

# 一种基于本体的并行网络流量分类方法

陶晓玲<sup>1,2</sup>, 韦毅<sup>2</sup>, 王勇<sup>2,3</sup>

(1. 桂林电子科技大学认知无线电与信息处理省部级共建教育部重点实验室 广西 桂林 541004;

2. 桂林电子科技大学广西高校云计算与复杂系统重点实验室 广西 桂林 541004;

3. 桂林电子科技大学广西可信软件重点实验室 广西 桂林 541004)

**【摘要】**海量网络流量数据的处理与单一节点的计算能力瓶颈这一矛盾导致数据分类效率低,无法满足现实需求。为解决这一问题,结合本体与MapReduce技术各自在海量异构数据描述与处理方面的优势,提出一种基于本体的并行网络流量分类方法。该方法基于MapReduce并行计算架构,根据网络流量本体结构,对网络流量本体并行化构建;通过并行知识推理完成基于流量统计特征的网络流量分类。实验结果表明,集群环境下基于MapReduce的网络流量本体构建效率明显高于单机环境,而且适当增加计算节点使得加速比线性提升;并行知识推理的分类方法能够有效地提高大规模网络流量的分类效率。

**关键词** 知识推理; MapReduce; 网络流量分类; 本体; 并行化

**中图分类号** TP393 **文献标志码** A **doi**:10.3969/j.issn.1001-0548.2016.02.0018

## An Ontology Based Parallel Network Traffic Classification Method

TAO Xiao-ling<sup>1,2</sup>, WEI Yi<sup>2</sup>, and WANG Yong<sup>2,3</sup>

(1. Key Laboratory of Cognitive Radio and Information Processing, Guilin University of Electronic Technology Guilin Guangxi 541004;

2. Guangxi Colleges and Universities Key Laboratory of Cloud Computing and Complex Systems,

Guilin University of Electronic Technology Guilin Guangxi 541004;

3. Guangxi Key Laboratory of Trusted Software, Guilin University of Electronic Technology Guilin Guangxi 541004)

**Abstract** The contradiction between the processing of mass network traffic data and the computing bottleneck of a single node leads to low efficiency of data classification. To address this challenge, we propose an ontology based parallel network traffic classification method by integrating the advantage of ontology and MapReduce in dealing with the description and processing of mass heterogeneous data. Our approach makes use of MapReduce, a framework of parallel computing. Firstly, it uses the ontology to describe and manage network traffic data, and constructs the layered and parallel network traffic ontology. Then it builds the classification model by employing the decision tree algorithm, by which the inference rule set is generated. Network traffic classification based on traffic statistical features is completed by utilizing parallel knowledge reasoning. Implementation results show that data classification efficiency of the proposed approach in group environment is higher than in stand-alone scenario. The speedup ratio increases linearly when increasing the quantity of compute nodes. In addition, the new method is able to improve the classification efficiency of large-scale network traffic significantly.

**Key words** knowledge reasoning; MapReduce; network traffic classification; ontology; parallelization

作为管理和优化各类网络资源的关键技术,网络流量分类<sup>[1]</sup>广泛应用于网络监控、服务质量(quality of service, QoS)管理、网络安全、态势分析等领域,是高效实现网络管理、流量控制以及安全检测的重要环节。随着Web技术的发展和企业信息化需求的不断提高,许多新型网络应用模式和需求应运而生,网络环境也升级为高速、大规模、复杂

网络,随之而来的网络流量数据呈现出新的特点:海量(数量惊人、信息丰富)、多源(数据源分布在离散的,彼此可以通信的多个网络节点上)、异构(格式异构、语法异构、语义异构),致使网络流量分类面临严峻的挑战。

其一,各个网络节点传感器使用不同的流量采集系统收集网络数据包,在数据生成方式、存放方

收稿日期:2015-11-21; 修回日期:2016-03-15

基金项目:国家自然科学基金(61163058, 61363006); 广西可信软件重点实验室开放课题(KX201306); 广西高校云计算与复杂系统重点实验室开放课题(14104)

