

# 实时网络流量分类研究综述

柏 骏 夏靖波 吴吉祥 任高明 赵小欢

(空军工程大学信息与导航学院 西安 710077)

**摘 要** 实时流量分类技术能够按照应用类型对在线网络流量分类,它对网络管理、流量控制以及网络相关研究具有重要意义。首先从不同层次上简单分析了实时流量分类技术的研究现状;给出了流量分类的实时性概念及其指标;然后从流量统计特征和机器学习算法两个方面综述了实时流量分类的主要技术及研究进展,并进行了实时性分析;最后根据未来网络发展对实时流量分类技术提出的新要求,展望了该领域未来的研究发展方向。

**关键词** 流量分类,实时性,流特征,机器学习

**中图法分类号** TP393 **文献标识码** A

## Survey on Real-time Traffic Classification

BAI Jun XIA Jing-bo WU Ji-xiang REN Gao-ming ZHAO Xiao-huan

(Information and Navigation College, Air Force Engineering University, Xi'an 710077, China)

**Abstract** Real-time traffic classification, identifying and classing online traffics according to their applications, is of fundamental importance to network managements, operations and correlative studies. The paper first analyzed the research statusquo of real-time traffic classificaiton from different levels. The definition and evaluation metric of real-time classifcaiton were given to address the situation that there is no measurable and accepted metrics. Next, the main techniques and research progresses of real-time traffic classifi-cation were summarized while the real-time property of eath techniques were analyzed from flow features and machine learning shcemes. Finally, we pointed out possible future research directions in real-time traffic classification according to the demand of future network.

**Keywords** Traffic classification, Real-time property, Traffic features, Machine learning

## 1 引言

随着网络规模日益庞大、业务种类日渐繁多、用户数不断攀升、网络行为日趋复杂,如何有效地进行网络管理控制、流量入侵检测以及网络规划建设,是当前所有网络服务提供商(ISPs)以及网络管理者亟需解决的问题。

实时流量分类能够按照应用类型对在线网络流量分类,可为上述问题的解决提供有力的技术支持。通过实时流量分类,ISPs可随时获知当前占据主要带宽的流量类型,并依此快速做出反应以支持其不同的网络运行目标,提供尽可能高的QoS;分析流量态势,为网络优化以及进一步的网络规划提供策略支持;企业或校园网络管理者可以在流量高峰时期根据需要实时管控网络中非关键流量(如P2P、VoIP、PPS等),以保证网络畅通;发现可疑网络流量,进行入侵检测,提高网络安全性能。

自20世纪末以来,学术界对网络流量识别、分类进行了大量研究,并取得了丰硕的成果。对于已有的网络流量分类方法,按照所用技术的不同,可将其分为基于端口识别技术<sup>[1]</sup>、深层数据包载荷检测技术(DPI)<sup>[2,3]</sup>、基于传输行为模

式的流量分类技术(BLINC)<sup>[4]</sup>和基于机器学习(Machine Learning)的流量分类技术<sup>[5]</sup>。上述方法各有优缺点,应用范围也不尽相同。

实时流量分类技术更加注重网络流量的在线识别、分类,是对流量识别技术研究的进一步深入、拓展。关于实时网络流量识别、分类的研究,可以从数据包级(Packet-level)和流级(Flow-level)两个层面上来展开分析。

在数据包级的层面上,只有基于端口的流量识别技术能做到在线实时分类,但是随着端口跳变技术、伪装技术的出现,该方法已不再可靠<sup>[16]</sup>,其识别准确率仅为50%~70%<sup>[2]</sup>,在P2P大量应用的场合其准确性更低,因此主要用作流量识别、分类的辅助手段。基于载荷的流量识别技术则需要对每个数据包进行深度检测,其时间、空间复杂度较高,在高速网络中难以满足实时性要求。因此,由于基于端口的方法在准确性上的不足以及基于载荷的方法在时间、空间复杂度上的苛刻要求,目前鲜有数据包层面上的实时流量识别技术研究。

在网络流级的层面上,上述几种方法均可用于实时流量识别、分类。由于只需要对网络流(Flow)第一个或前几个数据包进行分析就能识别出其应用类型,因此基于端口和基于

到稿日期:2012-11-21 返修日期:2013-03-27 本文受陕西省科技计划自然科学基金重点项目(2012JZ8005),全军军事学研究生课题(2010XXXX-488)资助。

柏 骏(1985—),男,博士生,主要研究方向为网络管理与信息安全,E-mail:peking1985-2005@163.com;夏靖波(1963—),男,博士后,博士生导师,主要研究方向为通信网络管理与评估。

载荷的流识别技术都具有非常好的实时性,但是由于新出现的网络应用大多采用规避检测技术,该类方法的流量识别准确性逐年下降。而机器学习的流量识别方法,尤其在 Bernaille<sup>[17,18]</sup>和 T. Nguyen<sup>[19]</sup>提出利用网络子流的统计特征进行流量分类之后,不但具备较高的流量分类准确性,而且具备较好的实时性。这个特性引起了学者的广泛关注,自 2006 年起,国内外涌现出大量关于实时网络流量分类的研究成果。

本文详细综述了 2006 年至 2012 年以来国内外关于实时流量识别、分类技术的研究进展。第 2 节提出以准确性和实时性两个指标对实时流量分类方法进行性能评价,并给出了实时性的定义及具体评价指标;第 3 节从网络流量统计特征和机器学习算法两个重点研究方向出发,阐述了目前主要的实时流量分类技术,并进行了实时性分析;最后分析了未来网络发展对流量分类技术研究提出的新要求,并给出了下一步的研究方向。

## 2 实时流量分类技术的评价指标

目前,在衡量流量分类技术的准确性方面,已有了具体量化的、公认的指标,但是对于流量分类的实时性,既没有准确、具体的概念,又没有可量化、公认的评价指标。为了更好地评估一个实时流量分类方法的性能,本文沿用了常用的准确性指标,提出了流量分类技术的实时性概念,并给出了具体指标。

### 2.1 准确性评价指标

准确性是流量分类技术的关键评价指标,通常以召回率(recall)和精度(precision)两项指标来评价分类结果。召回率和精度的计算方法如下:

$$recall = \frac{TP}{TP + FN}$$

$$precision = \frac{TP}{TP + FP}$$

式中,TP、FN、FP、TN 分别代表真正(true positive)、假负(false negative)、假正(false positive)、真负(true negative)。表 1 阐明了四者之间的关系。TP 表示属于 X 类的样本被正确分到 X 类的概率;其余亦然。显然  $TP + FN = 100\%$ ,于是有  $recall = TP$ 。

