基于运营商网络的大带宽违规分析检测研究

Reasearch on Detection of Large Bandwidth Violation Based on the Operators' Network

田　园，汪襄南，胡学良（中讯邮电咨询设计院有限公司，北京 100048）
Tian Yuan，Wang Xiangnan，Hu Xueliang（China Information Technology Designing & Consulting Institute Co.，Ltd.，Beijing 100048，China）

摘　要：
主要阐释了何谓大带宽违规，并从违规的接入方式入手，通过研究大带宽违规的接入特征、流量特征，从用户向外发出访问的流量、用户向外发出的请求包数量、用户流量中VPN流量的占比、用户流量中异常协议等几个维度，对大带宽违规现象进行识别检测，最终筛查出疑似违规的用户流量，从而为规范互联网接入市场提供有力支撑，为提升公司收入保驾护航。

关键词：
IDC；大带宽违规；外访流量；HTTPGET；VPN；异常应用
doi：10.12045/j.issn.1007-3043.2019.08.016
中图分类号：TN919.2
文献标识码：A
文章编号：1007-3043（2019）08-0074-03

Abstract：
The definition of the large bandwidth violation is explained. From the illegal access mode，the large bandwidth violations is identified and detected by the study of the access characteristics and traffic characteristics of large bandwidth violation，the amount of traffic sent from the user，the number of request packets sent by the user，the proportion of VPN traffic and exception protocols in the traffic. Therefore，it provides great support for standardizing the Internet access market and increases the company's income.

Keywords：
IDC；Large bandwidth violation；Traffic sent from the user；HTTPGET；VPN；Exception protocols

引用格式：田园，汪襄南，胡学良. 基于运营商网络的大带宽违规分析检测研究[J]. 邮电设计技术，2019(8)：74-76.

1 研究背景及意义

互联网和信息技术的发展在经济和生活的各个领域正在迅速普及，其地位日益重要，整个社会对网络的依赖程度越来越大。与此同时，也产生了各种各样的问题。网络带宽的违规私接日益猖獗。这种不正当市场竞争不仅严重损害各大运营商的利益，而且不法分子还可能利用私接宽带发送垃圾邮件、搭建不良信息网站、甚至从事网络违法活动，给网络安全监控及治理带来了不便和难题。能否通过监控分析用户流量数据，自行发现违规行为，从源头上扼制违规私接的发生，已成为各大运营商关注的重点。

收稿日期：2019-07-03

2 大带宽违规

IDC机房正常业务模式为宽带用户通过公网请求访问IDC机房内的服务器业务，IDC服务器给予应答，具体如图1所示。

所谓大带宽违规，则是宽带用户通过某IDC机房出网访问互联网上其他资源，IDC机房内某用户将带宽引出，为其他宽带小区提供转租。

下面以北京联通和某省联通、某省电信IDC机房为例，讲解大带宽违规私接时宽带用户访问互联网数

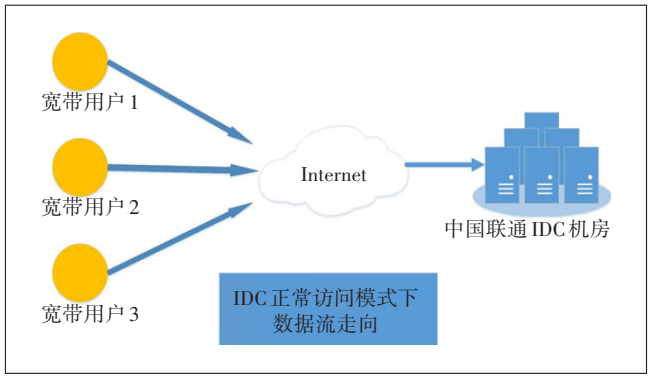

图 1　IDC 机房正常业务模式图

据包流向。私接方通过某省电信为用户做接入服务，以某省电信到某省联通，再到北京联通建立 VPN 通道，使得某省电信接入用户的访问互联网流量实际通过北京联通 IDC 机房作为网络出口，帮助了某省电信宽带市场的发展，影响当地联通宽带及楼宇专线等业务发展，具体如图 2 所示。