作者简介:陶晓玲(1977-),女,副研究员,主要从事云计算、网络安全及机器学习方面的研究。

式和处理方式上呈现多样化,出现网络流量数据格式不一,类型不同,且不同数据存在语义区分等问题。因此,多个数据源提供的异构网络流量数据之间会存在数据不一致性问题,从而影响网络流量分类结果的准确性。

其二,目前主流的网络流量分类方法是基于流量统计特征的机器学习方法<sup>[2]</sup>,因为此类方法需提取高维的流量统计特征,通过复杂的计算构造分类模型,面对海量的网络流量数据,处理时间开销较大,分类效率不高,这使得其不能满足高速网络流量分类的需求。

本体<sup>[3]</sup>在信息系统中被定义为一种能在语义和知识层次上描述特定知识领域的形式化技术,具有良好的概念层次结构,对逻辑推理无缝支持,为信息资源规范、无二义性和可扩展性描述问题提供了有效的解决途径。文献[4]提出采用本体作为网络流量信息资源的统一描述的思路。

并行处理技术MapReduce能够为可划分的大规模数据并行计算处理问题提供充分的并行计算语义,已经被普遍接受。该技术为提高网络流量分类中海量数据处理效率问题提供了新方法。因此,基于文献[4],本文借助并行处理技术MapReduce,提出一种基于本体的并行网络流量分类方法。该方法将发挥MapReduce在海量异构数据处理方面的优势,为本体的构建、知识管理及推理提供计算资源,用于海量网络流量数据并行处理、分类,为高速大规模复杂网络环境下的网络流量分类提供新思路 and 理论方法依据。

## 1 相关工作

### 1.1 本体及其应用

本体本质上是针对一个特定领域,对领域知识的标准化描述,以便相关领域学者在概念层面上达成共识,相互之间实现知识共享。本体作为一种知识表达的工具<sup>[5]</sup>逐渐成为国内外的研究热点。文献[6-8]主要集中在利用本体进行信息资源描述的模型一致性、逻辑一致性和关系一致性三方面。由此可见,本体为异构数据提供了统一的概念接口,并且独立于数据模式,可以对异构数据进行丰富的语义描述。

本体也被应用于决策支持系统中管理领域知识,并被许多推理机支持,用于实现知识推理<sup>[9]</sup>。目前,基于本体的知识推理功能也被应用于分类问题。文献[10]提出了一个基于本体的海洋卫星图像分

类模型,构建了基于决策树和专家定制规则的图像本体分类器,该模型的分类准确率达到92.49%;文献[11]将本体应用于轻度认知障碍(MCI)诊断,提出一种本体驱动的利用磁共振成像(MRI)自动诊断MCI的方法,并通过对比发现决策树算法更适用于构建推理规则集。

近年来,有少量研究者已尝试将本体应用于网络流的定义及分类模型的建立。文献[12]构建了一个基于本体范例的分类树,首次尝试利用本体对网络流量类别进行标准化定义。文献[13]设计了一个基于流轮廓和本体的在线、自学习网络流量分类模型,此模型比较复杂,文中也未给出该模型的具体实现。

### 1.2 海量数据并行处理技术

面向海量数据的处理问题,考虑到单一节点的计算能力瓶颈,研究者们纷纷采用分布式或并行处理的方式来解决此类问题,其中基于并行处理技术MapReduce的海量数据处理方面的工作颇有成效。文献[14-15]分别提出了一种基于MapReduce的分布式ELM学习模型及训练框架,实验证明,在大数据的学习及训练方面是有效的。

MapReduce技术也被研究者们用于解决海量数据的分类问题。文献[16]提出一种基于类别的集成技术用于分类概念漂移数据流,采用了基于MapReduce的技术提高分类方法的效率和鲁棒性。文献[17]指出MapReduce具有易于开发的可扩展性和容错性,进而提出了一种基于MapReduce的随机森林方法来处理非平衡大数据的分类问题。