表 1 TP、FN、FP、TN 之间的关系

Classified as →	X	$\bar{X}$
X	TP	FN
$\bar{X}$	FP	TN

需要指出的是,大多数文献中关于准确性的指标都是基于流(Flow)的,即被正确分类的流数占网络所有流数的百分比。然而,Erman 等人在文献[20]中给出的数据集表明,0.1%的大流占整个流量字节总数的 46%。这说明即便是流的准确性达到了 99.9%,如果没有识别出 0.1%的大流,那么仍将损失 46%的字节准确性。因此在实际的流量分类效果评估中,在给出流的准确性的同时,最好也给出字节的准确性。

### 2.2 实时性定义及评价指标

#### 2.2.1 实时性定义

实时性能够反映流量分类技术在线、快速识别网络流量应用类型的能力。其定义如下:

定义 1(实时性) 利用流量分类技术对网络流进行在线

识别,假如从流的第 1 个数据包开始到识别出该流应用类型所用的时间远远小于网络流平均存在时间或通过某网络节点的时间,那么称该流量分类技术具有实时性。

针对当前研究中出现的准实时(Near Real-time)<sup>[21]</sup>流量分类,给出其非正式定义,以供参考。利用流量识别技术对网络流进行在线识别,假如在流快结束或结束后较短的时间内识别出该流的应用类型,称其具有准实时性。

#### 2.2.2 实时性评价指标

当前绝大多数的实时流量分类研究都基于机器学习算法,故根据定义,建立了基于机器学习的流量分类实时性评估模型。考虑到国内外关于流量识别、分类研究的环境差异性,尤其是所采用的数据集、实验设备对实验结果有较大影响,实时性评估模型基于以下假设:

(1)采用的数据集相同,数据包到达的间隔时间相同,记为  $t_0$ ;

(2)采用的实验设备相同,使用某种算法对样本进行识别的时间与流特征个数  $m$  成正比,记为  $mt_1$ ,其中  $t_1$  表示分类算法处理一个特征所用的时间。

实时性主要是通过在线流量应用识别的时间来反映的,因此采用从第 1 个数据包到达识别出该流所用时间  $t$  作为实时性指标。整个流量实时识别过程分为两个阶段——流特征的在线获取以及基于分类模型的应用类型识别,如图 1 所示。

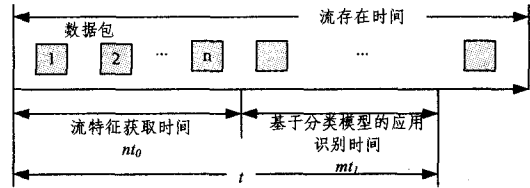


图 1 实时流量识别过程

于是,实时流量识别时间为

$$t = nt_0 + mt_1 \quad (1)$$

式中, $n$  为子流的数据包个数(网络流的前  $n$  个数据包)。更具体地说, $t_1$  受所选机器学习算法  $s$  影响,即  $t_1 = \varphi(s)$ 。因此,又有

$$t = nt_0 + m\varphi(s) \quad (2)$$

显然,根据假设,数据包到达时间  $t_0$  是相同的;对于算法  $s$  的选取,可认为总是选择最优的算法,即处理一个流量特征的时间是相同的。那么,实时流量识别时间  $t$  主要受子流数据包个数  $n$ 、流特征个数  $m$  影响,即  $t \approx f(n, m)$ 。

综上所述,本文将子流数据包个数  $n$ 、流特征个数  $m$  作为评价流量分类技术实时性的关键指标,记为  $(n, m)$ ,其值越小,实时性越好。

## 3 实时流量分类技术

虽然基于端口和基于载荷的识别技术也可以实现实时流量识别、分类,但是该类方法由于自身技术的限制,已不再是学术界研究的重点,且近年来并无相关方面的实时性研究,故本文以基于机器学习的流量分类技术作为重点研究对象,对实时流量分类技术进行详细阐述。通过第 2 节的实时性指标  $t = nt_0 + m\varphi(s)$  可以发现,流量分类方法的实时性由流量特征  $(n, m)$  和机器学习算法  $s$  决定。因此,本文主要从流量统计

特征以及机器学习算法两个方面来展开实时流量分类技术的讨论。

### 3.1 网络流量统计特征

基于机器学习的流量分类技术利用从被传输的网络流量(traffic)中获取的流(flow)统计特征达到区分流量应用类型的目的,这些特征可以是基于数据包(packet)的,也可以是基于网络流(flow)的。基于数据包的统计特征是以网络流中的数据包为研究对象,主要涉及数据包的大小、到达间隔、速率等等;基于网络流的统计特征以网络流为研究对象,其特征主要包括流(子流)大小(数据包个数、比特数)、持续时间、标志位个数等等<sup>[22]</sup>。

#### 3.1.1 基于数据包的特征

数据包在流的不同阶段表现出不同的特征,例如处于TCP连接协商阶段的数据包顺序是预先定义的,且不同应用的数据包大小是明显可区分的。Bernaille等人<sup>[17,18,23]</sup>正是利用这个特性,选择网络流前 $n$ 个数据包(不包含TCP控制报文,如SYN、ACKs等)的大小和方向作为流特征进行流量应用类型识别研究,取得了较好的分类效果,流准确性均达到了80%以上。Sena等人<sup>[24]</sup>也选择网络流前几个数据包的大小和方向作为流量特征,只是个数稍有不同,为5或者6个。彭建芬等人<sup>[25]</sup>不仅选择TCP流前3个数据包的大小作为流特征,还将服务器端口号作为辅助特征,并对上述特征进行组合,当选取特征1、2、4、1、2、3、4和1、2时,分类准确性都超过97%。

P. Branch等人<sup>[26]</sup>研究了Skype的数据包特性,发现对于字节数小于80的数据包,其大小总是某一个固定的值:44、45、48、51、58、60、65、73或者74,而且这些数据出现的频率较高。另外,Skype客服端产生数据包的间隔是16毫秒(ms)的倍数,即16、32、48或64ms,这个特性被证明是非常可靠的Skype识别特征。实验结果表明,单独使用上述两个特征时能达到90%左右的分类准确性,联合使用时则能达到95%以上的准确性。