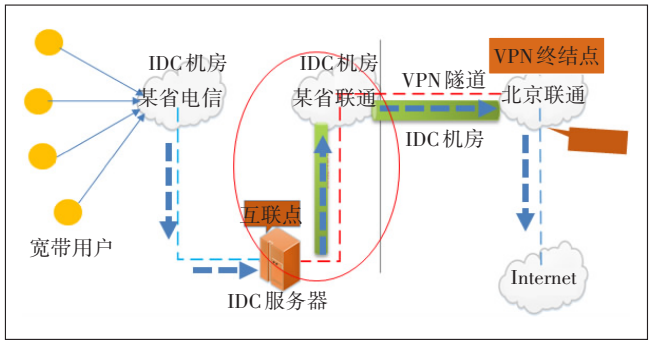

图 2　大带宽违规示意图

3 大带宽违规的分析及检测

3.1 数据获取的实现

实现大带宽违规私接的分析检测，需获取数据中心的全部网络流量，并在此基础上，利用深度数据包和大数据技术，进行流量的分析和检测。

网络流量数据可以通过数据分流提取和代码植入等方式获取。

代码植入方式是在数据中心所有客户端网页中嵌入代码，以监测访问客户网络的流量情况，然后将网站访问情况反馈至分析系统。数据分流提取方式是在运营商的数据中心网络出口部署数据分流提取监测设备，并将分流提取出的流量送至分析系统处理。鉴于运营商网络的数据中心出口带宽较大，客户众多，数据分流提取方式更为可行。

3.2 大带宽违规的分析方法

3.2.1 IDC 机房向外访问流量

在不考虑 IDC 机房存在 CDN 服务器的情况下，传统 IDC 业务模式为终端用户通过公网请求访问 IDC 机房内的服务器业务，不会存在 IDC 机房的服务器向外网请求访问业务的情况。如存在大量从 IDC 机房内部服务器发出的向外访问的请求，则怀疑此 IDC 用户流量具有公众业务流量的特征，认为存在违规嫌疑（见图 3）。

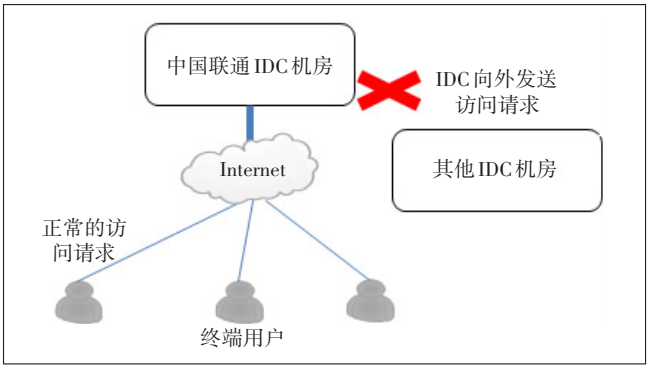

图 3　IDC 机房异常向外访问流量图

3.2.2 VPN 流量及 VPN 流量成分

如果发现某 IDC 机房用户 VPN 隧道的流量较多，且 VPN 隧道流量中存在大量其他运营商的 IP 地址，或者 IP 地址所产生的流量非常大，这都是非正常 VPN 业务，可以认为存在违规嫌疑（见图 4）。

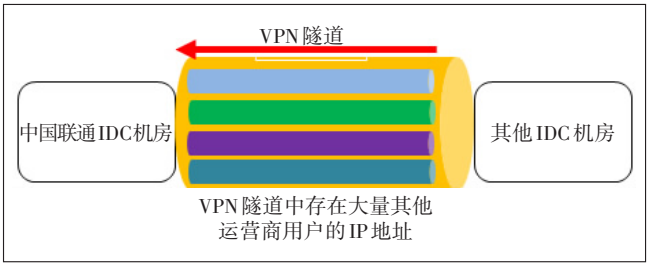

图 4　VPN 隧道流量示意图

3.2.3 向外访问的特殊应用

一般而言，IDC 用户业务模式相对统一。如某 IDC 用户业务为传统 Web 网站服务，则其服务器发出的流量多为 Http 流量；又如某 IDC 用户业务为视频服务，则其服务器发出的流量协议多为 P2P 协议。如果 IDC 机房内某用户向外访问的流量中，应用协议类型较多，且多为 QQ、微信、购物、游戏等应用，则认为其具有终端用户的访问特征，存在违规嫌疑（见图 5）。

