doi:10.3969/j.issn.1000-436x.2014.z1.019

基于流记录的非对称路由检测

兰浩良，丁伟，夏震
(东南大学 计算机科学与工程学院，江苏 南京 211189)

摘 要：针对同时使用多个 ISP 接入服务的校园网，网络边界路由的错误配置导致非对称路由现象发生，提出了一种基于流记录的非对称路由检测(FARD, flow-based asymmetry routing detection)方法。该方法利用 TCP 面向连接的传输特性结合 IP 地址的归属，基于边界路由器提供的流记录数据定位网络中可能存在非对称路由的 IP 地址。算法基于 CERNET 2 个主节点的接入路由器流记录进行了验证。

关键词：TCP；流记录；TCP 单向流；非对称路由；路由检测
中图分类号：TP393.0 **文献标识码**：A **文章编号**：1000-436X(2014)Z1-0098-05

Asymmetric routing detection based on flow records

LAN Hao-liang, DING Wei, XIA Zhen
(School of Computer Science and Engineering, Southeast University, Nanjing 211189,China)

Abstract: The misconfiguration of border routers may cause asymmetric routing in campus networks which connect to multiple ISP. A method FARD was proposed according to the phenomenon. This method uses TCP connection-oriented transmission characteristics combined with the home IP address, locating the possible asymmetric routing IP address in the network based on the flow records provided by the router. Experiment on the access router flow records from the main point of CERNET network to demonstrate the proposed method.

Key words: TCP; flow records; TCP unidirectional flows; asymmetric routing; routing detection

1 引言

非对称路由指当源主机 A 与目标主机 B 进行数据传输时，从主机 A 到主机 B 的数据分组选择了特定的路径 R1(如 Chinanet 链路)，而从主机 B 返回到主机 A 的数据分组却因为各种原因选择了另外一条不同的路径 R2(如 CERNET 链路)，当此情形出现时，就认为非对称路由发生。在一般情况下，非对称路由不会对传输产生太大的影响。但在某些特殊情况下，如传递的数据分组经过一些状态检测防火墙时，非对称路由可能会导致网页只能部分加载以及应用程序不能正常工作等问题[1]。除此之外，邬海涛[2]等人通过数学建模以及仿真的方式证明非对称路由可能会导致信道 TCP 吞吐量的减少。另外，由于往返路径延迟、带宽等链路质量的不同，非对称路由的发生也增加了网络行为的不确定性，对实时音频、视频应用非常不利，服务质量得不到保证。在当前的 IP 网络环境中，非对称路由是不希望出现的一种情形。因此在非对称路由发生时，能及时地发现并找到导致非对称路由发生的原因，对网络管理员及运营商都具有重要意义。

引起非对称路由的原因有很多，其中路由器错误配置导致非对称路由的情况比较常见。在互联网目前的体系结构下，一个企业网或校园网如果使用多个 ISP 提供的接入服务，则需要在不同的接入信道上使用不同 ISP 提供的 IP 地址，这种情况在 CERNET 校园网环境中普遍存在。在实际情况中，企业网或者校园网一般通过边界路由器支持这项工作，如果边界路由器错误配置了不同运用商的 IP 地址，将导致非对称路由的出现，但此时网络仍然可以正常工作，因此，这样的配置错误很难被发现。

本文研究工作围绕这个问题展开，提出了一

收稿日期：2014-10-14

种基于流记录的非对称路由检测方法 FARD (flow-based asymmetry routing detection)。利用 TCP 面向连接的传输特性，通过接入点路由器提供的流记录[3]，筛选出存在非对称路由行为的 TCP 单向流，结合 IP 地址归属，定位到可能存在非对称路由的接入单位。

2　场景描述

在同时使用多个 ISP 接入服务的校园网(如校园网 X)中，校园网边界路由配置错误导致非对称路由的情形之一可以描述为：校园网 X 将其自身使用的属于 ISP A(如 CERNET)的 IP 地址集合 IPsetA 错误地指向了 ISP B(如 Chinanet)。此时从校园网 X 发出的以 IPsetA 为源地址的流量直接进入了 ISP B 的网络。这样的配置会导致非对称路由的出现。