数据包比例(Packet Ratio)也是非常重要的数据包特征。J. But等人<sup>[27]</sup>在研究Bittorrent流量识别方法时将载荷为5字节或者17字节的数据包定义为Bittorrent特征数据包,并将Bittorrent特征数据包个数占总数据包个数的比例定义为Bittorrent特征比( $r_{bit}$ ),此外还定义了小数据包比( $r_{small}$ )、大数据包比( $r_{large}$ )、两个方向上较小的载荷方差( $\sigma_{small}$ )。为了实现快速识别Bittorrent流量,作者选择150个数据包的网流子流作为研究对象,采用上述数据包特征,并使用C.45构建分类模型,达到了98.2%的召回率和96.5%的精度。

Weirong Jiang等<sup>[28]</sup>在研究多媒体流量实时分类时也只考虑了数据包级的特征。文中指出每个数据包有数十种可观察的特征,其中大部分包含在包头中,并最终选择IP协议、包大小以及TCP/UDP端口号作为流量特征对VoIP、即时通讯(IM)、IPTV等多媒体流量进行识别、分类,达到了99%的准确率。

#### 3.1.2 基于网络流的流特征

在工具软件Natemate<sup>[29]</sup>中,定义了40种网络流量属性特征,阳爱民教授在文献[30]中基本上沿袭了这些流量属性。Moore等在文献[31]中提出了249种流量属性(包括应用类型),但其中很大一部分是通过傅里叶变换得来的。

实时流量分类技术研究致力于发现一种在线识别、分类网络流量应用类型的方法。上述的流量特征将不能完全适用于实时流量分类,因为某些特征是无法通过网络流前 $n$ 个数据包就能提取出的。表2列举了常用于实时流量分类的主要流量特征,其最大特点是可以随时通过在线计算得到,且时间、空间复杂度低。

表2 基于网络流的流量特征

流量特征	TCP/UDP	时间复杂度	空间复杂度
前(后)向数据包个数	TCP/UDP	$O(n)$	$O(1)$
前(后)向数据包总(最小、最大、平均)长度	TCP/UDP	$O(n)$	$O(1)$
前(后)向数据包平均(最小、最大)到达间隔	TCP/UDP	$O(n)$	$O(1)$
子流持续时间	TCP/UDP	$O(n)$	$O(1)$
前(后)向标识位(PSH、URG)个数	TCP	$O(n)$	$O(1)$

上表共包含了21个流量特征,但并不是在所有的流量分类方法中都会用到。Li Wei等人在文献[32]中分别利用网络流前5个和25个数据包的10项特征对WWW、FTP、Email等流量进行了分类,分类结果表明,后者的分类准确性能达到95%左右。文献[33]在25个候选流特征的基础上采用后向贪婪特征选择算法<sup>[34]</sup>选择总数据包个数、大小、平均包大小等11个流量特征,使用半监督学习方法对29种网络流量应用进行了分类,准确率达到了90%以上,同时具备较好的实时性。张剑等人<sup>[35]</sup>采用FilterSubset算法从38个流量属性特征中选取了最佳的4个特征,并通过聚类获得了约94%的全局准确度。

网络流起始位置的数据包处于网络连接的协商阶段,含有丰富的特征信息。通过理论分析和实验观察,Sun等人<sup>[36]</sup>将连接请求者在接收到第1个ACK前接收到的总的长度定义为ACK-len<sub>ab</sub>,并将为连接响应者在接收到第1个ACK前接收到的总数据长度定义ACK-len<sub>ba</sub>。在实验中,利用C4.5决策树算法,仅通过上述两个属性进行分类,准确率达到了99%,并取得了良好的实时效果。

Yang等人<sup>[37]</sup>指出在BT连接中的MSE(Message Stream Encryption)握手阶段,其前3条信息在长度上有易区分的特点,但是这些信息被分成了多个有效载荷为100 bytes左右的数据包,需重组这些数据包。首先,按照四元组将TCP数据包形成网络流;随后,只保留网络流中含有标志位PSH的数据包,并将同向数据包合成一条信息,其长度为同向数据包之和,若有反向包,则为另一条信息;最后,对前3条信息按照方向赋予其正负值获得信息流,将该信息流量作为流特征来判断其是否是BT流。该方法与基于载荷的流量识别方法相结合,准确率能达到95%以上。

#### 3.1.3 流特征选取方法比较

总的来说,使用基于数据包的特征用于流量分类能够取得很好的分类效果,并且在时间、空间复杂度上具有较大的优势,可方便、高效地获取相关流量特征,尤其适用于实时流量分类。然而其最大的不足在于特征选取相对单一,其中数据包大小和到达间隔是非常重要的特征指标,即严重依赖于数据包到达顺序,又非常容易被伪造,面对恶意流量时该方法束手无策。

与基于数据包的特征相比,基于网络流的统计特征包含的可选择特征更多,特征信息也更加丰富,也更能反映出网

络流的统计特性,因此利用其进行实时分类能够取得更高的准确性,而且稳定性也要好于前者。但是基于网络流的统计特征在时间、空间复杂度上明显要大于前者。不过 Sun 和 Yang 的研究给了我们新的启示,亦即可通过寻找新的、更简单有效的流特征来满足更高的实时性要求。上述两类流特征选取方法比较如表 3 所列。

表 3 流特征选取方法比较

流特征选取方法	研究对象	流特征	时空复杂度	分类准确性
基于数据包的流特征选取方法	数据包 (packets)	数据包的大小、到达间隔、速率等	较低	一般
基于网络流的流特征选取方法	网络流量 (flow)	流大小、持续时间等	较高	高

### 3.2 机器学习算法

在实时流量分类技术中,机器学习算法用于训练、构建分类模型并根据在线获取的流特征对网络流进行分类,其执行分类的速度与准确性直接影响到实时流量分类方法的性能。对于机器学习算法,按照用于构建流量分类模型的样本是否需要预先标记,可将流量分类方法分为非监督(聚类)方法、监督方法和半监督方法。

#### 3.2.1 非监督学习

非监督流量分类方法主要包括两个步骤:聚类和标识。聚类,通过非监督机器学习(聚类)算法将流样本映射到相似度最高的簇当中;标识,对映射到该簇的流进行应用类型标记。

##### 1) 基于 K-Means 的实时流量聚类方法

K-Means 算法<sup>[38]</sup>比较简单,即在选择  $K$  个初始质心后将每个样本点指派到最近的质心,形成  $K$  个簇,并重新计算每个簇的质心,然后不断迭代此过程,直到质心不再变化,最后使得簇内相似度高、簇间相似度低。