复杂网络环境下,网络链路中不断增加的待处理数据与基于单机的计算机系统处理能力的矛盾日益突出,有少数的国内外学者也将MapReduce技术应用于网络流量的处理。文献[18]提出了一个基于Hadoop平台的网络流量分流并行处理结构,实验验证对于大数据进行分流时,该结构优势显著。文献[19]提出基于高斯混合模型-隐马尔可夫模型的网络流量分类方法,模型使用了两个数据包级属性来构建,实现了一个基于MapReduce的并行分类架构,并验证了模型具有灵活性。

## 2 基于MapReduce的网络流量本体构建方法

本文采用文献[4]设计的分层的网络流量本体结构。首层包括网络流量采集节点信息和流量信息。其中,流量采集节点的下一层记录流量采集节点的相关信息,包括网络软硬件设备信息、流量采集工

具及节点的配置参数; 流量采集工具的下一层记录各流量采集节点的各种流量采集工具名称以及流量采集信息格式。而流量信息的下一层记录网络流量的相关信息, 包括网络流量实例集合、流量统计特征集合、流量所属应用类型及协议; 流量所属应用类型的下一层描述各种应用类型。

### 2.1 网络流量本体并行构建模型

MapReduce的任务分解规约的分布式计算模式在 Google 系统上得到很好的验证, 而且以 MapReduce思想设计的语义推理算法也已经被证明是有效的<sup>[20]</sup>。因此, 面向海量的网络流量数据本体构建, 本文采用MapReduce作为并行处理技术, 实现实时、在线的网络流量数据并行集成。基于MapReduce的并行化网络流量本体构建模型如图1所示。采用基于MapReduce的语义映射方法, 根据网络流量本体结构, 实现网络结点流量数据到本体的完整映射, 为网络流量分类模块提供访问和操作数据的统一接口。

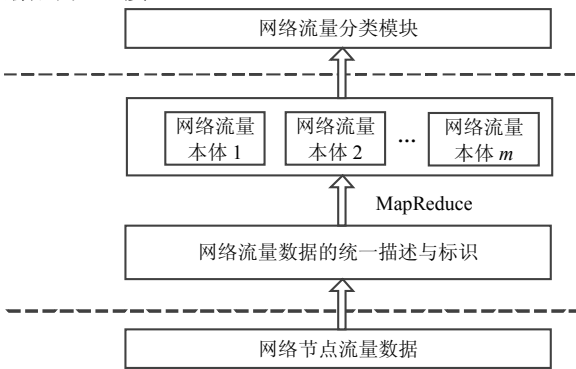


图1 基于MapReduce的网络流量本体构建模型

### 2.2 方法的设计与实现

为实现从网络流量信息采集到流量信息资源本体描述流程一体化, 依照MapReduce的架构, 将Map函数设计成流量信息采集器, Combiner设计成流量过滤与整合器, 将Reduce函数设计成网络流量本体构建器, 通过整个MapReduce完成从网络采集节点到网络流量本体的完整映射, 实现对网络流量本体构建并行化。

该方法实现的具体步骤如下。

令 $N_i(1 \leq i \leq n)$ 表示第 $i$ 个网络节点ID,  $IP_i$ 表示第 $i$ 个网络节点的IP,  $I_i$ 表示第 $i$ 个网络节点的相关信息,  $F_j(1 \leq j \leq m)$ 表示第 $j$ 条网络流量标识,  $O_j$ 表示第 $j$ 条网络流量的本体。  $M_{NF}$ 表示从网络节点流量数据到网络流量的映射,  $R_{FO}$ 表示从网络流量到网络流量本体的映射。

1) 根据每个网络节点启动对应的Map函数, 其中, 每个Map函数以键值对 $\langle N_i, IP_i \rangle$ 作为输入;