根据配置错误的主机 IPa (属于 IPsetA)交互的对象 IPb 是否与 IPa 属于同一网络，非对称路由有 2 种更具体的情况（这里假设 IPb 的配置是正确的）。1) IPb 不属于同一网络，具体情形如图 1 所示。此时，在网络边界的观测点上只能观测到入方向的正常流量，这些流量报文的 SIP∈IPb, DIP∈IPa。2) IPb 属于同一网络，具体情形如图 2 所示。此时，在网络边界的观测点上只能观测到入方向的正常流量，这些流量报文的 SIP∈IPa, DIP∈IPb。

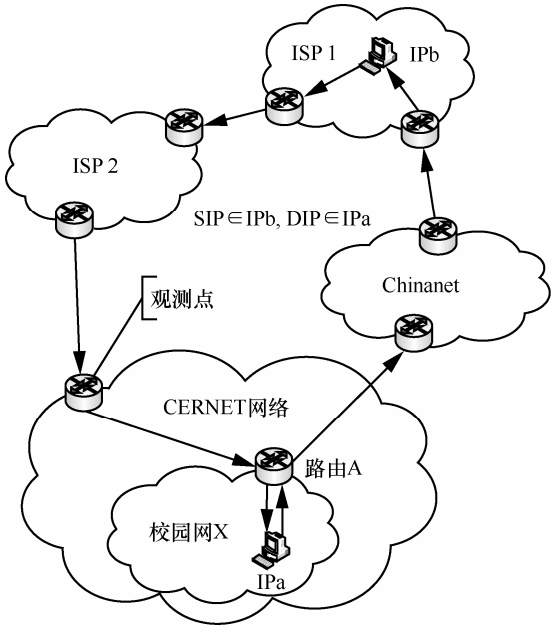

图 1　情况 1 下的非对称路由

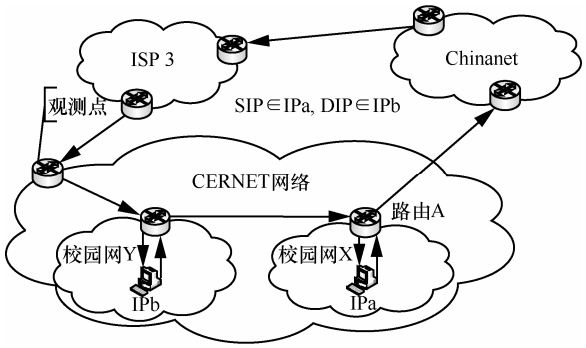

图 2　情况 2 下的非对称路由

3　非对称路由检测方法 FARD

如果观测点设在 IPa 所属的网络边界处，通过场景描述可以看出，配置错误引起的非对称流量呈现 2 个特征：一是在观测点为入方向的单向流；二是交互双方的流量构成一个闭合环路。因此，如何筛选这种类型的流量将成为本文所提方法 FARD 的关键。为得到这种类型的流量，首先在观测点筛选出入方向的单向流，然后再从这些单向流中过滤掉不满足特征 2 的流量，得到本文用于检测非对称路由的网络流。

分析流经观测点的单向流，影响特征 2 闭合环路形成的因素有以下 2 点。

1) UDP 流及 ICMP 流等不总是需要数据分组接收端产生应答的网络流，即使在观测点呈现单向性，交互双方的流量也可能不会构成环形，因此，算法在入出 2 个方向同时排除这类流量，仅使用具有闭合环路特征的 TCP 流量作为检测依据。

2) 源 IP 欺骗流量；这类流量也会在观测点呈现单向性，但不能构成闭合环路。特别是当网内正常配置的地址假冒错误配置的地址向网外地址发送流量的情况发生时（如假冒源地址的 DDoS 攻击），会在观测点上出现到出方向流量，这些流量会影响算法对单向流量的判定。因此，算法必须将这类流量筛选出来过滤掉。

通过以上的处理，得到用于非对称路由检测的有效 TCP 单向流，FARD 算法的核心思路是从完整的流量信息中过滤出这些流量，并依据这些流提供的数据，定位网络中可能存在非对称路由的 IP 地址。