Bernaille 等人在文献<sup>[17]</sup>中提出使用 K-Means 算法对 edonkey、http、ftp 等基于 TCP 的应用流量进行实时分类。将网络流前  $P$  个数据包的大小记为  $P$  维向量用于表示该流,同时为了表述出数据包的方向性,赋予了包大小的正负值,其值为正表示该包由服务器发往客服端,反之为负。在离线学习阶段,先利用 K-Means 算法建立分类模型,将网络流归入欧式距离最小的簇中,随后采用载荷分析工具<sup>[39]</sup>确定网络流的应用类型,对不同簇里的流进行应用类型标记;在在线分类阶段,在线捕捉网络流并形成流特征进行分簇映射,根据该簇中大多数流的应用类型对新流进行标记。结果表明,使用 TCP 流前 5 个数据包大小对网络流量进行分类可以达到 80% 的准确率。

##### 2) 基于 GMM 的加密流量聚类方法

高斯混合模型(GMM)就是利用高斯概率密度函数(正态分布曲线)准确地量化事物,将一个事物分解为若干个基于高斯概率密度函数形成的模型。

针对当前较多类型的网络应用采用载荷加密的方法来规避检测的情况,文献<sup>[40]</sup>提出了使用网络流前 3 个数据包大小即可实现加密流量实时检测的方法。该方法首先利用 GMM 建立 SSL 流量分类模型,通过计算网络流量属于某个簇的概率来将其映射到最大概率簇,并对其进行 SSL 标记;随后分析、检测 SSL 流量的第一个应用数据包,结果表明,

SSL 应用数据开始于第 3 到第 12 个数据包中的任意数据包;最后利用第一个应用数据包开始的 3 个数据包的大小及方向进行加密流量具体类型识别。对 SSL 流量进行标识,能达到较高的准确性,其中 SSLv2 和 SSLv3 的识别率达到 80% 以上。

#### 3) 基于 DBSCAN 的在线流量聚类方法

DBSCAN (Density-Based Spatial Clustering of Applications with Noise)<sup>[41]</sup>与划分和层次聚类方法不同,它将簇定义为密度相连的点的最大集合,能够把具有足够高密度的区域划分为簇,并可在噪声的空间数据库中发现任意形状的聚类。

张剑等人<sup>[35]</sup>提出一种基于 DBSCAN 算法的在线流量识别方案。该方案以网络数据流的若干初始数据包作为子流,提取子流的统计特征,应用基于滤波器算法的属性相关性算法提取子流最佳特征子集,并提出基于密度的在线带噪声空间聚类算法对子流特征向量进行聚类,采用优势概率业务实现聚类和应用类型的映射。实验结果表明,该方案具备识别新应用类型和加密数据流的功能,且能实现在线的网络流量分类,分类准确率达到 90% 以上。

#### 4) 流标记方法

文献<sup>[18]</sup>介绍了两种标记方法:优势法(Dominant)和聚类+端口方法(Cluster & Port)。

虽然一个簇中可能包含多种应用类型的流,但是大多数簇只有一个应用类型。优势法即将新加入的网络流标记为该簇中占最大优势的应用类型。此方法虽然简单,但是当簇中含有不止一个应用类型时非常容易将不占优势的网络应用标记错误。

通过端口号虽然不能识别所有网络应用,但很多应用依赖于端口号,尤其对于客户/服务器类型的应用。聚类+端口方法定义了标准的客户/服务器应用端口号: $S = \{21, 22, 25, 80, 110, 119, 443, 995\}$ ,以及与该端口相对应的标准应用类型: $std(S) = \{FTP, SSH, SMTP, HTTP, POP3, NNTP, HT-TPS, POP3S\}$ 。该方法描述如下:如果网络流  $x$  使用了标准端口号,且所对应的应用类型属于该簇,那么将流  $x$  标记为该应用类型;否则将该流标记为伪装(masquerade)。如果网络流的端口号是非标准的且簇所包含的应用均为非标准的,则用优势法对该流进行标记;否则将该流标记为伪装。实验结果表明,聚类+端口方法的准确性优于优势法。

#### 3.2.2 监督学习

基于监督机器学习算法的流量分类主要包括数据预处理、训练分类模型和基于分类模型的流量分类。

##### 1) 基于 C4.5 的实时流量分类方法

在继承决策树基本构造思想的基础上,C4.5<sup>[42]</sup>算法用信息增益率作为选择分裂属性的标准。C4.5 能够较好地处理连续型属性及缺失数据,并引入了树剪枝技术,在精简决策树的同时提高了分类的准确率。

Jun Li 等人<sup>[43]</sup>使用决策树算法(C4.5、REPTree)对 P2P 流量进行实时识别,分类准确率达到 95% 左右,其准确性远高于文献<sup>[17]</sup>中基于 K-Means 的流量识别方法。实验表明,基于决策树的流量分类模型建模时间仅需要 0.2s,处理  $10^6$  个流样本仅需要 1s;该方法具备较强的健壮性和灵活性,可根据 P2P 流量共同的行为特征、传输特性等识别未知的 P2P

流量。文献[36]将定义的 ACK-len ab、ACK-len ba 作为流特征,利用 C4.5 决策树算法对 Moore 数据集和自采数据集中的 WWW、Mail、Ftp、P2P 等流量进行实时流量分类。对于 Moore 数据集,分类准确率达到 99%;对于后者,分类准确率也达到了 97%。

### 2) 基于 Naïve Bayes 的实时流量分类方法

Naïve Bayes(朴素贝叶斯分类方法)<sup>[44]</sup>是一种利用贝叶斯定理对属性集合和类型变量之间概率关系进行建模的方法。该方法因计算高效、精确度高并具有坚实的理论基础而得到广泛的应用。

T. Nguyen 等<sup>[19]</sup>提出了基于网络流最近的  $N$  个数据包(滑动窗口)的实时流量分类方法,该方法不需要捕捉流的开始,使得在流生命周期任意时刻识别其应用类型成为可能。从网络流的不同位置选取两条或两条以上子流(每个子流包含连续的  $N$  个数据包)来代表该流,提取数据包到达时间、包大小等相关子流特征,并通过 Naïve Bayes 训练流量分类模型。实验结果表明,基于完整流的分类器性能不如基于多个子流的,即使用网络流中段的 25 个数据包作为子流来进行分类,也可以达到 95% 的召回率和 98% 的准确率。文献[32]提出了一个基于 Naïve Bayes 的可实现准实时分类的结构框架——ANTc。ANTc 包括数据包捕获、流形成、流特征提取和分类 4 个模块,并利用 C 语言通过 JNI<sup>[46]</sup>调用 WEKA<sup>[47]</sup>实现。他们的实验结果表明,ANTc 不但能够实现实时流量分类,而且进行在线实时流量分类的准确性和离线分类的准确性相差无几。