2) Map函数根据 $IP_i$ 操作网络节点, 收集网络节点相关信息 $I_i$ , 并调用网络流量采集工具捕获网络数据包, 然后将采集到的所有资源传入Combiner中间结果;

3) Combiner根据过滤规则提取所需网络流量信息, 并将数据包整合成网络流量 $F_j$ , 以键值对 $\langle I_i, F_j \rangle$ 的形式传给Reduce函数, 此时, 传向Reduce函数的每个键值对就对应着一条完整的网络流量信息。该步骤完成网络节点到网络流量的映射 $M_{NF}: N_i \rightarrow F_j$ ;

4) Reduce函数根据接收到的键值对计算流量统计特征, 并用本体语言OWL做统一资源描述, 借助本体建模工具Protégé的API, 完成网络流量本体的构建。该步骤实现网络流量到网络流量本体的映射 $R_{FO}: F_j \rightarrow O_j$ 。

## 3 基于知识推理的并行分类方法

### 3.1 知识推理的并行分类框架

针对上一节构建的网络流量本体, 提出一种基于知识推理的并行网络流量分类方法。该方法利用本体支持知识推理的特性, 通过知识推理挖掘出本体中概念、属性间的隐含关系; 考虑到大规模复杂网络下网络流量是海量的, 要实现在线流量实时分类, 则必须借助并行处理技术, 建立基于MapReduce的并行知识推理引擎, 实现网络流量实例与所属应用类型之间的对应关系, 即对网络流量进行分类。该方法的框架如图2所示。

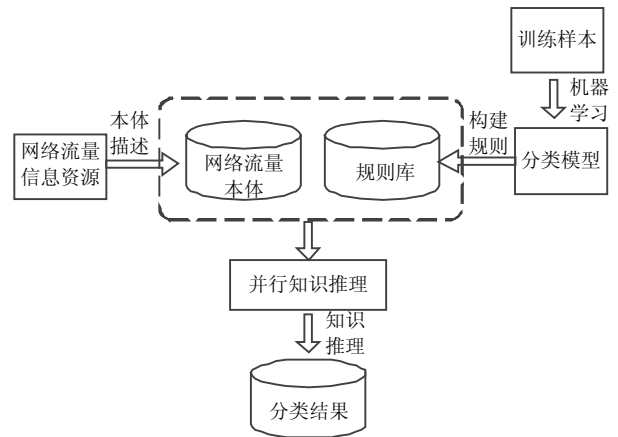


图2 基于知识推理的并行网络流量分类框架

由图2可以看出, 方法采用的知识推理是一种基于规则的推理。首先通过机器学习算法训练传统分类模型, 接着分析分类模型的内部结构, 将其转换成可供本体做知识推理的规则集形式, 然后以网络流量本体作为推理对象, 将本体和规则集一并输入并行知识推理引擎, 使得本体基于规则集做出推理, 自动对本体中流量实例的应用类型进行标注, 最终

得出分类结果。

### 3.2 并行知识推理引擎的设计与实现

为了能处理海量的网络流量本体推理,高效地执行推理过程,将采用MapReduce并行处理技术构建并行知识推理引擎。该引擎可以直接处理网络流量本体,即以未被标记应用类型的网络流量本体为输入,通过结合推理规则集对本体进行推理,得到网络流量的应用类型,最终将属于同一应用类型的网络流量作为输出,完成网络流量本体到网络流量类别的映射,实现对网络流量的分类。

并行知识推理的实现步骤如下。

令 $O_j(1 \leq j \leq n)$ 表示第 $j$ 个网络流量本体分片, $FI_l(1 \leq l \leq p)$ 表示第 $l$ 个网络流量实例标识(对应于第 $l$ 条网络流量 $F_l$ ), $S$ 表示推理引擎中的规则集, $L_k(1 \leq k \leq m)$ 表示第 $k$ 类(指应用类别)流量标签, $C_k$ 表示第 $k$ 类已分类流量集。 $MR_{OC}$ 表示从网络流量本体到网络流量类别的完整映射。

