文章编号: 1006-4729(2006)01-0063-04

报文分类算法的设计与实现

汪 伟[1], 孙 翌[2]

(1. 中国地质大学 信息工程学院, 北京 100083;
2. 上海交通大学 图书馆, 上海 200030)

摘 要: 报文分类是网络交换设备的基础操作之一. 针对目前的报文分类算法主要为小规则集或低维的分类问题, 提出了一种新的多维快速报文分类算法. 该算法通过压缩、分割和索引操作, 设计了压缩分割规则表和索引列表结构表这两类数据存储结构, 能支持上万条多维分类规则, 分类速度快、存储空间小、支持快速更新, 适合于软硬件实现.

关键词: 报文; 分类算法设计; 压缩; 分割; 索引
中图分类号: TP393 文献标识码: A

Design and Implementation of Packet Classification Algorithm

WANG Wei[1], SUN Yi[2]

(1. *School of Information Engineering, China University of Geosciences, Beijing 100083, China*;
2. *School of Library, Shanghai Jiaotong University, Shanghai 200030, China*)

Abstract: Packet classification is a fundamental operation performed in networking equipment such as switchers and routers. Many packet classification algorithms exist focusing on a few rules or low dimension classification. With the booming of the requirement of network management, network intrusion detection and QoS, a new fast scalable multi-dimension packet classification algorithms, based on PCBNP, RFC and regional partition algorithm, is given in this paper. By designing two storage data structure, compress partition rule table and index list table, this algorithm can support ten thousand multi-dimension rules, and have fast rate of packet classification, low expected storage complexity and can be implemented in software or hardware.

Keywords: packet classification; design of algorithm; compress; partition; index

报文分类系统广泛应用于路由器、防火墙, 以及入侵检测系统和流量监控系统中. 报文分类必须以线速转发报文, 因而要求路由器能够及时转发从线路进入的最小分组(64字节). 这对上述的各种网络应用来说是至关重要的, 否则路由器就有可能在未确定报文的重要性之前将一些重要的报文丢掉. 报文分类的本质是快速查找, 它的困难在于分类的多维性(报文分类的字段)和规则数的日益增多. 因此, 直接简单套用快速查找的算法是行不通的[1], 必须研究分类规则的特性及其相互之间的关系等诸多因素才有可能找到更有效的解决方案, 从而为各种网络技术的发展奠定坚实的基础[1]. 与现有的报文分类算法相比较, 本算法是一种综合性能较好的报文分类算法, 因而具备较强的软硬件通用能力, 可适用于较广的应用领域, 如网络智能小区建设等[2].

收稿日期: 2005-12-02

1 压缩分割索引算法

本算法最大的特点是不直接存储原规则集,而是把原规则集转化为压缩分割规则表和索引列表结构表两类数据存储结构,达到存储空间需求量小、分类速度快、能支持上万条多维分类规则的目的.

(1)删除冗余规则并以规则完全匹配的优先顺序排序　研究表明,在分类器中 8%的规则是冗余的,如果报文分类算法能够识别冗余规则并删除这些规则,则可提高报文分类的性能.然后以规则的完全匹配的优先顺序排序,排序的目的是要达到在顺序查找时第一个完全匹配的规则为最优的匹配规则.

(2)生成并存储压缩分割规则表　本算法不直接存储原规则集,而是存储由原规则集压缩分割生成的压缩分割规则表,目的是减少存储容量,加快分类的速度[3].

由规则集生成压缩分割规则表时,要进行压缩和分割操作.

(3)生成并存储索引列表结构表　生成并存储索引列表的操作我们称之为索引操作.索引操作是用 P 个特定的索引列表结构表对压缩分割规则表的 P 个分割字段进行映射,建立 P 个索引列表结构表.按分割字段的值作为索引值对索引列表结构表进行查找,就可以找到与该分割字段对应的所有域匹配的规则,这就加快了分类速度.同时我们给出了算法的整体映射图,见图 1.

图 1中的索引列表结构表是由压缩分割规则表生成的.