### 3) 基于 SVM 的实时流量分类方法

支持向量机(Support Vector Machine, SVM)<sup>[45]</sup>以结构风险最小化为原则,在 VC 维的意义下将原始的样本集稀疏化为小规模的支持向量集。与其他学习机相比,支持向量机具有良好的推广能力,在处理非线性识别和小样本学习方面具有良好的特性。

Zhu Li 等人<sup>[48]</sup>选择优化 SVM(支持向量机)对包含 7 类应用类型的 10000 个网络流建立分类模型并进行分类实验。通过实验发现,选择不同数量的流特征只会造成建模时间的不同(当选择 19 个流量特征时建模时间是 60s,当选择 9 个流量特征时建模时间是 30s),而不会造成分类准确性的下降(二者都达到了 99%),并且每分钟能处理 60000~80000 条网络流,能够满足实时性要求。虽然都是采用 SVM 建立流量分类模型,但与文献[48]不同的是,文献[24]选用前 5 或 6 个数据包的大小及方向作为流量特征,致使后者的准确性明显低于前者,仅为 80% 左右。

文献[49]在研究加密流量实时分类时对上述 3 种方法进行了比较,实验结果表明 C4.5 较其它机器学习方法能达到更高的分类准确率,在分类准确性上  $C4.5 > SVM > Naïve Bayes$ 。Tavallaee 等在文献[52]中对几种分类算法的准确性进行比较的同时还对分类模型训练的时间进行了比较,结果显示 Naïve Bayes  $> C4.5$ 。

### 3.2.3 半监督学习

虽然相对于非监督学习流量分类方法,监督学习方法的优劣不言而喻,但仍有两大问题困扰着监督学习方法:首先,获取并标记大量的具有代表性的流量样本是非常耗时的,其代价也十分昂贵;其次,训练集中可能无法囊括所有类型的流

量,以至于无法识别出未知网络流量。

为了解决上述问题,Erman 等<sup>[50]</sup>首次提出了半监督学习的网络流量分类方法。该方法利用 K-Means 算法将少量标记过的流样本和大量未标记的流样本按照相似性聚类并形成簇,随后通过簇中标记过的流样本获得该簇与已知网络应用类型的映射关系。作者使用了最大似然估计  $\frac{n_{jk}}{n_k}$  来寻找标记流到簇的映射关系,其中  $n_{jk}$  表示被分到簇  $k$  的标记为  $j$  的流样本个数,  $n_k$  表示被分到簇  $k$  的标记过的流样本总数。不存在映射关系的簇可以标记为未知,用以发现新的未知流量。实验结果表明,即便用于实验的未标记流样本为 64000 个,只需要设置 400 个簇且每个簇含有两个标记流样本就能达到 94% 的流准确率。

文献[52]将基于载荷和机器学习流量分类方法的优点结合到一起,提出了混合实时流量分类方法。该方法先利用基于载荷的流量分析软件 MeterFlow<sup>[53]</sup>识别网络流量,对于未能识别的流量使用基于机器学习的流量分类方法处理。为了增强可靠性,Tavallaee 等定义了为期一天的训练间隔,即在一个训练间隔后利用已识别的、确定的网络流训练基于机器学习的流量分类模型,然后利用该模型对基于载荷方法不能识别的流量进行分类。实验显示,基于 C4.5 的混合流量分类方法的分类准确率最高,达到 99.27%,且模型训练时间也相对较短,仅为 43.38s。

### 3.2.4 3 种方法的比较

非监督机器学习方法的优点在于无需预先对流样本进行应用类型标记,而且能以较大的概率发现新的网络应用,但缺点也较为明显:聚类形成的“簇”与流量“类型”并不存在一一对应的关系,需要进一步标记“簇”中的流;不同的非监督机器学习算法,甚至同一算法在选择不同的参数或面对不同网络环境时都可能对结果产生较大影响,这导致该类方法健壮性较差,可扩展、移植能力不佳,训练时间较监督学习方法更长,且分类准确性也不如后者。因此,在目前已有的实时流量分类技术研究中,监督学习方法占了主导地位,尤其是决策树 C4.5,无论是在模型训练时间上还是分类准确性上都非常适用于实时流量分类研究。但随着研究的深入,监督学习方法的缺点已变得不可忽视。建立分类模型需要对流样本进行应用类型标记,这是一项非常困难的工作,大大限制了监督学习方法的进一步应用,同时不断变化的网络环境使得该方法难以直接用于实际应用。半监督(混合)机器学习方法可较好地综合二者的优点,在只需少量标记样本或无需样本的情况下能以较高的准确性完成流量分类,但不足之处在于其实时性受到影响。目前,基于半监督(混合)机器学习算法的流量分类技术成为学术界研究的热点之一。3 种方法的比较如表 4 所列。

表 4 3 种机器学习方法的比较

机器学习方法	样本是否需要标记	训练时间	准确性	实时性	能否发现未知流量
非监督机器学习方法	否	长	一般	高	能
监督机器学习方法	需要	短	高	高	否
半监督(混合)机器学习方法	少量标记或不需	一般	高	一般	能

### 3.3 实时性分析

根据第2节的实时性指标建模、分析,可以发现流量特征的选取直接决定了流量分类方法的实时性。在此,本文利用实时性指标 $(n, m)$ 对近年来具有代表性的研究成果进行实时性分析、比较。

通过比较可以发现,基于数据包级流量特征的分类方法在实时性上优于基于网络流级流量特征的分类方法。因为前者只需要采集前几个数据包的大小及方向即可(如文献[17, 18, 23],其实时性分别为 $(5, 5)$ 、 $(4, 4)$ 和 $(3, 3)$ ),而后者则需要采集前 $n$ 个数据包的 $m$ 个流量统计信息(如文献[32]的实时性为 $(25, 10)$ ,文献[49]的 $(12, 10)$ ),显然 $(5, 5) < (25, 10)$ ,即前者的实时性优于后者。

