路由,也就是说这样的路由存在异常,我们把这样的路由称为异常路由。显然,要很好地利用这个思想,如何构造 Internet 模型以及定义相关的异常检测规则就显得非常重要。在本小节,主要讨论两个 Internet 模型及相关的异常检测规则,它们被用于 ISP-Health 系统中。

3.1　Internet 层次关系模型及相关的异常检测规则

3.1.1　Internet 层次模型

整个 Internet 在自治系统级别可以划分为三个层次:核心层、转发层和边缘层[8]。下面给出了对这三个层次的讨论以及得到的推断算法。

方法 A:一般认为,顶级服务提供商的骨干网形成了 Internet 的核心,称为 DFZ 区域(default-free Zone)。为获得整个 Internet 的连通性,各顶级服务提供商之间相互建立同级对等(peer-peer)商业互连关系。因此,推断 Internet 核心层这个问题可定义为:对于 Internet 的自治系统拓扑图 G,求图 G 中的最大全互连集。显然,这是个 NP 难问题。本文利用 Internet 拓扑中的自治系统的度信息,采用启发式方法推断核心层的组成与结构,具体过程如下:

输入:全部 AS-PATH 集(把全部 AS-PATH 集看成图 G)

输出:核心层自治系统集 Tier1 _ AS _ SET

(1) Tier1 _ AS _ SET←Ø;

(2) 计算图 G 中每个结点 v 的度,并把结果存放在一张信息表中;

(3) 得到图 G 的最大度结点集;

(4) $\max\_degree\_nodes(G) = \{v \mid d(v) = \max(d(v_1), d(v_2), \cdots),\ v_1, v_2, \cdots \in V\}$;

(5) 若 $|\max\_degree\_nodes(G)| = 1$,设 z 为 $\max\_degree\_nodes(G)$ 的唯一元素;

(6) 若 $|\max\_degree\_nodes(G)| \neq 1$,那么查看信息表以选出一个元素 z,其中 $z \in \max\_degree\_nodes(G)$,且 z 在信息表中历史记录的度不比其它元素小;

(7) Tier _ AS _ SET←Tier1 _ AS _ SET∪ {z}；

(8) Neighbor _ set←得到图 G 中结点 z 的邻居集；

(9) 从图 G 中得到结点集为 Neighbor _ set 的导出子图 G'；

(10) $G \leftarrow G'$；

(11) 若图 G 满足条件 $|E(G)| \geqslant \alpha \cdot^{1/2} \cdot (|V(G)| - 1) \cdot |V(G)|)$，则退出；否则，返回到第(2)步(其中，$|E(G)|$ 是 G 中的边数，$|V(G)|$ 是 G 中的结点数，α 是用来控制 Tier1 集合中连接稀疏程度的系数，若 $\alpha = 1$ 则为全连接图)。

方法 B 构造边缘层的方法是：若一个自治系统不为其它任何自治系统转发网络流量，则它称为边缘自治系统，它位于边缘层(最底层)。某自治系统若是边缘自治系统，则它在 AS-PATH 集中只会出现在 AS-PATH 的尾部。因此，对于某个自治系统，通过扫描所有 AS-PATH 就可以判断是否属于边缘自治系统集；利用每个自治系统的判别结果，就可得到边缘自治系统集，具体构造方法是：

输入：全部 AS-PATH 集

输出：边缘自治系统集 STUB _ AS _ SET

(1) STUB _ AS _ SET←Ø；

(2) 得到自治系统列表 AS _ LIST；

(3) 对于 AS _ LIST 表中的每个自治系统 v，重复(4) ~ (6)步；

(4) Flag = 0；

(5) 检查所有 AS-PATH 集，若 v 不在 AS-PATH 的尾部，则 Flag = 1；

(6) 若 Flag = 0，则把 v 加入边缘自治系统集 STUB _ AS _ SET。

方法 C 构造转发层的方法是：通过方法 A 识别出核心层和方法 B 识别出边缘层后，剩下的自治系统都归为转发层，具体方法是：

输入：全部 AS-PATH 集

输出：转发层自治系统集 TRANSIT _ AS _ SET

(1) 获得核心层自治系统集 Tier1 _ AS _ SET(利用方法 A)；

(2) 获得边缘自治系统集 STUB _ AS _ SET(利用方法 B)；

(3) 得到 Internet 中所有自治系统集 AS _ SET；

(4) TRANSIT _ AS _ SET←AS _ SET－Tier1 _ AS _ SET－STUB _ AS _ SET。

3.1.2 特殊异常检测规则

根据学到的 Internet 层次关系模型检测异常。在正常情况下，一条满足 Internet 层次特性的路径应该是先从低一层次爬升到高一层次，然后，从高一层次降到低一层次；若一条路径通过核心层，由于核心层中的自治系统是全互连关系，该路径只通过一跳就可穿过核心层。因此，只要满足下面规则之一便违背了层次特性，称为违背 Internet 层次关系模型的异常判定规则：

