

Analysis and design on key techniques of DPI traffic inspection

LIU Chuang, XIN Yang

(Information Security Center, Beijing University of Posts and Telecommunications, Beijing 100876, China)

Abstract: With the wide range of applications such as P2P and VoIP network technology, a large number of P2P, VoIP protocols and application software began to appear, such as uTorrent, BitTorrent, BitComet, Thunder, SinaTV, Skype, PPStream, PPLive, etc., Although it makes our life more convenient, it also brings a lot of serious problems. Such as taking up a lot of network bandwidth, spreading network viruses. How to limit and manage P2P traffic is becoming the current research focus which is also a problem that need to be resolved by network operators and network managers. In this paper, we realize the protocol analysis of network traffic by researching the key technologies of the DPI traffic detection. We also improve traffic identification technology by proposing encryption rule and combination of keyword and packet length rule, design the network traffic detection process, do experiment which shows that the proposed scheme can improve the recognition rate of network traffic.

Key words: P2P; VoIP; DPI; protocol analysis; traffic inspection

DPI 流量检测关键技术的分析与设计

刘 创, 辛 阳

(北京邮电大学 信息安全中心, 北京 100876)

摘 要: 随着 P2P 和 VoIP 等网络技术的广泛应用, 大量 P2P 和 VoIP 协议及其应用软件开始出现, 如 Utorrent、BitTorrent、BitComet、Thunder、SinaTV、Skype、PPStream、PPLive 等, 其在给人们带来生活上的方便的同时, 也带来很多严重的问题, 如占用大量的带宽传播网络病毒。目前如何有效的限制和管理 P2P 流量成为当前网络的研究热点, 也是运营商和网络管理者急需解决的问题。本文通过研究 DPI 流量检测关键技术, 来实现网络流量的协议分析, 对流量识别技术进行了改进, 提出了加密规则和关键字包长结合体规则, 设计了网络流量检测流程, 并进行了实验, 实验结果表明所提方案具有提高网络流量识别率的作用。

关键词: P2P; VoIP; DPI; 协议分析; 流量检测

1 引言

目前, 随着 Internet 日益普及和网络结构的日益复杂, 网络的安全性、可管理性及其传统应用的可用性收到了严重挑战。虽然 P2P 和 VoIP 等已经

成为重要的网络应用, 其流量已占据互联网流量 50%~80%, 它不论在知识产权的保护方面还是在防病毒抗攻击等方面都存在很多问题, 但是制约 P2P 和 VoIP 等发展的关键是盗版问题和流量控制问题, 此外, 如何测量和评估 P2P 和 VoIP 等应用的性能以及如何使网络提供满意的网络服务, 是一个

资助信息: 中央高校基本科研业务费专项资金资助(2012RC0215, 2012RC0216)。

急需解决的问题。在 P2P 环境下,方便的共享和快速的选路机制,为某些网络病毒提供了更好的入侵机会^[1]。总之,网络流量的分析与控制,对网络流量管理具有非常大的意义。首先流量检测技术是流量监控的实现的先决条件,必须能对网络流量的识别率达到非常高的水平之后才能对网络流量进行监控。流量检测技术对网络数据流量进行协议类别的分类,这种技术可以将网络流量的分类精确到协议级别,即可以检测到用户正在使用什么样的协议或网络软件,如可以把网络流量按照协议分成 P2P(如 BT、迅雷等下载类软件),VoIP(如 skype 等语音通话软件),IM(如 QQ、ICQ 等即时通信软件),PeerCasting(如 ppfilm 等视频播放器),流量检测的重点在于全面性和准确率。目前常见的流量识别技术有三种,一种是以 DPI(Deep Packet Inspection,深度包检测)技术为代表的特征字检测方法;一种是以 DFI(Deep Flow Inspection,深度流检测技术)为代表的基于数据流特征的检测方法;还有一种就是前两种的结合 DPI&DFI 形式。