对于同是基于数据包级流量特征的分类方法而言,文献[17]的实时性不如文献[18];对于基于数据包级流特征的流量分类方法而言,文献[49]优于文献[32]。

同时还发现,半监督学习方法<sup>[33, 51, 52]</sup>虽然不需在样本预处理上花费过多精力,且准确性较高,能达到90%以上,但实时性受到影响,不如监督学习方法和非监督学习方法,如文献[51]的实时性为 $(100, 17)$ ,明显差于上述文献[17, 32]的实时性。

当然,并不是实时性越好,分类方法的性能就越好,实时性与准确性是相互影响的。通常情况下,实时性越好,即数据包和流量特征个数越少,其对流量特点的描述就越不准确,其分类的准确性就越差,如基于数据包级流量特征的分类方法的分类准确性就明显低于基于流级流量特征的分类方法。

利用上述指标进行实时性比较时也发现一些不足,该指标只能对一般方法的实时性进行比较,在面对诸如 $(5, 20)$ <sup>[43]</sup>、 $(25, 10)$ <sup>[32]</sup>之间的实时性比较时就显得无能为力。为更准确地描述流量分类的实时性,拟通过分析数据包达到间隔 $t_0$ 与单个流特征处理时间 $t_1$ 之间的大致关系,进一步研究实时流量分类技术的实时性评价指标。

### 4 下一步研究工作

近年来,诸如P2P、网络视频、在线游戏等新(非传统)网络应用的不断涌现,使得网络流量骤增,尤其是在高峰时期,网络运行不堪重负;随着下一代网络的部署,网络速度将急速提升,流量空前巨大。这对实时流量分类技术提出了新的要求,主要表现在:

(1)实时流量分类技术应兼具高分类准确性和高实时性,以满足网络流量管理、控制的需求;

(2)应具备更快的网络流量处理速度,以适应高速网络的发展趋势;

(3)应具备更强的网络环境适应能力,以应对不断变化的网络环境;等等。

鉴于网络发展对实时流量分类技术提出的新要求以及实时流量分类技术的研究现状,下一步研究工作主要包括以下方面:

(1)研究新的半监督(混合)机器学习方法,在保证流量分类准确性的条件下,进一步提高实时性。半监督(混合)机器学习方法虽然能较好地弥补监督学习方法在样本标记方面的不足,且保持较高的准确性,但实时性会受到影响,如何提高其实时性是混合机器学习方法需要解决的问题。

(2)通过研究新的方法来解决流量分类面临的“概念漂移”问题(Concept Drift)。所谓概念漂移,即是在时刻 $t$ 得到的最佳分类模型 $y_t$ 与前一时刻 $t_1$ 得到的最佳分类模型 $y_{t_1}$ 不一致。分类模型是通过机器学习算法对流特征向量进行训练得到的,因此,为了解决概念漂移问题,可以从两点出发:第一,研究新的机器学习算法,如Hulten等<sup>[54]</sup>提出的CVFDT就能很好地解决因为概念漂移导致的预测、分类性能下降问题;第二,研究更加稳定的流特征,该特征不随网络环境的变化而变化或变化较小。

(3)研究高速网络环境下的实时流量分类方法。高速是未来网络的发展趋势,但就当前实时流量分类的研究现状来说,主要面临两个难题。首先,如何利用非常稀缺的网络资源在高速网络中获取海量数据流;其次,面对如此海量的流量数据,如何以更快的处理速度进行实时分类。

(4)研究基于硬件的实时流量分类方法。硬件在处理速度上的优势是软件所无法比拟的,目前国内外已有学者开展了这方面的研究。Weirong Jiang等<sup>[28, 55]</sup>提出了基于FPGA的网络流量分类方法,桂林电子科技大学的王勇教授也申请了国家自然科学基金项目《基于FPGA的分级并行高速网络流量分类方法研究》。

(5)关于实时流量分类的研究从实验走向实际应用。目前实时流量分类研究大都停留在实验阶段,尚无如Meter-Flow<sup>[15]</sup>、Tstat2.0<sup>[56]</sup>等实用的流量分类工具。研发实时流量分类工具既有助于网络流量分类研究的进一步发展,又可推动在实时流量分类基础上的流量控制、管理研究,具有重要的理论意义和实用价值。

**结束语** 实时流量分类技术可广泛地应用于流量控制、异常检测、实时态势分析等实际网络管控,是流量分类技术研究的一个必然方向。针对这一研究领域的发展趋势,本文在介绍当前主要流量分类技术的基础上综述了2006年至2012年以来国内外关于实时网络流量识别、分类技术的研究进展,并结合未来网络发展对实时流量分类技术的要求,给出了该领域下一步的研究发展方向。为了更准确、直观地描述、比较实时性,提出了流量分类的实时性概念以及实时性评价指标,并对当前的实时流量分类方法进行了实时性分析。

关于流量分类技术的研究看似已经相当成熟,部分方法甚至达到99%的分类准确性,其实不然,关于实时性的研究才刚刚起步,离实用还有相当的距离。随着技术的不断进步以及研究的进一步深入,新的流量分类方法必将继续涌现。

### 参考文献

- [1] Internet Assigned Numbers Authority (IANA). <http://www.iana.org/assignments/port-numbers>, August 28, 2010
- [2] Moore A W, Papagiannaki D. Toward the accurate Identification of network applications[C]//Proc. 6th Passive Active Measurement Workshop(PAM), 2005, 3431: 41-54
- [3] Sen S, Spatscheck O, Wand D. Accurate, scalable in-network identification of P2P traffic using application signatures[C]//Proceedings of the 13th International World Wide Web Conference on Alternate Track Papers & Posters(WWW'04). New York, NY, USA, ACM, 2004: 512-521
- [4] Karagiannis T, Papagiannaki K, Faloutsos M. BLINC: Multilevel traffic classification in the dark [C]//ACM SIGCOMM. Phila-