本文分析数据源是网络边界（图 1 和 2 中的观测点）路由器提供的流记录，在实际运行环境中流记录通常基于抽样报文产生。抽样比的存在不能保证每个

报文均被抽样，这会导致误判的发生，比如，在观测点的双向 TCP 流，若出方向上的报文未被抽中，而入方向的被抽中了，就会在观测点形成入方向有效的 TCP 单向流，造成误判。为此，FARD 算法给出的解决策略：1) 进行 IP 地址聚合；2) 为入方向的有效 TCP 流量设定面向报文数的阈值，只有单向有效报文达到一定数量后，方可认定是有效流量。通过这样的方法降低流抽样带来的不良影响，减少误判。

3.1 源 IP 欺骗流量过滤

FARD 算法的核心思路是通过检测出有效的单向 TCP 流量来定位可能存在的配置错误导致的非对称路由。这里有效的 TCP 流量指使用 TCP 协议的网络应用在正常的交互中产生的流量。

TCP 提供的是一种面向连接的可靠传输，正常的 TCP 连接的建立需要经过 3 次握手。使用欺骗 IP 地址不可能完成这个过程。如图 3 所示，若存在源 IP 欺骗，则伪造数据分组的客户端将无法收到服务器端返回的 SYN+ACK 数据分组，也就无法完成第 3 次 ACK 数据分组的发送。这样，FARD 算法通过过滤掉 SYN 单分组流，来避开源 IP 欺骗的影响。其次，为了提高算法的效率，同时避免异常 FIN 及 RST 攻击，算法也过滤掉 FIN 单分组流及 RST 单分组流。

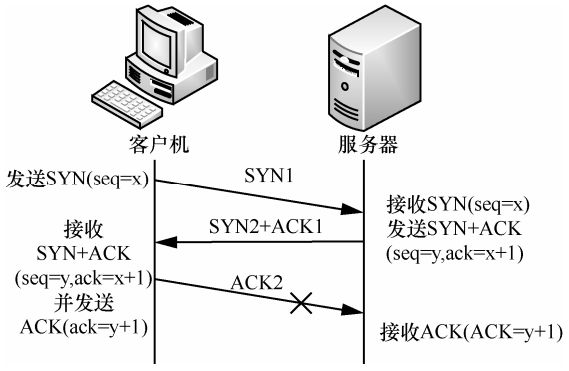

图 3 源 IP 欺骗示例

3.2 IP 地址聚合粒度

为了避免因抽样导致的误判，综合考虑路由操作及地址分配策略[4]，FARD 算法采用聚合地址空间的方法来解决这个问题，即将一个连续的地址空间作为一个处理单位，来进行有效单向 TCP 流检测。首先将这个连续的地址空间定为 CIDR/24 地址块。做出这个选择的一个重要原因是认为对一个校园网而言，一个 CIDR/24 地址空间要么部分配置错误，部分配置正确（或部分配置错误，部分不配置）；要么全部配置正确（或部分正确配置，部分不配置）；或全部配置错误。同时，使用 CIDR/24 作为单位也降低了 DHCP（动态主机分配协议）及抽样比对检测带来的不良影响。结合场景描述中陈述的情况，体现到 FARD 算法中的逻辑为：一个 CIDR/24 为非对称的当且仅当该 CIDR/24 内的部分 IP 地址只存在超过阈值的入方向上的具有情形 1 或情形 2 所述特征的有效 TCP 单向流量。

3.3 FARD 算法描述

为获得用于非对称路由检测的 TCP 单向流，首先要移除本质上可能存在非对称路由的网络流，如 UDP 流、ICMP 流等不总是需要数据分组接收端产生应答的网络流；接着利用 TCP 面向连接的传输特性，过滤掉建立了完整 TCP 连接的双向流以及 TCP 背景辐射流量[5](借助 SYN、RST、FIN 等标志位发起的网络扫描、探测、攻击等所形成的 TCP 单分组信令流)，尽管这类网络流比特数较少，但在流数上也占有相当的比例[4]，因此，不可忽略。