2 流量检测技术的相关工作

2.1 DPI 技术

DPI(Deep Packet Inspection)即深度数据包检测技术,是一种相对简单、高效的数据流检测技术,它基本上是一种基于“特征字”的识别技术^[2]。深度数据包检测技术中所谓深度是与普通检测技术比较得出的,普通报文检测仅包含五元组信息,即源 IP 地址、源端口、目的 IP 地址、目的端口和传输层协议类型。而深度报文检测不但包含五元组信息的检测还包含应用层的分析,识别各种应用情况。DPI 检测技术通过对数据流中特定数据报文中的特征信息的检测以确定数据流承载的应用,结构化的 P2P 网络资源搜索机制采用了随机图的组织方式,这个特征可能是固定的端口,对应软件应用特征的字符串,或者有某种规则的 bit 序列。DPI 技术就是通过对应用流中的数据报文内容有规律的进行探测,最终确定报文的应用意义。典型的 DPI 检测技术有以下几类:

(1) 端口检测技术^[3]

端口检测技术是根据 TCP 数据包或 UDP 数据包首部的源端口或目的端口号是别一些常见协议的流量。这种检测方法是属于普通报文检测,这类

端口检测大都是基于 IANA 注册的知名的端口,例如 HTTP 协议的端口是 80,TCP 流端口 22 是远程登录协议,如 SSL 对应的端口就是 22,这些端口通常小于 1024,通过这些知名的端口我们可以容易判断相应的应用程序。有些应用程序的端口号大于 1024,但他们端口号是不变的,这可以辅助我们判断网络流量的对应的协议和软件应用,如一些游戏的端口基本上是不变的,例如当前流行的三国杀游戏的端口就是 41006,这样可以结合其他特点确定这段游戏应用产生的流量。

(2) 应用层特征匹配检测技术^[4]

不同的应用层协议有不同的协议信息,这些信息可以作为应用层协议的检测特征。这种方法又称为基于“特征字”的识别技术,可以理解为不同协议通常会有不同特征字,不同的特征字把网络流量协议一一区分开,如 SinaTV 软件产生的流量数据报文负责部分又“Sina”“TV”等字段,这些特征和其他检测技术配合可以确定这种软件产生的流量。基于“特征字”的识别技术又可以分为固定位置特征字匹配,不定位置的特征匹配以及状态匹配三种。

(3) 基于包长特征的检测技术

由应用层数据编码的结构体中可以看出,数据包的负载部分往往会有报文长度信息。大部分互联网软件在报文表现形式上都会体现报文长度信息。

(4) 应用层网关识别技术^[5]

应用层网关识别技术某些业务的控制流和业务流是分离的,业务流没有任何特征。这种情况下,我们就需要采用应用层网关识别技术。应用层网关需要先识别出控制流,并根据控制流的协议通过特定的应用层网关对其进行解析,从协议内容中识别出相应的业务流。对于每一种协议,需要有不同的应用层网关对其进行分析。如 SIP/H.323 协议都属于这种类型。SIP/H.323 通过信令交互过程,协商得到其数据通道,一般是 RTP 格式封装的语音流^[6]。也就是说,纯粹检测 RTP 流并不能得出这条 RTP 流是通过哪种协议建立的。只有通过检测 SIP/H.323 的协议交互,才能得到其完整的分析。在实际的报文分析中,这种情况并不少见,因为互联网流量中 P2P 和 VoIP 软件的流量占有很大一部分,许多 VoIP 软件语言通话的流量识别都需要这种技术。

2.2 DFI 技术

DFI(Deep Flow Inspection,深度流检测)技

术^[7]。不同于 DPI 常用的基于数据包的特征字检测思路,DFI 技术通过据流之间的共性,以此作为业务识别的依据。分析整条数据流中部分或者所有的数据包,找出其中数据包之间的关系,或者数据流之间的共性,以此作为软件流量识别的依据。比如某软件产生的数据流中前三个报文成等差数列,或者这条流中所以报文长度均在 15~30 之间,这条流中数据包之间的这种规律就可以作为识别这条流的依据之一。对于一些比较复杂的协议,单独使用 DPI 技术和 DFI 很难识别,必须将二者结合起来能进行识别。在 DPI 和 DPL 技术的识别互联网流量的关键之一,互联网有成千上万的软件,怎么样把他们的流量一一映射的识别出来,这需要后台特征数据库设计的非常合理才能达到这一目的^[8]。