1) 根据每个计算节点的性能以及网络流量本体中所描述的网络流量实例的数据规模,对已构建好的网络流量本体进行分割,得到多个网络流量本体分片 $O_j$ ,将网络流量本体分片上传至Hadoop分布式文件系统,并对每一个网络流量本体分片中描述的网络流量实例标记为 $FI_l$ ,以键值对 $\langle FI_l, O_j \rangle$ 作为Map函数的输入;

2) 启动多个Map函数并行地调用Jena推理机,推理机利用规则集中的各条规则对网络流量本体分片 $O_j$ 中描述的与网络流量实例 $FI_l$ 有关的各种信息资源进行知识推理,得出 $FI_l$ 的类标签 $L_k$ ,将 $\langle L_k, FI_l \rangle$ 作为键值对输出到Reduce函数;

3) Reduce函数根据 $L_k$ 按类型合并 $F_j$ ,形成已分类流量集 $C_k$ ,至此完成流量本体集到已分类流量集的映射 $MR_{OC}: O_j \rightarrow C_k$ 。

## 4 实验与结果分析

### 4.1 实验环境及数据

本实验建立在Apache的开源项目Hadoop<sup>[21]</sup>系统之上,搭建的Hadoop平台由4台机器(即4个节点)构成。节点配置信息如下:4核CPU(Intel I7-3770, 3.4 GHz)、4 G内存、1 TB硬盘,运行64位Ubuntu系统。

采用文献[22]采集并公开的真实网络流量数据集作为本文的实验数据,称之为摩尔数据集。摩尔数据集中的每个网络流量样本都是完备的传输控制协议(TCP)双向流量,共有248个网络流量统计特征。

### 4.2 加速比

为了能较精确地衡量提出的方法采用并行化技术MapReduce所带来的性能方面的提升,使用加速比 $R$ 作为评价指标:

$$R = T_s / T_p \quad (1)$$

式中, $T_s$ 表示单机环境下方法的运行时间; $T_p$ 表示并行环境下方法的运行时间。

### 4.3 网络流量本体并行化构建效率评估实验

为验证网络流量本体构建采用MapReduce并行化处理执行效率,针对不同网络流量数据规模,在单机和多台机情况下对比网络流量本体构建时间,对比结果如图3所示。

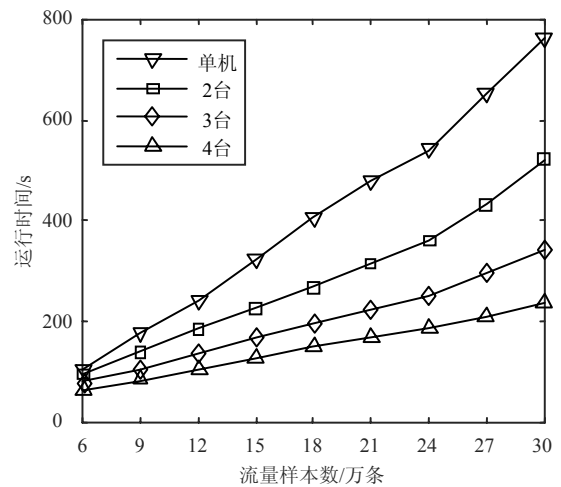


图3 单机环境和集群环境下网络流量本体的构建时间对比

从图3可以看出,当网络流量样本数较少时,不同个数的计算节点构建网络流量本体所需的时间差距不大。随着网络流量样本数据规模的增大,网络流量本体构建所需的时间几乎呈线性增长。由此可知,本体构建时间与网络流量样本数据规模呈线性关系,不受流量样本间相互关系的影响,比较适合并行化处理。另外,图3也显示多台机环境下比单机环境下的网络流量本体构建时间少,也说明本方法体现了并行化处理的优点。