规则 s1：若一条路由从高一层次降到低一层次后又回到高一层次，则这条路由为异常路由；

规则 s2：若一条路由通过核心层用了两跳以上，则这条路由为异常路由。

3.2 ISP 商业关系模型及相关的异常检测规则

3.2.1 ISP 商业关系模型

一般认为，Internet 的自治系统之间存在三种基本的商业互连关系：提供商－客户关系(provider-customer)，客户－提供商关系(customer-provider)，同级对等关系(peer-peer)等[7]。下面给出了我们用于推断这个 ISP 商业关系模型的算法。方法 D：

输入：全部 AS-PATH 集，其中的路径 p 由自治系统序列组成，记为 $p = \alpha_1 \alpha_2 \cdots \alpha_i \cdots \alpha_n (1 \leqslant i \leqslant n)$

输出：自治系统对 $< \alpha, \beta >$ 的关系集 Relation _ SET，其中 α, β 是 AS-PATH 中出现的任意自治系统号

(1) 利用方法 A 得到核心层 AS 集 Tier1 _ AS _ SET，若 $\alpha、\beta \in$ Tier1 _ AS _ SET，则 $< \alpha, \beta >$ 记为 peer-peer 关系；

(2) 从 AS-PATH 集提取出含有 Tier1 _ AS _ SET 集合中元素的路径，这些路径构成集合 Core _ AS-

PATH;

(3) 若 $p \in$ Core _ AS-PATH,设 $\alpha_i \in$ Tier1 _ AS _ SET

(3 – 1) 把 p 中 α_i 左侧的所有 AS 对 $< \alpha_{j-1}, \alpha_j > (j \leqslant i)$ 记为 customer-provider 关系;

(3 – 2) 把 p 中 α_i 右侧的所有 AS 对 $< \alpha_j, \alpha_{j+1} > (j \geqslant i)$ 记为 provider-customer 关系;

(4) 若 $p \in$ AS-PATH-Core _ AS-PATH

(4 – 1) 若 p 中 AS 对 $< \alpha_{i-1}, \alpha_i >$ 和 $< \alpha_j, \alpha_{j+1} > (i < j)$ 为 customer-provider 关系,则把它们中间的所有 AS 对 $< \alpha_r, \alpha_{r+1} > (i \leqslant r < j)$ 记为 customer-provider 关系;

(4 – 2) 若 p 中 AS 对 $< \alpha_{i-1}, \alpha_i >$ 和 $< \alpha_j, \alpha_{j+1} > (i < j)$ 为 provider-customer 关系,则把它们中间的所有 AS 对 $< \alpha_r, \alpha_{r+1} > (i \leqslant r < j)$ 记为 provider-customer 关系;

(4 – 3) 重复(4 – 1)和(4 – 2)直到没有发现新的 customer-provider 或 provider-customer 关系对;

(5) 若 $p \in$ AS-PATH-Core _ AS-PATH;

(5 – 1) 若 p 中 AS 对 $< \alpha_{i-1}, \alpha_i >$ 为 customer-provider 关系和 $< \alpha_j, \alpha_{j+1} >$ 为 provider-customer 关系 $(i < j)$,则把它们中间的所有 AS 对 $< \alpha_r, \alpha_{r+1} > (i \leqslant r < j)$ 记为 peer-peer 关系;

(5 – 2) 若 p 中 α_i 右侧的所有 AS 对 $< \alpha_j, \alpha_{j+1} > (j \geqslant i)$ 为 provider-customer 关系且 α_i 左侧的所有 AS 对 $< \alpha_{k-1}, \alpha_k > (k \leqslant i)$ 还没标记关系,则把 $< \alpha_{k-1}, \alpha_k > (k \leqslant i)$ 都记为 peer-peer 关系;

(5 – 3) 若 p 中 α_i 左侧的所有 AS 对 $< \alpha_{j-1}, \alpha_j > (j \leqslant i)$ 为 customer-provider 关系且 α_i 右侧的所有 *AS* 对 $< \alpha_k, \alpha_{k+1} > (k \geqslant i)$ 还没标记关系,则把 $< \alpha_{k-1}, \alpha_k > (k \geqslant i)$ 都记为 peer-peer 关系。

3.2.2　特殊异常检测规则

Internet 的连通性并不等价于可达性。例如,在实践中通常一个 ISP C 有两个提供商 A 与 B,尽管 A 通过 C 与 B 相连,但 B 与 C 的流量不能通过 A 相互。若存在这样的路由使得 B 可通过 A 到达 C,这种路由显然违背了 A、B 与 C 之间的商业关系[10]。

因此,根据上小节的 ISP 商业互连关系推断算法可定义下面四条特殊异常检测规则。若一条路由满足下面规则之一便违背了 ISP 商业互连关系约束,该路由就是异常路由。

规则 s4:若一条路由在通过提供商到客户的正向边后又通过一条对等边,则该路由为异常路由;

规则 s5:若一条路由在通过提供商到客户的正向边后又通过一条客户到提供商的逆向边,则该路由为异常路由;

规则 s6:若一条路由在通过一条对等边后又通过一条客户到提供商的逆向边,则其为异常路由;