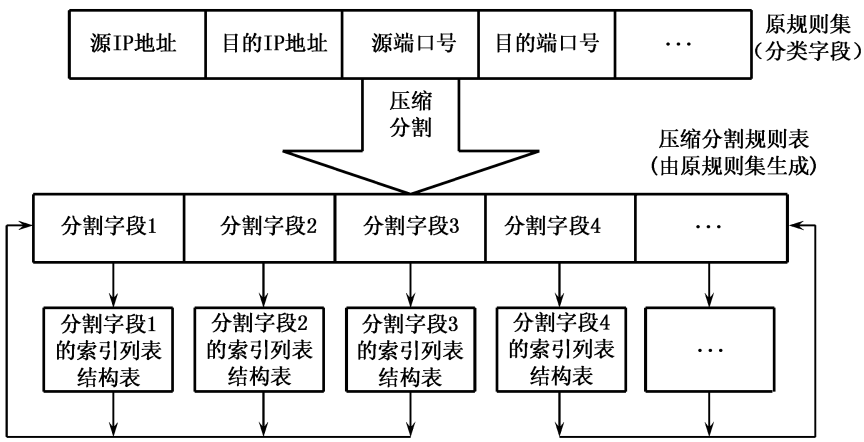

图 1　算法的整体映射示意

(4)查找过程　当接收到一个报文后,取出报文头中的要分类的所有字段,对分类字段用相同的压缩分割算法生成 P 个 16位分割字段(构成分割字段组),然后按照这 P 个分割字段的值分别在 P 个索引列表结构表中查找并各在一次访存中取出对应的列表结构(行).

2 算法举例

分类时必须取报文头字段.取报文头字段是指按分类组合的要求从进入的报文中取出相应的字段.常见的多维分类字段主要有两种组合[1]:源、目的 IP地址(各 32位)的两维分类(共 64位);源、目的 IP地址(各 32位),传输层协议类型号(8位),源、目的端口号(各 16位)的五维分类(共 104位)[4].

下面以五维分类为例对本算法设计了一个具体的例子.

2.1 由原规则集压缩分割生成压缩分割规则表

由规则集生成压缩分割规则表时,要进行分类域宽的压缩操作和压缩后分类域的分割操作.

(1)压缩操作　压缩操作类似于 RFC算法的第一次递归,但不采用启发式算法,而是采用以分析分类字段的特征为主、分析规则分布的统计规律为辅的固定的压缩算法.

设协议类型的值为 x_8,源端口号的值为 y_{16},

目的端口号的值为 z_{16}, 并按照前面的规定压缩后的值分别为 x_4, y_6, z_6, 则压缩后的宽度为 16位的 m 的值为 $m = x_4 + 2^4 y_6 + 2^{10} z_6$

这样 3个字段压缩为 16位的 1个字段, 原来的 40位缩减为 16位, 压缩率为 60%. 映射用的 16位字段如图 2所示.

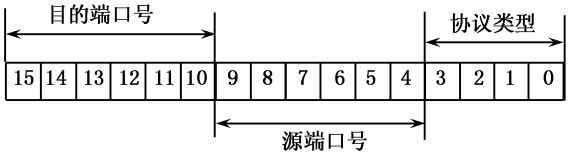

图 2　协议类型、源、目的端口号映射到一个16位字段内的压缩分配图例

源、目的 IP地址字段可以用类似方法压缩.

(2)分割操作　分割操作的思想类似于区域分割算法, 但要求与区域分割算法有区别: 一是本例中要求把压缩后的规则分割为 5个 16位的字段, 称为分割字段; 二是没有区域分割算法要求严格, 本算法对分割字段匹配的规则个数只要求能分割到在一次内存访问中取出即可. 假设内存中一次能取出的规则数是 7, 在本例中为小于 7个即可, 同时要求前面的分割字段匹配的规则个数尽量少. 具体做法是对源、目的端口号和协议类型合并为 16位, 加上 IP源、目的地址各 32位, 共 80位. 在建立压缩分割规则表时把 80位分为 5个 16位的分割字段, 而每个分割字段对应一个 64k行的索引列表结构表[5].

本算法给出的分割生成压缩分割规则表见表 1和表 2, 并由表 1经分割生成了表 2

表 1　压缩后的规则集

规则名称	源 IP地址字段	目的 IP地址字段	端口协议合并字段
规则 0	160. 25. 249. 18 0xa019f912e	161. 15. 248. 158 0xa108879e	0x0218
规则 1	164. 66. 39. 25 0xa427ee19	168. 25. 105. 24 0xa8196918	0x0008
规则 2	162. 24. 153. 159 0xa218999f	160. 62. 23. 130 0xa03e1782	0x1472
规则 3	162. 19. 53. 84. 57 0xa2135439	167. 224. 93. 7 0xa7e05d07	0x0f12
…	…	…	…

表 2　压缩分割规则表

规则名称	分割字段 1	分割字段 2	分割字段 3	分割字段 4	分割字段 5
规则 0	0x01f9	0x0887	0x0218	0xa012	0xa19e
规则 1	0x27ee	0x1969	0x0008	0xa419	0xa818
规则 2	0x1899	0xe178	0x1472	0xa29f	0xa082
规则 3	0x1354	0xe05d	0x0f12	0xa239	0xa707
…	…	…	…	…	…

2.2　由压缩分割规则表生成索引列表结构表

(1)索引列表结构表的数据结构　5个分割字段对应生成 5个索引列表结构表, 也就是说共有 5个 64k表.