图4给出了当集群环境分别为采用2、3、4台机,即计算节点分别为2、3、4时,本方法的加速比曲线图。

如图4所示,从3条加速比曲线之间的间隔变化来看,每增加一个计算节点,加速比都会有相应幅度的提升,且提升幅度较为稳定,这意味着集群中的计算节点可以平衡地分担本体构建任务,不同的计算节点均独立完成构建任务,各节点完成自身任务时相互之间不需要进行信息资源交互。因此,适当增加计算节点可以得到与节点个数成比例的加速

比提升。综上所述, MapReduce并行处理技术可以有效地提高网络流量本体构建的效率。

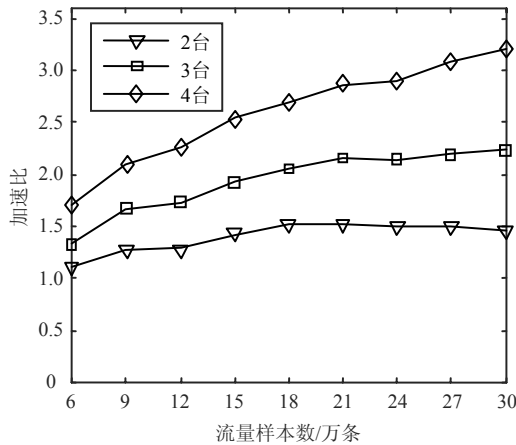


图4 加速比曲线

#### 4.4 并行知识推理的分类执行效率评估实验

文献[4]证明了采用决策树算法建立分类模型并转化成知识推理规则集, 能够很好地继承决策树算法在网络流量分类上的性能优势, 实验表明在准确率、召回率、F1-Measure 3个性能评价指标均优于SVM、BP神经网络和贝叶斯网络, 从而获得更佳的网络流量分类性能。

为验证并行知识推理的分类方法的执行效率, 采用不同网络流量数据规模, 对单机环境和集群环境下知识推理分类时间进行对比, 对比结果如图5所示。

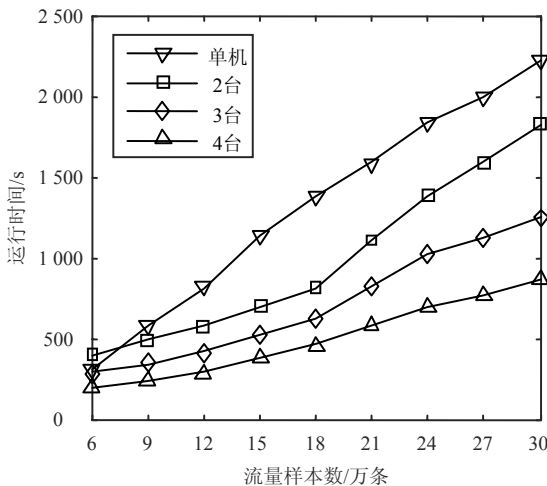


图5 单机环境和集群环境下网络流量分类时间对比

从图5中可看出, 当网络流量实例个数较少时, 集群环境下的分类时间与单机环境的分类时间差距较小。在流量样本数只有6万条的小规模分类任务中, 单机环境所需分类时间甚至低于只开启了两个节点的集群环境, 逼近于开启了3个节点的集群环

境, 这是因为当数据量较少时, MapReduce的过程中调度任务以及分割和重组数据等步骤需要耗费一定的时间, 由此可知对于小规模数据的处理, hadoop平台的优势无法展现。随着网络流量样本数的增长, 单机与集群环境的分类方法运行时间的差距也越来越大, 此时MapReduce的额外开销逐步趋于稳定, 方法中并行处理的优势就逐渐显现出来, 体现了并行模型的高效性。