3 特征数据库的设计

3.1 数据包流量检测的解析

假设一条数据流 L 有 n 个数据包, $L = (b_1, b_2, b_3, b_4, \dots, b_n)$, 互联网网络流量协议分类为 $\text{Class} = (c_1, c_2, c_3, c_4, \dots, c_n)$, 则可以将流量检测的定义为一个从流 L 到协议 Class 的一一映射 f , 使得 $f: L \rightarrow \text{Class}$ ^[9]。对于基于 DPI 流量检测分析, 数据流的 n 个数据包组成, 每个数据包由包头 header 和负载 load 组成, 在包头信息中有源 IP 地址 IP_{src} , 源端口 PORT_{src} , 目的 IP 地址 IP_{dst} , 目的端口 PORT_{dst} 和连接状态 Connection status , 在负载部分 load 含有协议特征 $\text{Protocol characteristics}$ 和交互数据 data 。对于数据包检测可以理解为对于任意一个数据包存在一个协议与之对应, 即

$$\forall b \in L, \exists p = \text{header} + \text{load}$$

其中 $\text{header} = \{\text{IP}_{\text{src}}, \text{PORT}_{\text{src}}, \text{IP}_{\text{dst}}, \text{PORT}_{\text{dst}}, \text{Connection status}\}$

$$\text{load} = \{\text{Protocol characteristics}, \text{data}\}$$

即 $\forall b \in L, \exists p = \{\text{IP}_{\text{src}}, \text{PORT}_{\text{src}}, \text{IP}_{\text{dst}}, \text{PORT}_{\text{dst}}, \text{Connection status}\} + \{\text{Protocol characteristics}, \text{data}\}$

这是数据包层次上分析, 一条流由 n 个数据包组成, 普通检测主要检测 header 中的信息, DPI 检测技术主要分析 load 中协议特征, 比如特征字信息, 包长信息^[10]。一般一条流前几个数据包包含负载信息, 也是体现这条流于其他流不同的地方, 后面大多数数据包是交互的数据并没有明显信息, 我

们可以以前 m 个数据包的特征来确定流的特征。

一条数据流 L 有 n 个数据包, $L = (b_1, b_2, b_3, b_4, \dots, b_n)$, 那么前 m 个数据包的特征可以用以上分析方法得到特征为 $p_1, p_2, p_3, \dots, p_m$ 。那么这条流的协议特征就是前 m 个数据包特征的交集

$$P_L = p_1 \cap p_2 \cap p_3 \cap p_4 \cap \dots \cap p_m$$

一个互联网软件由很多功能和场景, 每个功能和场景都可能会有不同的流, 所以我们分析协议的由软件功能场景为单位划分。假如一个互联网软件场景 S 有 n 条数据流, $S = (L_1, L_2, L_3, L_4, \dots, L_n)$, 一条数据流 L 有 n 个数据包, $L = (b_1, b_2, b_3, b_4, \dots, b_n)$, 由以上分析方法可以得到每条流的特征为 $p_{L1}, p_{L2}, p_{L3}, p_{L4}, \dots, p_{Ln}$ 。那么这个软件场景的协议特征为这 n 条流特征的交集。

$$p_S = p_{L1} \cap p_{L2} \cap p_{L3} \cap p_{L4} \cap \dots \cap p_{Ln}$$

如果 $p_S = p_{L1} \cap p_{L2} \cap p_{L3} \cap p_{L4} \cap \dots \cap p_{Ln} = \emptyset$, 这些流中找不到共同的特征把所有流都包含, 这时候应该把这些流分类成若干部分, 比如分成两部分 $S_1(L_1, L_3, L_5, \dots, L_{n-1}), S_2(L_2, L_4, L_6, \dots, L_n)$, 分配的原则是有共同的特征集合在一起。这样就有