索引列表结构表的索引值就是对应的报文的分割字段值. 每个索引列表结构的第一项 (列)是分割字段匹配的规则的个数, 其他项 (列)为指向压缩分割规则表的指针 (存储的是规则的序号值), 并且按规则的序号从小到大顺序存放. 这样, 就能争取一次内存访问全部取出所有本分割字段匹配的规则.

另外, 考虑到在有些特殊情况下, 字段匹配的规则的个数可能在一次内存访问中不能全部取出所有本字段匹配的规则, 因此, 每个索引列表结构的最后一项设计为两选一, 即指向压缩分割规则表的指针 (存储的是规则的序号值), 或者是指向本索引列表结构的下一个列表结构 (存储在索引列表结构链接表中)的序号指针. 两选一是由分割字段匹配的规则的个数项 (列)确定的, 如果字段匹配的规则的个数超过一次内存访问所能取出的本字段匹配的规则数, 那么最后一项就是指向本索引列表结构的下一个列表结构的指针 (存储的是链接库的列表结构的序号, 即行值)[6].

索引列表结构表 (64k表)的具体数据结构如图 3所示.

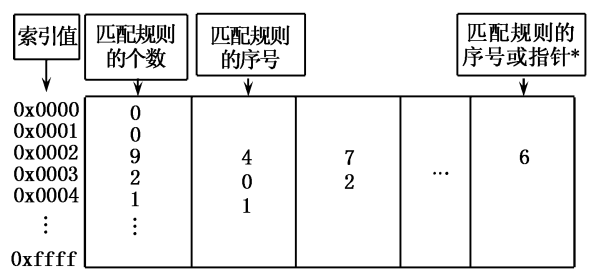

图 3　索引列表结构表 (64k表)的数据结构的具体实现示意图

(2)索引操作 是指一个过滤规则首先被压缩分割到压缩分割规则表中,每一个过滤规则被压缩分割为 5个分割字段,然后根据过滤规则的 5个分割字段的值的大小被插入到相对应的 5个索引列表结构表的某一列表结构(索引列表结构表的行)中.

进行索引操作的方法为:设某一个分割字段的值为 x,则压缩分割规则表的第 y 条规则的索引操作是指将该分割字段对应的索引列表结构表的第 x 行(索引列表结构表的一个列表结构)的第一列(第一项)的值 m 加 1,然后再在 $m+2$ 项中存储值 y;其他项则保持不变[7]. 其他分割字段的索引操作与此类似,这里就不赘述了.

2.3 查找过程

查找过程是当接收到一个报文后,取报文头中的要分类的所有字段,对分类字段用相同的压缩分割算法生成 5个 16位的分割字段(构成分割字段组),然后按照这 5个分割字段的值依次分别在 5个索引列表结构表中查找并各在一次访存中取出对应的列表结构(行). 在列表结构(行)中,第一个是对应分割字段匹配的规则的个数,如果为零,我们就立即返回(表示没有完全匹配的规则);如果不为零,当少于或等于还没有查找过的索引列表结构表的个数与一个给定数目 K(比如 $K=3$)的和(即 $K+5-i$, i为已经查找过的索引列表结构表的个数)时,利用其他的项,即指向分割表的指针,直接到压缩分割规则表中查找是否为完全匹配的规则,只要找到一个就可以立刻得到最优的完全匹配的规则. 本算法与其他算法的比较见表 3. 表 3中 N 为规则数, D 为分类的维数.

表 3 与其他算法的比较

算法名称	时间复杂度	空间复杂度	说 明
RFC	D	N^d	需要并行计算能力,不适合于大的规则集,需要大的存储空间
Hierarchical cuttings	D	N^d	存储空间需要量大
Tuple-space search	N	N	在最坏情况下,时间复杂度为 $O(N)$,速度太慢
PCBNP	D	N^2	多维,速度快,不需要并行计算能力,空间需求量较小,不适合于大的规则集
This	D	DN	多维,速度快,不需要并行计算能力,存储空间需求量小,支持上万条规则容量

从表 3中可以看到,本算法与已有的算法相比,能支持上万条规则,并多维、速度快、存储空间需求量小. 适合于软硬件实现,不需要并行计算能力,但支持并行计算能力. 在万条左右的规则的情况下,本算法是一种综合性能比较好的报文分类算法.