规则 s7:若一条路由在通过一条对等边后又通过一条对等边,则该路由为异常路由。

4　ISP-Health:一个基于规则的监测系统

利用前面的异常检测技术,我们实现了一个监测系统原型——ISP-Health。该系统根据已有的网络基本信息(如 AS、IP 块的分配信息等)、定义的规则以及上面的模型构造算法,检测网络的异常行为和可能的路由攻击。在图 3 中展示了 ISP-Health 的系统结构。

检测异常的过程主要包括:利用常规异常检测规则对采集到的路由数据进行一般性的检测和利用特殊异常检测规则进一步检测两个阶段。特别地,在整个构造模型与异常检测期间,ISP-Health 都要对取自多个被监测点的路由信息进行处理。如图 3 所示,当路由数据从一个监测点采集后,ISP-Health 首先执行一般性检测并把可疑的路由送入数据库,以备后来进一步分析;然后,它进行特别检测并把异常路由也存放到数据库;最后,它根据数据库中的信息生成一份安全报告。

与通常的检测工具不同,ISP-Health 不但能检测出常见的问题路由,如含有私有前缀、AS 号的路由,而且能识别出许多特别的异常路由,如违背商业关系的路由等(如表 1 所示)。所有这些都是基于多视图的,因此比其他工具更加全面和准确地判定异常路由。

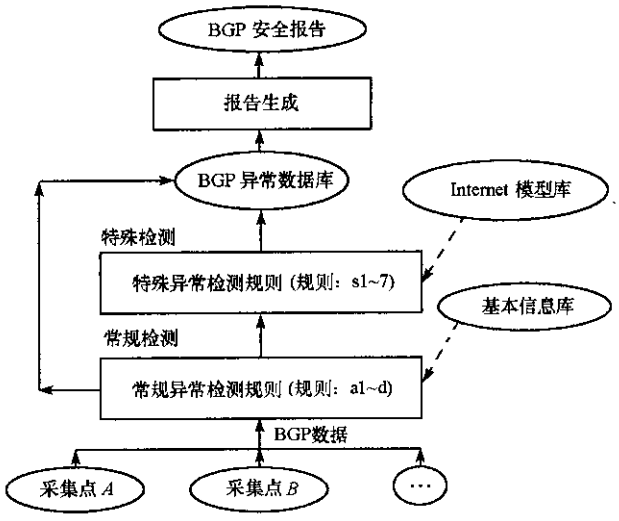

图 3 ISP-Health 系统结构

Fig.3 The architecture of ISP-Health

表 1 ISP-Health 的能力

Tab.1 The capabilities of ISP-Health

异常路由的类型	能力	异常路由的类型	能力
含有私有前缀	√	含有自治系统环	√
含有未分配前缀	√	伪造路由	√
含有保留前缀	√	不匹配 Internet 层次模型	√
含有私有自治系统号	√	违背 ISP 商业关系约束	√
含有未分配自治系统号	√	MOAS 冲突	√

5 结束语

如今域间路由系统存在许多问题，如 BGP－4 易配置错误、面临恶意攻击威胁等。本文提出一个基于规则的框架来保证域间路由系统安全，主要讨论规则引擎的设计、异常检测规则及相关 Internet 模型。应用该异常检测技术构造一个基于规则的检测系统 ISP-Health，并展示了它的几个优势：可根据定义的规则有效检测异常路由，规则引擎易于扩展，以及容易实现等。下一步，将主要考虑 Internet 的其它特性，如地理特性、小世界特性及密率特性等，以用于域间路由安全领域。

参 考 文 献：

[1] Halabi B. Internet Routing Architectures[M]. Indianopolis：Cisco Press，Second Edition，2001.

[2] Kent S，Lynn C，Seo K. Secure Border Gateway Protocol (Secure-BGP)[J]. IEEE Journal on Selected Areas in Communications，2000，18(4)：582－592.

[3] Murphy S. Border Gateway Protocol Security Analysis[EB/OL]. IETF Internet Draft，draft-murphy-bgp-vuln-00.txt. Nov. 2001.

[4] Cowie J，Ogielski A，Premore B，et al. Global Routing Instabilities During Code Red Ⅱ and Nimda Worm Propagation[EB/OL]. http：//www.renesys.com/projects/bgp_instability.

[5] Misel S A. Wow，AS7007！NANOG Mail Archives[EB/OL]. http：//www.merit.edu/mail.archives/nanog/1997－04/msg00340.html.

[6] Mahajan R，et al. Understanding BGP Misconfiguration[A]. ACM SIGCOMM'2002[C]，2002.

[7] Gao L. On Inferring Autonomous System Relationships in the Internet[A]. In IEEE Global Internet Symposium[C]，2000.

[8] Subramanian L，Agarwal S，Katz R H. Characterizing the Internet Hierarchy from Multiple Vantage Points[A]. INFOCOM 2002[C]，2002.

[9] Broido A，Nemeth E，Claffy K. Internet Expansion，Refinement and Churn[A]. ETT Janu 2002[C]，2002.

[10] Mahajan R，Wetherall D，Anderson T. Understanding BGP Misconfiguration[A]. In ACM SIGCOMM[C]，2002.