- [5] Roughan M, Sen S, Spatscheck O, et al. Class-of-service mapping for QoS: A statistical signature-based approach to IP traffic classification [C]//Proceedings of ACM SIGCOMM Internet Measurement Conference, Taormina, Sicily, Italy, 2004
- [6] Moore A W, Zuev D. Internet Traffic Classification Using Bayesian Analysis Techniques [C]//Proceedings of ACM SIGMETRICS International Conference on Measurement and Modeling of Computer Systems, New York, USA, 2005
- [7] 徐鹏, 刘琼, 林森. 基于支持向量机的 Internet 流量分类研究[J]. 计算机研究与发展, 2009, 46(3): 407-414
- [8] 徐鹏, 林森. 基于 C4. 5 决策树的流量分类方法[J]. 软件学报, 2009, 10(20): 2692-2704
- [9] McGregor A, Hall M, Lorier P, et al. Flow Clustering Using Machine Learning Techniques [C]//Proceedings of PAM'04, Antibes Juan-les-Pins, France, 2004
- [10] Zander S, Nguyen T, Armitage G. Self-Learning IP Traffic Classification Based on Statistical Flow Characteristics [C]//Proceedings of PAM'05, Boston, USA, 2005
- [11] Erman J, Arlitt M, Mahanti A. Traffic Classification Using Clustering Algorithms [C]//Proceedings of SIGCOMM Workshop on Mining Network Data, Pisa, Italy, 2006
- [12] Dainotti A, Pescapé A, Sansone C. Issues and Future Directions in Traffic Classification[J]. IEEE Network, 2012, 26(1): 35-40
- [13] Gu Cheng-jie, Zhang Shun-yi, Sun Yan-fei. Real-time Encrypted Traffic Identification using Machine Learning [J]. Journal of Software, 2011, 6(6): 1009-1016
- [14] Hurley J, Garci-Palacios E, Sezer S. Classifying network protocols: A 'two-way' flow approach[J]. Communications, 2012, 5(1): 19-89
- [15] Hifn company[OL]. <http://www.hifn.com>, September, 2008
- [16] Karagiannis T, Broido A, Brownlee N. Is P2P Dying or Just Hiding? [C]//GLOBECOM '04, Dallas, USA, November 2004
- [17] Bernaille L, Teixeira R, Akodkenou I, et al. Traffic classification on the fly[J]. ACM Special Interest Group on Data Communication (SIGCOMM) Computer Communication Review, 2006, 36(2)
- [18] Bernaille L, Teixeira R, Salamati K. Early Application Identification [C]//Proceedings of Conference on Future Networking Technologies 2006 (CoNEXT 06), Lisbon, Portugal, 2006
- [19] Nguyen T, Armitage G. Training on multiple sub-flows to optimise the use of Machine Learning classifiers in real-world IP networks[C]//Proc. IEEE 31st Conference on Local Computer Networks, Tampa, Florida, USA, November 2006
- [20] Erman J, Mahanti A, Arlitt M F. Byte me: A case for byte accuracy in traffic classification[C]//Sen S, Sahu S, eds. Proc. of the 3rd Annual ACM Workshop on Mining Network Data (MineNet 2007). New York: ACM Press, 2007: 35-37
- [21] Mohammed A, Nor S M. Near Real Time Online Flow-based Internet Traffic Classification Using Machine Learning (C4. 5) [J]. International Journal of Engineering (IJE), 3(4): 370-379
- [22] 鲁刚, 张宏莉, 叶麟. P2P 流量识别[J]. 软件学报, 2011, 22(6): 1281-1298
- [23] Bernaille L, Teixeira R. Early Recognition of Encrypted Applications[J]. Lecture Notes in Computer Science, 2007, 4427: 165-175
- [24] Sena G, Belzarena G P. Early traffic classification using support vector machines[C]//Proceedings of the 5th International Latin American Networking Conference, LANC 09, 2009: 60-66
- [25] 彭建芬, 周亚建, 王枫, 等. TCP 流量早期识别方法[J]. 应用科学学报, 2011, 29(1): 73-77
- [26] Branch P, Heyde A, Armitage G. Rapid Identification of Skype Traffic Flows [C]//NOSSDAV '09, Williamsburg, Virginia, USA, 2009: 3-5
- [27] But J, Branch P, Le T. Rapid Identification of BitTorrent Traffic [C]//Local Computer Networks (LCN), 2010: 536-543
- [28] Jiang Wei-rong, Gokhale M. Real-Time Classification of Multimedia Traffic using FPGA [C]//Proceedings of the 2010 International Conference on Field Programmable Logic and Applications, FPL'10, 2010: 56-63
- [29] NetMate [OL]. <http://sourceforge.net/projects/netmate-meter/>, 2007
- [30] 阳爱民, 周咏梅, 邓河, 等. 一种网络流量分类特征的产生及选择方法[J]. 山东大学学报: 工学版, 2010, 40(5): 1-7
- [31] Moore A W, Zuev D, Crogan M. Discriminators for use in flow-based classification, RR-05-13 [R]. London: Queen Mary University of London, 2005
- [32] Wei Li, Abidin K, Dann R, et al. Approaching real-time network traffic classification [R]. Technical Report RR-06-12, Department of Computer Science, Queen Mary, University of London, December 2006
- [33] Erman J, Mahanti A, Arlitt M, et al. Offline/Realtime traffic classification using semi-supervised learning[J]. Performance Evaluation, 2007, 64(9-12): 1194-1213
- [34] Guyon I, Elisseeff A. An Introduction to Variable and Feature Selection[J]. Journal of Machine Learning Research, 2003, 3: 1157-1182
- [35] 张剑, 钱宗珏, 寿国础, 等. 在线聚类的网络流量识别[J]. 北京邮电大学学报, 2011, 34(1): 103-106
- [36] Sun Mei-feng, Chen Jing-tao. Research of the traffic characteristics for the real time online traffic classification[J]. The Journal of China Universities of Posts and Telecommunications, 2011, 18(3): 92-98
- [37] Yang Zhe, Li Ling-zhi, Ji Qi-jin, et al. Cocktail method for BitTorrent traffic identification in real time[J]. Journal of Computers, 2012, 7(1): 85-95
- [38] Tan Pang-ning, Steinbach M, Kumar V. Introduction to Data Mining[M]. Posts & Telecom Press, 2011
- [39] Qosmos[OL]. [www.qosmos.com](http://www.qosmos.com)
- [40] Bernaille L, Teixeira R. Early Recognition of Encrypted Applications[J]. Lecture Notes in Computer Science, 2007, 4427: 165-175
- [41] Ester, Martin, Kriegel H P, et al. A Density Based Algorithm for Discovering Clusters in Large Spatial Data bases with Noise [C]//Proceedings of the 2nd International Conference on Knowledge Discovery and Data Mining (KDD-96), Portland, Oregon: [s. n.], 1996
- [42] Quinlan R J. C4. 5: Programs for machine learning [M]. San Mateo, Calif: Morgan Kaufman, 1993
- [43] Li Jun, Zhang Shun-yi, Lu Yan-qing, et al. Real-time P2P Traffic Identification [C]//Global Telecommunications Conference, 2008: 1-5