$$p_{S1} = p_{L1} \cap p_{L3} \cap p_{L5} \cap \dots \cap p_{Ln-1}$$

$$p_{S2} = p_{L2} \cap p_{L4} \cap p_{L6} \cap \dots \cap p_{Ln}$$

$$p_S = p_{S1} \cup p_{S2}$$

在这种分析方式中, 多条流横向比较就会用的 DFI 的检测技术, 如每条流的前三个包依次为 10, 12, 15。有时也会用到 DPI 的应用层网关识别技术, 一个软件场景的两条流往往是控制流和业务流的关系, 需要解析出一条流然后关联出另一条流。

假如一个软件 Software 有 n 个场景, $\text{Software} = (S_1, S_2, S_3, \dots, S_n)$

$$P = p_{S1} \cup p_{S2} \cup p_{S3} \cup \dots \cup p_{Sn}$$

假如一个软件 Software 有 n 个不同版本, $\text{Software} = (V_1, V_2, V_3, \dots, V_n)$

$$P = p_{V1} \cup p_{V2} \cup p_{V3} \cup \dots \cup p_{Vn}$$

以上我们从数据包, 数据流, 软件的场景, 软件的整体, 软件的不同版本, 分析了流量特点以及分析方法。这些是后台特征数据库设计的依据和基础。

3.2 流量检测识别特征数据库的组建

特征数据库是流量特征的一种正则表达式, 是体现软件流量特征的数据集合, 协议特征数据库是整个流量监控系统的基础, 流量监控系统主要点是引擎将应用层数据采取字符串匹配算法与特征数

据库匹配,从而将识别结果反馈到流量控制模块。在特征数据库中结构是根据上文分析的流量特点设计的。因为一个软件有很多版本,很多场景,所以一个软件对于的协议往往不是一条特征就能描述完整的,往往由很多特征规则组合在一起的。即一个协议是由许多规则组成,假如一个协议为 p , 协议 p 由规则 $ruld_1, ruld_2, ruld_3, \dots, ruld_n$ 组成。 $p = ruld_1 \cup ruld_2 \cup ruld_3 \cup \dots \cup ruld_n$

即每条规则都是这个软件产生的流量的反应,只是可能代表不同场景,不同版本,不同平台流量识别,或者也可能是同种场景中产生的流量用一条规则不能完全表示。这就需要规则 $ruld$ 的组成因子要足够的多,具有区别力。我们把组成规则 $ruld$ 的因子称为关键因子 KEY 。比如一个规则由 $KEY_1, KEY_2, KEY_3, \dots, KEY_n$ 组成。

$ruld = KEY_1 \cap KEY_2 \cap KEY_3 \cap \dots \cap KEY_n$ 关键因子越多,规则 $ruld$ 越具有区分力,也称规则强,否则称为弱规则,容易产生误报,由 DPI,DFI 技术的研究中我们可以得到关键因子的选择。首先端口和 IP 可以作为关键因子的选择,为了加强区分维度增加了源端和目的端的区别,这样源端口 $PORT_{src}$ 和目的端口 $PORT_{dst}$,源 IP 地址 IP_{src} 和目的 IP 地址 IP_{dst} 产生了四个关键因子。由 DPI 技术分析中得出关键特征字和包长也可以作为关键因子。

3.3 流量识别规则的设计

为了更好地区别流量协议,我们把协议类型分成多种类型,比如单包规则、多包规则、等差多包规则、加密规则、深度解析规则、其他疑难规则,这样分类有助于网络流量的识别,也帮助协议分析工程师更好地研究网络协议的特点。