经过以上几个步骤的处理，再结合场景描述中的 1)、2) 2 种情况，利用算法 1 输出配置错误的 CIDR/24 地址段。

算法 1 基于流记录的非对称路由检测方法 FARD

输入 网络边界路由器提供的网络流记录 F_s，待判定的网络地址空间子集 ADR_{as} 及观测时长 T。

输出 ADR_{as} 中可能存的配置错误的 CIDR/24 地址段。

数据结构 构造整形数组 IP[n]、OUT[n]、packets[n,2]和 traffic[n,2]、scene[n]，其中，n 为待判定的 IP 地址空间 ADR_{as} 所包含的 CIDR/24 地址段数量，所有数组初值为 0。其中，IP[i]记录确定配置错误的 CIDR/24；OUT[i]=0/1，对应出方向的某个 CIDR/24 是否存在有效 TCP 流量；packets[i, j]，记录 i 所对应的 CIDR/24 在入方向的有效 TCP 流量的报文数，j=0/1 分别对应 SIP 和 DIP；traffic[i, j]记录记录 i 所对应的 CIDR/24 在入方向的有效 TCP 流量字节数，j=0/1 分别对应 SIP 和 DIP；scene[i]=0/1 分别对应场景描述中所述情况 1 宿地址配置错误及情况 2 源地址配置错误。

算法描述如下。

Step1 读入一个流记录 Fs，判断其是否有效，有效则继续执行，否则读取下一条流录。

Step2 若该流记录的源、宿地址均属于 ADR_{as}，且源地址对应的 CIDR/24 的 IP[i]=0，则令

OUT[i]=0，scene[i]=1，IP[i]=源地址对应 CIDR/24，并修改 packets[i,0]和 traffic[i,0]数组的对应位置；若该流记录的源地址不属于 ADR_{as} 而宿地址属于 ADR_{as}，且宿地址对应的 CIDR/24 的 IP[i]=0，对出方向有效流量则置对应 OUT[i]=1，对入方向有效流量修改 packets[i,0]和 traffic[i,0]数组的对应位置；否则读取下一条流记录。

Step3　重复 Step1 与 Step2 至观测时常 T 终止。

Step4　输出所有(OUT[i]=0∪IP[i]≠0)∩ (packets[i,j]>阈值)的 CIDR/24 和对应 packets[i,j]、traffic[i,j]及 scene[i]，此时，该 CIDR/24 存在错误配置。其中，有效 TCP 流的判定条件为：不包含 TCPflag 为 SYN、ACK、FIN 或 RST 的单分组流。

4　实验结果与分析

实验借助 NBOS(network behavior observation system)平台完成。NBOS 是在 CERNET 211 三期工程支持下开发的一个面向流记录的综合处理平台，它安装在 PoP 的接入边界，其主要功能是根据流记录的流终止时间，以 5 min 为时间粒度对来自接入路由器的流记录进行整理后输出。整理工作包括对 5 元组相同的流进行合并、对乱序情况进行处理等操作。其输出的数据格式中还会给出源宿地址的归属信息和流量的方向，其中，地址归属信息的编码中包含 PoP 编号。

实验观测时间为 2014 年 8 月 1 日到 2014 年 8 月 7 日。根据以往的经验及知识储备，有针对性的选取了 6 所规模小，管理能力相对较弱以及 6 所规模相对较大、管理能力相对较强的校园网进行观测，观测点构建在 CERNET 省网边界信道上。其中，6 所在江苏省网内，分别记为 JS-A、JS-B、JS-C、JS-D、JS-E 及 JS-F。另外 6 所属于另一个接入点 POP-A，分别记为 POP-A-1、POP-A-2、POP-A-3、POP-A-4、POP-A-5 及 POP-A-6。实验结果见表 1 所示，其中，Univ 表示被观测校园网，Num 表示校园网拥有的 C/24 地址段个数，Pro 表示 FARD 检出的配置错误的 C/24 地址段个数，SIP 表示属于源地址配置错误的 C/24 地址段个数，DIP 表示属于目的地址配置错误的 C/24 地址段个数，Region 表示配置错误的 C/24 的地址。