图6给出了当集群环境采用2、3、4台机, 即计算节点分别为2、3、4时, 本方法的加速比曲线图。

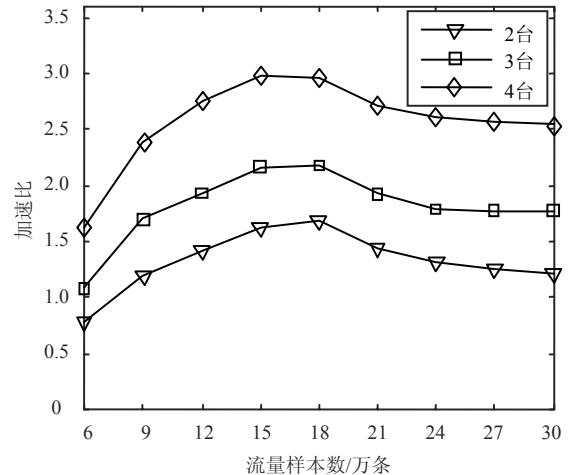


图6 加速比曲线

如图6所示, 当流量样本数一定时, 随着计算节点的增加, 其加速比呈现阶跃式变化; 随着样本数的增加, 加速比在增大到一个最大值之后减小, 之后趋于稳定。经过对各个节点运行状态的观测与分析可知, 当流量样本数较小时, 集群的资源利用率不高, 各计算节点的资源没有被有效利用; 随着样本数的增加, 加速比呈现上扬趋势, 逐渐增加到最大值, 此时集群的资源利用率达到最高, 集群中各个节点的资源均能被很好地调度; 随着样本数继续增加, 加速比慢慢减小, 然后趋于平稳, 这是因为集群资源的利用已达到瓶颈, 集群的调度器开始调整调度策略, 最终达到一个稳态。

综上所述, 并行知识推理的分类方法采用MapReduce并行架构能够有效地提高大规模网络流量的分类效率。

## 5 结束语

针对海量、多源、异构网络流量的分类, 本文结合了本体和云计算技术的优势, 提出了一种基于本体的并行网络流量分类方法。该方法建立在MapReduce并行计算架构之上, 结合网络流量本体

结构,设计基于MapReduce的网络流量本体构建方法,及并行知识推理的网络流量分类方法。实验表明,MapReduce并行处理技术可以有效地提高网络流量本体构建的效率及网络流量的分类效率。下一步我们将研究对未知网络流量的描述以及本体知识库的自我管理与更新,实现在线的网络流量本体并行化构建方法和网络流量的实时并行分类方法,以适应大规模复杂网络流量的实时、准确分类。