(1) 单包规则

图 1 和图 2 是 P2P 下载软件迅雷(版本号 v7.1.8.2298)下载产生的流量中报文上行流中的两个包的特点。

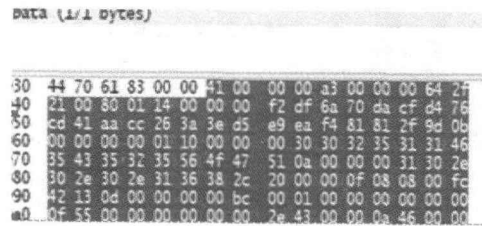


图 1 数据包 1

由图 1 和图 2 明显看出,从负载偏移 0 的位置

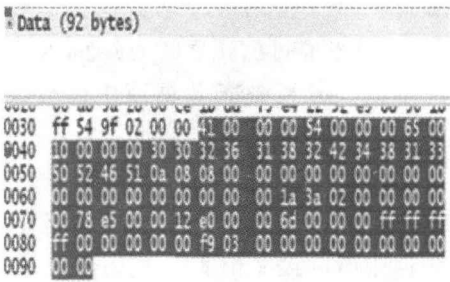


图 2 数据包 2

开始后四个字节都是 41 00 00 00 是不变的,经过验证其他上行流的报文也有这个特征,因为应用层携带的数据并非是纯粹的数据,应用层数据的开头分一般还包含协议软件的一些信息,再观察发现第一个包长为 171,负载偏移 4 位后的十六进制是 a3,即十进制 163,即负载偏移四位的数字加上数字 8 即为包长。数据包 2 也是符合这样的规律即十六进制 54 等于十进制 84,再加上数字等于 92。在验证其他上行流的报文也是符合这一规律。这种方法在一条流里比较前后数据包在一个数据包中提取的规则,也称单包规则^[11]。流量可以通过这样的规则识别。

(2) 多包规则^[12]

当协议特点不是单个数据包体现的,是有多个数据包体现的时候,多包是由多个单包规则构成,单包规则之间的关系为是否顺序、是否方向和是否连续。所谓顺序是指数据包特征按照规则顺序出现。所谓方向是指所有单包规则体只能出现在同一方向。所谓连续是指所有单包规则体必须连续出现。

(3) 深度解析规则

网络中需要传输的流量比较大的数据时往往会先使用一小流量来交互建立连接,通常用 TCP 流量,交互流称为信令流,通常传输层协议为 TCP,因为 TCP 确保数据传输准确。然后根据交互得到的三元组的 IP 端口,和传输层协议进行数据传输。数据传输的这条流称为数据流,通常为传输层协议为 UDP。整个过程的流程图如图 3 所示。

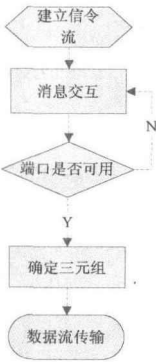


图 3 深度解析协议流程图

数据流一般没有什么特征很难直接从数据流中提取协议,但信令流通常通过分析出数据流的三元组信息,三元组信息一般会在特征字符串后体现^[13]。

(4) 统计规则

有些网络流量特征不是很明显,一条流中一个或者两个数据包中很难发现规律和特征。这时候就需要用统计的方法对一条流中的前 32 个数据包进行统计分析。同向固定包长序列、双向固定包长序列、同向固定包长集合、双向固定包长集合、同向连续包长范围、双向连续包长范围、同向连续包长平均值、双向连续包长平均值、同向连续包长求和、双向连续包长求和、指定包长复现、同向连续包数统计、同向包数统计。一条统计识别规则的统计条件可以是上面几种方式的组合,各条统计条件之间可以是“与”或者“或”的逻辑关系。

(5) 加密规则

有些情况由于传输的需要数据流的加密比较简单,加密算法通常是与 AND,或 OR,异或 XOR。

有两种加密方式静态加密和动态加密,两者不同的是动态加密时,密钥是从报文中获取的。

例 1:一条流中前三个包的特点:

包一 00 00 1f 23 21 00 78 89 c1 00 21 b1 00 97 21 98

包二 12 98 cf 45 09 09 87 b2 b3 91 43 f2 0c 12 43 92 32 09 00 98

包三 c1 32 5f 12 00 24 93 77 29 35 e3 53 93 9c 87 c3 00 00 21 33 5f

这三个包按照常规的单包和多包是不能找到其规律的,但是负载偏移 2 个字节后的第一个字节与 0x2F 做与运算结果为 0x00,这个静态加密规则的加密算法是与运算 AND,密钥串是 0x2F,密钥长度为 1,匹配负载开始位置为偏移 2 的字节,运算结果为 0x00。