通过表 1 可以看出，JS-A 存在源地址配置错误的情况，在 JSERNET 省网边界通过对该段地址采集的报文进行 DPI 分析，发现该校园网的这个 CIDR/24 中的 4 个 IP 地址出现如图 2 所示的情况 2 的配置错误。除 JS-A 外江苏省网内其余 2 所高校 JS-B、JS-C 出现了场景描述中情况 1 所述的配置错误，配置错误的 CIDR/24 如表 1 中的 Region 列所示。本文同样用 DPI 的方法对 JS-B、JS-C 相关地址段进行了分析，证实确实存在相关 IP 地址的配置错误。而江苏省网内的另外 3 所规模相对较大的校园网 JS-A、JS-B 及 JS-C 在观测时间内，没有出现配置错误引起非对称路由的现象。FARD 在另外一个主节点 POP-A 的运行发现 POP-A-1、POP-A-2、POP-A-3 存在相关 IP 地址的非对称流量及情况 1 所述的配置错误问题，配置错误的 CIDR/24 如表 1 中的 Region 列所示。而 POP-A-4、POP-A-5、POP-A-6 等 3 所规模相对较大的校园网在观测时间内没有发现配置错误引起非对称路由的现象。

表 1　　**实验结果**

Univ	Num	Pro	SIP	DIP	Region
JS-A	16	1	1	0	180.*.32.0/24
JS-B	24	4	0	4	219.*.188-191.0/24
JS-C	24	2	0	2	210.*.92-93.0/24
JS-D	80	0	0	0	-
JS-E	98	0	0	0	-
JS-F	64	0	0	0	-
PoP-A-1	24	4	0	4	211.*.186-189.0/24
PoP-A-2	16	1	0	1	121.*.57.0/24
Pop-A-3	24	3	0	3	211.*.140-142.0/24
PoP-A-4	208	0	0	0	-
PoP-A-5	64	0	0	0	-
Pop-A-6	104	0	0	0	-

5　结束语

本文对网络中可能存在的路由配置错误导致的非对称路由现象进行了分析，在此基础上提出了一种基于流记录的非对称路由检测方法 FARD。在 NBOS 平台的支持下，算法原型在包括南京在内的 CERNET 的 2 个主节点上运行，并成功发现多个因校园网边界的错误配置导致的非对称路由现象。通过对 JSERNET 省网边界捕获的全报文的 DIP 分析，证明了算法的有效性。

后继的工作将从 2 个方面展开：一是在对算法

进一步完善的基础上将其正式集成入 NBOS 系统，在更大的范围内进行非对称路由检测；二是对算法进行全面的分析和优化，包括抽样对算法的影响、选择更加合适的 CIDR/n 作为分析粒度以及更加全面地过滤噪声流量等[6]。

参考文献：

[1] HE Y H, FALOUTSOS M, KRISHNAMURTHY S. Quant hy. quantifying routing asymmetry in the internet at the AS level[J]. IEEE GLOBECOM, 2004, (3): 1474-1479.

[2] WU H T, LONG K P, WU J, *et al*. Performance analysis of TCP over bandwidth asymmetry networks[J]. Journal of Beijing University of Posts and Telecommunications,2000, (23):30-33.

[3] RYU B, CHENEY D, BRAUN H W. Internet flow characterization: adaptive timeout strategy and statistical modeling[A]. Workshop on Passive and Active Measurement[C].2001.94-105.

[4] JOHN W, DUSI M., CLAFFY K. Estimating routing symmetry on single links by passive flow measurements[A]. ACM IWCMC, 2010. 473-478.

[5] BENSON K, DAINOTTI A, CLAFFY K, *et al*. Ga-ining Insight into AS-Level Outages through Ana-lysis of Internet Background Radiation[A]. IEEE TMA[C]. 2013.

[6] DAINOTTI A, BENSON K, KING A, *et al*. Diamantopoulos. estimating Internet address space usage through passive measurements[J]. SIGCOMM CCR, 2014, 44(1):42-49.

作者简介：

兰浩良（1986-），男，山东德州人，东南大学博士生，主要研究方向为网络测量与行为学。

丁伟（1962-），女，江苏南京人，东南大学教授、博士生导师，主要研究方向为网络测量与行为学。

夏震（1976-），男，江苏南京人，主要研究方向为网络测量、网络行为学等。