- [44] Witten I, Frank E. Data Mining: Practical Machine Learning Tools and Techniques with Java Implementations(Second Edition)[M]. Morgan Kaufmann Publishers, 2005
- [45] Vapnik V N. Statistical learning theory [M]. New York: Wiley, 1998
- [46] JNI[OL]. <http://java.sun.com>
- [47] Waikato Environment for Knowledge Analysis(WEKA)[OL]. <http://www.cs.waikato.ac.nz/ml/weka/>
- [48] Li Zhu, Yuan Rui-xi, Guan Xiao-hong. Traffic Classification-Towards Accurate Real Time Network Applications[C]// Human-Computer Interaction, Part IV, HCI 2007. LNCS 4553, 2007: 67-76
- [49] Gu Cheng-jie, Zhang Shun-yi, Sun Yan-fei. Real-time Encrypted Traffic Identification using Machine Learning[J]. Journal of Software, 2011, 6(6): 1009-1016
- [50] Erman J, Mahanti A, Arlitt M, et al. Offline/realtime network traffic classification using semisupervised learning[R]. Technical Report, Department of Computer Science, University of Calgary, February 2007
- [51] Bar-Yanai R, Langberg M, Peleg D, et al. Realtime Classification for Encrypted Traffic[C]// SEA 2010. LNCS 6049, 2010: 373-385
- [52] Tavallaei M, Lu W, Ghorbani A A. Online classification of network flows[C]// Proceedings of the 7th Annual Communication Networks and Services Research Conference. CNSR 2009, 2009: 78-85
- [53] Hifn company[OL]. <http://www.hifn.com>, September, 2008
- [54] Spencer H L, Domingos P. Mining timechanging data streams[C]// 7th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. ACM Press, 2001: 97-106
- [55] 王建东, 祝超, 谢应科, 等. 基于 FPGA 的万兆流量并行实时处理系统研究[J]. 计算机研究与发展, 2009, 46(2): 177-185
- [56] Tstat[OL]. <http://tstat.tlc.polito.it/index.shtml>
- 
- (上接第 7 页)
- [16] Ahn J, Hong S, Heidemann J. An adaptive FEC code control algorithm for mobile wireless sensor networks[J]. Journal of Communications and Networks, 2005, 7(4): 489-499
- [17] Srouji M S, Wang Z, et al. RDTS: A Reliable Erasure-Coding Based Data Transfer Scheme for Wireless Sensor Networks[C]// Proceedings of 17th IEEE International Conference on Parallel and Distributed Systems. 2011
- [18] Ganesan D, Ramesh G, Scott S, et al. Highly-resilient, energy-efficient multipath routing in wireless sensor networks[J]. ACM SIGMOBILE Mobile Computing and Communications Review, 2001, 5(4): 11-25
- [19] Deb B, Bhatnagar S, Nath B. ReInForM: Reliable information forwarding using multiple paths in sensor networks[C]// Proceedings of 28th Annual IEEE International Conference on Local Computer Networks. 2003
- [20] Felemban E, Lee C G, Ekici E. MMSPEED: multipath Multi-SPEED protocol for QoS guarantee of reliability and Timeliness in wireless sensor networks[J]. IEEE Transactions on Mobile Computing, 2006, 5(6): 738-754
- [21] Yahya B, Ben-Othman J. REER: Robust and energy efficient multipath routing protocol for wireless sensor networks[C]// Proceedings of IEEE Global Telecommunications Conference. 2009
- [22] Intanagonwiwat C, Govindan R, Estrin D. Directed Diffusion: A Scalable and Robust Communication Paradigm for Sensor Networking[C]// Proceedings of ACM MobiCom'00. 2000: 56-67
- [23] Lou W. An efficient N-to-1 multipath routing protocol in wireless sensor networks[C]// Proceedings of IEEE international Conference on Mobile Adhoc and Sensor Systems (MASS). Washington, DC, November 2005
- [24] Vuran M C, Akyildiz I F. Error control in wireless sensor networks: a cross layer analysis[J]. IEEE/ACM Transactions on Networking, 2009, 17(4): 1186-1199
- [25] Wen H, Lin C, Ren F Y, et al. Retransmission or redundancy: Transmission reliability study in wireless sensor networks[J]. Science China Information Sciences, 2012, 55(4): 1-10
- [26] Yahya B, Ben-Othman J. Relax: an energy efficient multipath routing protocol for wireless sensor networks[C]// Proceedings of IEEE International Conference on Communications. 2010
- [27] Courtade T, Wesel R. Optimal allocation of redundancy between packet-level erasure coding and physical-layer channel coding in fading channels[J]. IEEE Transactions on Communications, 2011, 59(8): 1-9
- [28] Kwon H, et al. A cross-layer strategy for energy-efficient reliable delivery in wireless sensor networks[J]. IEEE Transactions on Wireless Communications, 2006, 5(12): 3689-3699
- [29] Ahlswede R, et al. Network information flow[J]. IEEE Transactions on Information Theory, 2000, 46(4): 1204-1216
- [30] Ghaderi M, Towsley D, Kurose J. Reliability gain of network coding in lossy wireless networks[C]// Proceedings of the 27th IEEE Conference on Computer Communications. 2008
- [31] Chi K, Jiang X, Horiguchi S. Network coding-based reliable multicast in wireless networks[J]. Computer Networks, 2010, 54(11): 1823-1836
- [32] Cai N, Yeung R W. Network coding and error correction[C]// Proceedings the IEEE Information Theory Workshop. 2002
- [33] Koetter R, Kschischang F R. Coding for errors and erasures in random network coding[J]. IEEE Transactions on Information Theory, 2008, 54(8): 3579-3591
- [34] 张书奎, 樊建席, 崔志明. 无线传感器网络中可靠的数据协作传输机制[J]. 通信学报, 2010, 31(11): 30-40
- [35] Yu Rong, Zhang Yan, Song Ling-yang, et al. Joint Optimization of Power, Packet Forwarding and Reliability in MIMO Wireless Sensor Networks[J]. Mobile Networks and Applications, 2011, 16(6): 1-11
- [36] Stojmenovic I, Lin X. Power-aware localized routing in wireless networks[J]. IEEE Transactions on Parallel and Distributed Systems, 2001, 12(11): 1122-1133
- [37] Gomez J, Campbell A T, Naghshineh M, et al. Conserving transmission power in wireless ad hoc networks[C]// Network Protocols, 2001. Ninth International Conference on. IEEE, 2001: 24-34