例 2:一条流中的前三个包的特点:

包一 00 00 29 0b e3 cc b1 50 10 c1 70 06 39 00 12 00 98 00

包二 01 c3 12 c4 32 54 62 90 01 00 00 45 89 c9 b2 b3 cc

包三 23 b1 bc e2 c1 00 91 89 72 22 36 59 46 15 10 00 16 08 07 c2 07

这个报文的规则是比较难发现的,是属于动态加密的规则,负载偏移 4 个字节后第一个字节与静态密钥 0xAF 做异或运算的到动态密钥。三个包的

动态密钥分别为 $(0xE3) \text{ XOR } (0xAF) = 0x4C$, $(0x32) \text{ XOR } (0xAF) = 0x9D$, $(0xC1) \text{ XOR } (0xAF) = 0x6E$ 。负载偏移六个字节后的第一个字节分别与各自的动态密钥做与运算得到的结果都为 0x01。即 $(0x4C) \text{ AND } (0xB1) = 0x01$, $(0x9D) \text{ AND } (0x62) = 0x01$, $(0x6E) \text{ AND } (0x91) = 0x01$ 。这个动态加密规则的动态密钥匹配位置为偏移 4 的字节,动态密钥长度为 1,静态密钥为 0xAF,计算动态密钥的加密算法为异或运算,特征匹配位置为负载偏移 6 个字节,匹配长度为 1,加密算法为与运算,运算结果为 0x01。这就是这条动态加密规则完整的描述。

(6) 关键字和包长结合体规则

有些网络流量的报文特点是块状连接的,块块的大小往往是一些字节经过操作运算符得到,这些操作运算符为异或 XOR,与运算 AND,右移 RIGHT,左移 LEFT,加 ADD,减 SUB 等。这些包的特点是特征字后面的字节反应块块的大小,每一个块块前部的特征字的特点都有一定的联系。

例 1:一个数据包的特点:

a0 c9 00 01 bd 41 70 34 92 c8 00 02 22 33 76 98 87 67 21 12 81 ca 00 0a 12 32 12 12 23 34 01 0f 50 43 32 30 30 00 39 30 31 34 00 00 00 67 68 6e 73 74 72 61 22 00 00 41 6d 69 6e 69 02 00 0d 41

分析这个报文的发现这是又三块组成,a0 c9, 92 c8,81 ca,为块状前的特征部分,00 01,00 02,00 0a 分别体现之后数据块的大小,00 01 的大字节序为 01 00,即二进制 01 0000 0000,然后向右移 6 位得到 0100 即后面的数据块的大小 4 个字节,00 02 的大字节序为 02 00,即二进制 010 0000 0000,然后右移 6 位得到 01000 即后面的数据块为 8 个字节,00 0a 的大字节序为 0a 00,即二进制数 1010 0000 0000,然后右移 6 位得到 1010 00 即后面的数据块大小为 40。经验证符合规则规律。这个规则可以这样描述:特征出现位置为负载偏移 0,特征长度为 2,特征在 80c8 到 alca 之间,反应块块大小的负载偏移为 2,长度为 2 个字节,操作运算为右移 6 位。

4 网络流量检测流程设计

网络流量识别是通过引擎来实现的,引擎需要加载特征数据库,引擎的对流量的识别通过一系列识别算法来实现的。这些识别算法和特征数据库中的规则是对应的有端口识别算法、关联识别算法、特征识别算法和行为识别算法。报文在发送给