### 参 考 文 献

- [1] WANG Yu, XIANG Yang, ZHANG Jun, et al. Internet traffic classification using constrained clustering[J]. IEEE Transactions on Parallel and Distributed Systems, 2014, 25(11): 2932-2943.
- [2] CARELA-ESPAÑOL V, BARLET-ROS P, MULA-VALLS O, et al. An autonomic traffic classification system for network operation and management[J]. Journal of Network and Systems Management, 2015, 23(3): 401-419.
- [3] 刘凯鹏, 方滨兴. 基于社会性标注的本体学习方法[J]. 计算机学报, 2010, 33(10): 1823-1834.  
LIU Kai-peng, FANG Bin-xing. Ontology induction based on social annotations[J]. Chinese Journal of Computers, 2010, 33(10): 1823-1834.
- [4] 陶晓玲, 韦毅, 孔德艳, 等. 基于本体的网络流量分类方法[J]. 计算机工程与设计, 2016, 37(1): 31-36.  
TAO Xiao-ling, WEI Yi, KONG De-yan, et al. Network traffic classification method based on ontology[J]. Computer Engineering and Design, 2016, 37(1): 31-36.
- [5] HAUG P J, FERRARO J P, HOLMEN J, et al. An ontology-driven, diagnostic modeling system[J]. Journal of the American Medical Informatics Association, 2013, 20(e1): e102-e110.
- [6] OELLRICH A, WALLS R L, CANNON E K S, et al. An ontology approach to comparative phenomics in plants[J]. Plant Methods, 2015, 11(1): 10.
- [7] AZEVEDO C L B, IACOB M E, ALMEIDA J P A, et al. Modeling resources and capabilities in enterprise architecture: a well-founded ontology-based proposal for archimate[J]. Information Systems, 2015, 54(12): 235-262.
- [8] EBRAHIMPOUR V, YACOUT S. Ontology-based schema to support maintenance knowledge representation with a case study of a pneumatic valve[J]. IEEE Transactions on Systems, Man, and Cybernetics: Systems, 2015, 45(4): 702-712.
- [9] Gene Ontology Consortium. Gene ontology annotations and resources[J]. Nucleic Acids Research, 2013, 41(D1): D530-D535.
- [10] ALMENDROS-JIMENEZ J M, DOMENE L, PIEDRA-FERNANDEZ J A. A framework for ocean satellite image classification based on ontologies[J]. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 2013, 6(2): 1048-1063.
- [11] ZHANG Xiao-wei, HU Bin, MA Xu, et al. Ontology driven decision support for the diagnosis of mild cognitive impairment[J]. Computer Methods and Programs in Biomedicine, 2013, 113(3): 781-791.
- [12] PIETRZYK M, JANOWSKI L, URVOY-KELLER G. Toward systematic methods comparison in traffic classification[C]//Wireless Communications and Mobile Computing Conference(IWCMC), 2011 7th International. Istanbul: IEEE, 2011:1022-1027.
- [13] GU Cheng-jie, ZHANG Shun-yi, XUE Xiao-zhen. Online self-learning internet traffic classification based on profile and ontology[J]. Journal of Convergence Information Technology, 2011, 6(4): 81-91.
- [14] XIN Jun-chang, WANG Zhi-qiong, CHEN Chen, et al. ELM\*: Distributed extreme learning machine with map reduce[J]. World Wide Web, 2014, 17(5): 1189-1204.
- [15] CHEN Jiao-yan, CHEN Hua-jun, WAN Xiang-yi, et al. MR-ELM: a map reduce-based framework for large-scale elm training in big data era[J]. Neural Computing and Applications, 2016, 27(1): 101-110.
- [16] AI-KHATEEB T M, MASUD M M, KHAN L, et al. Cloud guided stream classification using class-based ensemble [C]//Proceedings of 2012 IEEE 5th International Conference on Cloud Computing (CLOUD). Honolulu, HI, USA: [s.n.], 2012: 694-701.
- [17] SARA DEL RÍO, VICTORIA LÓPEZ, JOSÉ MANUEL BENITEZ, et al. On the use of map reduce for imbalanced big data using random forest[J]. Information Sciences, 2014, 258(11): 112-137.
- [18] 郑天红. 基于Hadoop的网络流量分流并行化设计[D]. 呼和浩特: 内蒙古大学, 2012.  
ZHENG Tian-hong. Design and Implementation of the parallelization based on hadoop model network traffic diversion[D]. Huhehaote: Inner Mongolia University, 2012.
- [19] MU Xue-feng, WU Wen-jun. A parallelized network traffic classification based on hidden markov model[C]// Proceedings of 2011 International Conference on Cyber-Enabled Distributed Computing and Knowledge Discovery. Beijing: [s.n.], 2011: 107-112.
- [20] DAI Chao-fan, FENG Yang-he, ZHANG Peng-cheng. Retracted article: Research of ontology-based model representation method[C]//Proceedings of 2010 Second International Conference on Information Technology and Computer Science. Kiev, Ukraine: [s.n.], 2010: 364-367.
- [21] The Apache software foundation. Hadoop[EB/OL]. [2015-10-22]. <http://hadoop.apache.org/>.
- [22] MOORE A W, ZUEV D. Internet traffic classification using bayesian analysis techniques [EB/OL]. [2015-11-11]. <http://www.cl.cam.ac.uk/research/srg/netos/nprobe/data/papers/sigmetrics/index.html>.

编辑 蒋晓