图5　向外访问流量成分图

3.3 大带宽违规的检测手段

首先对机房向外访问流量、机房向外访问HTTP-GET数量、VPN疑似流量占比、异常应用统计等4个维度进行统计分析，分别加权打分；然后根据分值，对疑似私接IP进行整体汇总排名与审核。

3.3.1 机房外访问流量排名

“机房外访问流量”即为IDC机房内的服务器，主动向机房外部、公网上的站点发起的访问流量。大带宽违规检测的手段之一是针对此类流量进行排名、展示。可基于对应服务器IP（即其业务IP），进行外访流量统计与TOP100排名，前30个IP赋值疑似度60%，后70个IP赋值疑似度40%，如果发现有非本运营商的IP，其疑似度赋值直接提升至80%。

3.3.2 机房外访HTTPGET数据包数量排名

“机房外访HTTPGET数据包数量”即为IDC机房内的服务器主动向机房外部、公网上的站点发起的HTTPGET访问数据包总量。可基于对应服务器IP（即其业务IP），进行外访HTTPGET数据包数量统计与TOP100排名，前30个IP赋值疑似度60%，后70个赋值疑似度40%，如果发现有非本运营商的IP计入，其疑似度直接提升至80%。

3.3.3 VPN疑似流量占比排名

基于服务器业务IP，考查VPN流量在总流量中的占比，占比高于96%的，计入VPN疑似流量占比TOP100。VPN占比的计算公式如下。

VPN占比=VPN总流量/（进总流量+出总流量）

计入VPN疑似流量占比TOP100的IP，赋值疑似度75%，如果发现有非本运营商的IP计入，疑似度直接提升至85%。因VPN涉及到内外两层IP地址，外层IP用来判断本运营商的合法性，内层IP地址用来判断其他运营商IP归属，2种归属信息均需展示。

3.3.4 异常应用统计排名

通过识别分析IDC机房内某个服务器的业务IP流量中的QQ账号、邮件账号和UA数量，找出“1个IP承载多个应用账号标识”的IP信息，按照应用标识数量（如QQ账号、UA数量等），进行TOP100排名、展示。计入TOP100的IP，赋值疑似度75%，如果发现有非本运营商的IP计入，疑似度直接提升至85%。某些IP可能涉及到内外两层IP地址，这种情况下，内层IP地址需按照是否归属于其他运营商进行判断，外层IP地址需按照是否是本运营商的合法IP进行判断，2种归属信息均需展示。

3.3.5 疑似IP汇总审核

前述的4个分析维度，疑似度在60%以上的IP均列入疑似IP汇总表，按疑似度做TOP50排名。若有IP在几个维度之间交叉出现的，每交叉出现一次，对比取其最高的疑似度值，并在此基础上增加5%的疑似度，将结果纳入汇总表。最终，对疑似违规私接的IP地址疑似度重新排名，排名越靠前，违规的疑似度越大。

4 结束语

本文通过利用DPI及大数据技术，引入大带宽违规私接的监测模型，深度分析客户流量特征，辅助判断客户是否存在疑似违规接入的行为。大带宽违规私接的识别，为规范市场，治理客户业务提供了直观的技术支撑手段。

参考文献：

[1] BHAIJI Y. 网络安全技术与解决方案[M]. 北京：人民邮电出版社，2009.
[2] CONVERY S. 网络安全体系结构[M]. 北京：人民邮电出版社，2005.
[3] 陈兵，王立松. 网络安全体系结构研究[J]. 计算机工程与应用，2002，38(7)：138-140.
[4] 王立新，周元. 计算机网络安全技术的问题及解决办法[J]. 中国科技信息，2008(6)：94-95.
[5] 林波. 宽带用户行为分析与探究[J]. 电子测试，2014(21)：96-98.
[6] 张治宇. 互联网宽带用户行为分析系统的设计与应用研究[J]. 数字通信世界，2017(8)：90，99.
[7] 杨波. 通信运营商宽带用户行为分析的研究与应用[J]. 邮电设计技术，2014(11)：71-76.
[8] 延皓. 基于流量监测的网络用户行为分析[D]. 北京：北京邮电大学，2011.
[9] 杨庆. 宽带网络非法接入监控系统[J]. 经济视野，2013 (12).

作者简介：
田园，工程师，硕士，主要从事网络安全运营相关工作；汪襄南，工程师，主要从事骨干网络安全运营相关工作；胡学良，工程师，现主要从事骨干网络安全运营相关工作。