引擎进行识别之前,已经被做了包头解析,传送给引擎的实际上是一个结构体,数据包送入引擎后的识别过程如图 4 所示。

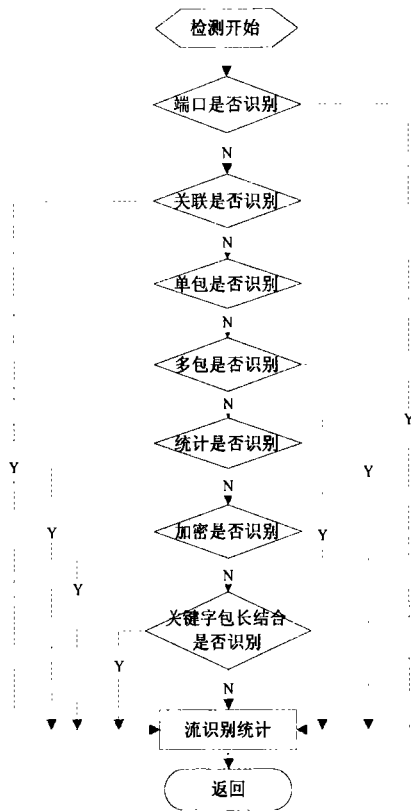


图 4 网络流量检测流程图

5 网络流量检测结果

网络流量的主要有 P2P、VoIP、IM 类等软件流量产生,本文检测互联网常用的软件检测结果如表 1 所示。

表 1 常见软件检测识别结果

软件名称	识别率	误报率
QQ	99.12%	0.00%
Thunder	98.23%	0.00%
SinaTV	97.12%	0.00%
BitComet	96.90%	0.00%
Skype	94.09%	0.00%
PPLive	96.23%	0.00%
UUsee	99.09%	0.00%
PPStream	95.89%	0.00%
utorrent	97.06%	0.00%
Maze	96.13%	0.00%
MSN	99.43%	0.00%
PPfilm	92.65%	0.00%
ICQ	96.34%	0.00%
YY	92.06%	0.00%

6 总结

根据本文所提出的基于 DPI 和 DFI 的技术设计的特征数据库,协议规则设计的更加详细规范,网络流量能更好的一一映射到特征数据库中的协议规则,进一步提高了网络识别率,对于一些比较棘手的网络流量分析给出了分析的办法,对特征数据库中的规则进行了详细的说明和实例分析。从表 1 中可以看出,网络常见的软件识别均到达 90% 以上,部分软件流量识别在 95% 以上。网络流量的识别和不识别是一个不断反复的过程,此外研究机器学习,减少人工分析报文,使用人工智能方法来实现报文识别也是进一步研究的重点。

参考文献

[1] P2P 的应用研究、面临的问题与前景展望. <http://www.acejoy.com/space/html/73/n-73.html>.
 [2] 王飞. IP 网数据综合监控系统体系结构及关键技术研究[D]. 北京:北京邮电大学, 2007.
 [3] 汤昊,李之棠. 基于 DPI 的 P2P 流量控制系统的设计与实现通信技术第 6 期.
 [4] 米淑云. IP 网络流量监控系统的设计与实现[D]. 北京:北京邮电大学, 2009.
 [5] 韩耀明. 基于 DPI 技术的 VoIP 流量检测系统的设计与实现. [D]. 北京:北京邮电大学, 2010.
 [6] ITU- T recommendation H. 323, PacketBasedMultimediaCommunicationSystem, 1998.
 [7] 李江涛,姜永玲. P2P 流量识别与管理技术电信科学 2005.
 [8] 王珊,陈松,周明天. 网络流量分析系统的设计与实现[J]. 计算机工程与应用. 2009.
 [9] J. Erman, A. Mahanti, M. Arlitt, C. Williamson. Identifying and Discriminating Between Web and Peer to Peer Traffic in the Network Core. In WWW'07, Ban, Canada, May 2007:883-892.
 [10] 王蛟. 基于行为的 P2P 流量及异常流量检测技术研究[D]. 北京:北京邮电大学, 2008.
 [11] 李万鹏. 网络流量控制及流量分析[D]. 北京:北京邮电大学, 2011.
 [12] 赵轶. 基于 P2P 流量监控系统的设计与研究[D]. 北京:北京邮电大学, 2010.
 [13] 张应. 基于综合特征的 P2P 流量识别与控制系统研究[D]. 上海:复旦大学, 2009.