

[本期目录](#) | [下期目录](#) | [过刊浏览](#) | [高级检索](#)

[\[打印本页\]](#) [\[关闭\]](#)

## 开发研究与设计技术

### 一种适用于应用层协议的特征提取算法

张玉冲, 蔡皖东, 丁要军

(西北工业大学计算机学院, 西安 710129)

**摘要:** 基于PrefixSpan算法, 提出一种适用于应用层协议的特征提取算法。通过加入关于位置的约束, 减少频繁序列模式的产生数量, 结合特征提取过程的实际情况加入约束条件, 从而在挖掘过程中减少投影数据库产生的个数, 在构建投影数据库过程中, 去除关于非频繁项的存储及投影数据库中序列数小于最小支持度的扫描过程。实验结果表明, 与原算法相比, 该算法的运行时间较短, 提取的特征具有较高的准确率和较低的误报率。

**关键词:** 序列模式挖掘 特征提取 PrefixSpan算法 关联规则

### Feature Extraction Algorithm for Application Layer Protocol

ZHANG Yu-chong, CAI Wan-dong, DING Yao-jun

(College of Computer, Northwestern Polytechnical University, Xi'an 710129, China)

**Abstract:** This paper proposes a feature extraction algorithm for application layer protocol based on PrefixSpan algorithm. It adds the constraints about position to reduce the number of frequent sequence mode, and is combined with the practical situation of the feature extraction process to join constraint conditions. In the mining process, it reduces the number of projection database. In construction of the projection database, it removes the storage about the frequent items and the scanning of which sequence number is less than the minimum support degree in projection database. Experimental results show that compared with the original algorithm, the running time of this algorithm is shorter, and the extraction features have higher precision and lower false alarm rate.

**Keywords:** sequential pattern mining feature extraction PrefixSpan algorithm association rule

收稿日期 2011-07-28 修回日期 网络版发布日期 2012-02-20

DOI: 10.3969/j.issn.1000-3428.2012.04.087

基金项目:


国家“863”计划基金资助项目(2009AA01Z424)

通讯作者:

**作者简介:** 张玉冲(1987—), 男, 硕士研究生, 主研方向: 网络与信息安全; 蔡皖东, 教授、博士生导师; 丁要军, 博士研究生

**通讯作者**E-mail: 444926690@qq.com

## 参考文献:

- [2] Zander S.[J].Nguyen1 T, Armitage1 G. Self-learning IP Traffic Classification Based on Statistical Flow Characteristics[C]//Proc. of PAM'05. Boston, USA: [s. n.].2005,;- 
- [3] 余 锋, 王小玲. 基于半监督学习的网络流量分类[J].计算机工程.2009, 35(12):90-91 [浏览](#)
- [4] Park B.[J].Won Y J, Kim M, et al. Towards Automated Application Sig- nature Generation

## 扩展功能

本文信息

▶ [Supporting info](#)

▶ [PDF\(277KB\)](#)

▶ [\[HTML\] 下载](#)

▶ [参考文献\[PDF\]](#)

▶ [参考文献](#)

## 服务与反馈

▶ [把本文推荐给朋友](#)

▶ [加入我的书架](#)

▶ [加入引用管理器](#)

▶ [引用本文](#)

▶ [Email Alert](#)

▶ [文章反馈](#)

▶ [浏览反馈信息](#)

## 本文关键词相关文章

▶ [序列模式挖掘](#)

▶ [特征提取](#)

▶ [PrefixSpan算法](#)

▶ [关联规则](#)

## 本文作者相关文章

▶ [张玉冲](#)

▶ [蔡皖东](#)

▶ [丁要军](#)

## PubMed

▶ [Article by Zhang, Y. C.](#)

▶ [Article by Ca, H. D.](#)

▶ [Article by Ding, Y. J.](#)

for Traffic Identification[C]//Proc. of NOMS'08. [S. l.]: IEEE Press.2008, :-crossref

[7] Pei Jian.[J].Han Jiawei, Behzad M A, et al. PrefixSpan: Mining Sequential Patterns Efficiently by Prefix-projected Pattern Growth[C]//Proc. of the 17th International Conference on Data Engineering. [S. l.]: IEEE Press.2001, :-crossref

[8] Lin Guanzhou.[J].Xin Yang, Yang Yixian. An Improved Prefix- Spanbased Signatures Mining Algorithm with Offset Con- straint[C]//Proc. of the 2nd International Workshop on Intelligent Systems and Applications. [S. l.]: IEEE Press.2010, :-crossref

#### 本刊中的类似文章

1. 张令杰, 徐维祥.基于时态约束的关联规则挖掘算法[J]. 计算机工程, 2012,38(5): 50-52
2. 李金凤, 王怀彬.基于关联规则的网络故障告警相关性分析[J]. 计算机工程, 2012,38(5): 44-46
3. 尚丽, 淮文军, 杜吉祥.具有Fisher判据约束的非负稀疏编码模型? [J]. 计算机工程, 2012,38(3): 176-177,179
4. 王强, 谢建军, 李华飙, 南凯.协同工作环境中基于用户行为的页面推荐[J]. 计算机工程, 2012,38(3): 249-251
5. 张宗郁, 张亚平, 张静远, 张晓君.改进关联规则算法在高校教学管理中的应用[J]. 计算机工程, 2012,38(2): 75-77
6. 潘冬生, 章昭辉, 代秀娟, 杨娟.基于层次梯度分析的协同数据挖掘算法[J]. 计算机工程, 2012,38(2): 72-74
7. 冯筠, 李刚, 孙霞, 冯宏伟.一种面向教学的知识点库自动生成方法[J]. 计算机工程, 2012,38(2): 201-203
8. 胡石, 梅雪.人体行为动作的形状轮廓特征提取及识别[J]. 计算机工程, 2012,38(2): 198-200
9. 贾洁, 王慧琴, 胡燕, 马宗方.基于最小二乘支持向量机的火灾烟雾识别算法[J]. 计算机工程, 2012,38(2): 272-275
10. 姚全珠, 李如琼, 王美君.项约束先过滤的最大频繁项集挖掘算法[J]. 计算机工程, 2012,38(04): 73-75

#### 文章评论

反馈人	<input type="text"/>	邮箱地址	<input type="text"/>
反馈标题	<input type="text"/>	验证码	<input type="text" value="2694"/>
<input type="text"/>			