

基于关联规则挖掘的社区网络数据爬虫算法

穆俊

(滇西科技师范学院 信息工程学院, 云南 临沧 677000)

摘要: 为了提高社区网络的联合推荐性能, 需要进行网络数据爬虫设计, 提出一种基于关联规则挖掘的社区网络数据爬虫算法. 构建社区网络的信息传递模型, 挖掘社区网络用户行为信息特征量, 根据数据的属性特征进行关联规则合并, 采用模糊指向性聚类方法进行社区网络的用户行为属性特征聚类处理, 对聚类输出的特征量采用自相关模板匹配方法实现信息融合, 实现社交网络数据的关联规则挖掘, 结合数据聚类分布属性进行网络信息爬虫, 实现社区属性混合推荐. 仿真结果表明, 采用该算法进行社区网络数据爬虫处理的准确度较高, 个性化匹配程度较好, 对社区网络联合推荐结果的置信度水平较高, 提高了社区发现能力.

关键词: 关联规则挖掘; 社区网络; 数据爬虫; 推荐

中图分类号: TP391

文献标识码: A

文章编号: 1000-7180(2018)08-0105-04

A Hybrid recommendation Model for Community Attributes of Social Networks Based on Association Rule Mining

MU Jun

(School of Information Science&Engineering, West Yunnan University, Lincang 677000 China)

Abstract: In order to improve the joint recommendation performance of the community network, we need to design the network data crawler, and propose a community network data crawler algorithm based on association rule mining. Constructing the information transfer model of the community network, mining the information characteristic quantity of the user behavior of the community network, combining the association rules according to the attribute characteristic of the data. The fuzzy directivity clustering method is used to cluster the user behavior attributes of the community network. The autocorrelation template matching method is used to realize the information fusion and the association rules mining of the social network data. The network information crawler is implemented by combining the data clustering and distribution attributes, and the community attribute mixed recommendation is realized. The simulation results show that the proposed algorithm has higher accuracy, better personalized matching degree and higher confidence level to the community network joint recommendation results, which improves the community discovery ability.

Key words: association rule mining; community network; data crawler; recommendation

作者简介:

穆俊男, (1979-), 硕士, 副教授. 研究方向为计算机软件与理论、计算机应用、数据挖掘.

E-mail: mujun511@126.com.