3 结束语

本文在 PCBNP算法[4]、RFC算法[7]和区域分割算法的基础上提出了一种新的可扩展的多维报文分类算法. 本算法首先对原规则集压缩和分割生成压缩分割规则表存储,然后通过建立多个具有 64k行的索引列表结构表这一数据存储结构,使该算法能支持上万条多维分类规则,并且分类速度快,存储空间需求量小. 总的来说本算法中第一阶段的压缩操作类似于 RFC算法的第一次递归,但不采用启发式算法,而是以分析分类字段的特征为主,分析规则分布的统计规律为辅,采用固定的压缩算法,以利于规则的快速更新、存储空间的减少和报文的快速分类. 分割操作类似于区域分割算法,但没有区域分割算法要求严格,本算法对分割字段匹配的规则个数及其各个规则只要求分割到能在一次访存中取出即可. RFC算法不适合于需要大的存储空间的大的规则集[8]. 因此,本算法第二阶段不采用类似 RFC算法的递归的方法,而是采用索引列表结构表(64k表)这一数据结构,以减少存储空间,支持大的规则集. 由于该算法数据结构固定,因而适合于软硬件实现,同时支持并行计算能力. 与现有的报文分类算法相比较,在几千到几万条的规则的情况下,本算法是一种综合性能比较好的报文分类算法,因此它具备较强的软硬件通用能力,可适用于较广的应用领域,如网络智能小区建设等[2].

参考文献:

[1] 田立勤,林 闯. 报文分类技术的研究及其应用[J]. 计算机研究与发展, 2003, (6): 765-775.

[2] 程启明. 基于 lonworks网络的智能小区概述[J]. 上海电力学院学报 2004, 20(1): 9-15.

[3] Ying Dar-Lin, Wei Huan-Yun, Wu kuo-Jui. Ordered Lookup with Bypass Matching for Scalable Per-flow Classification in Layer 4 Routers[J]. Computer Communications, 2001, (24): 667-676.

(下转第 70页)

随着分块数的增加,硬件复杂性越来越高,而PTS性能的改善却越来越小.所以在运用PTS法时,分块数 V 的取值要合适.

5 结束语

本文在总结目前降低峰均功率比方法的基础上,着重研究了信号扰码技术,对SLM法和PTS法进行了较为全面的分析比较.

到目前为止,还没有一种方法既能大幅降低峰均功率比,又能满足实现的复杂度低、编码速率高的要求.限幅和窗函数法会引起带外辐射和误码率的增加,但是与其他方法,如编码等结合,是很有研究价值的.PTS法虽然实现的复杂度较高,但是可以适用于任意数量的子载波数,而且星座调制的种类也不受限制,是一种比较好的降低峰均功率比的方法.

参考文献:

[1] Van Nee R, Prasard R. OFDM for Wireless Multimedia Communications[M]. Boston: Artech House, 1999.

[2] Davis J A, Jedwab J. Peak-to-mean Power Control in OFDM, Golay Sequences and Reed-Muller Codes[C], IEEE Trans Inform. Theory, 1999, 45: 2 397-2 417.

[3] Baum l R W. Reducing the Peak-to-average Power Ratio of Multicarrier Modulation by Selected Mapping[J], Electronics Letters, 1996, 32(22): 2 056-2 057.

[4] Stefan Muller H, Johannes Huber B. A Novel Peak Power Reduction Scheme for OFDM[J]. PIMRC'97, 1997, 1 090-1 094.

[5] Muller S H, Huber J B. OFDM with Reduced Peak-to-average Power Ratio by Optimum Combination of Partial Transmit Sequences[J]. Electronics Letters, 1997, 33(5): 368-369.

[6] 陈 琳, 郭振民, 华继钊, 等. 一种基于格雷互补序列抑制峰均功率比的方法[J]. 电视技术, 2003, (12): 25-27.

(上接第66页)

[4] 田立勤, 林 闯, 肖人毅, 等. 基于IXP1200的快速报文分类算法的设计与实现[J]. 计算机研究与发展, 2003, (11): 1 616-1 625.

[5] Baboescu F, Varghese G. Fast and Scalable Confict Detection for Packet Classifers[J]. Computer Networks, 2003, (42): 717-735.

[6] Pankaj Gupta, Nick McKeown. Packet Classification Using Hierarchical Intelligent Cuttings[J]. In Proc ACM Sigcomm, 1998, (9):

[7] 朱秋香, 陶 军. 流分类算法综述[J]. 小型微型计算机系统, 2004, (10): 1 802-1 810.

[8] 颜天信, 王永纲, 石江涛, 等. 区域分割包分类算法的优化实现[J]. 通信学报, 2004, (6): 80-88.