

# 软件学报 *Software* Journal of

2023 年 第 34 卷 第 3 期

大数据治理的理论与技术专题

科学出版社



主办



中国科学院软件研究所



中国计算机学会

出版



科学出版社



2013 年全国百强科技期刊



2015 年全国百强科技期刊



中国精品科技期刊



中国计算机学会会刊

# 软件学报

(Ruanjian Xuebao)

第 34 卷第 3 期  
2023 年 3 月

## 目次

### 大数据治理的理论与技术专题

大数据治理的理论与技术专题前言 .....	杜小勇 杨晓春 童咏昕 (1007)
面向列语义识别的共现属性交互模型构建与优化 .....	高珊 袁宛竹 卢卫 王兰 张静 杜小勇 (1010)
面向聚合查询的 Apache IoTDB 物理元数据管理 .....	赵东明 邱圆辉 康瑞 宋韶旭 黄向东 王建民 (1027)
基于多视角的多类型错误全面检测方法 .....	彭锦峰 申德荣 寇月 聂铁铮 (1049)
兼顾行列的时序数据质量规则发现 .....	丁小欧 李映泽 王晨 王宏志 李昊轩 (1065)
预训练语言模型实体匹配的可解释性 .....	梁峥 王宏志 戴加佳 邵心玥 丁小欧 穆添愉 (1087)
面向数据联邦的安全多方 $\theta$ -连接算法 .....	张媛媛 李书缘 史焯轩 周南 徐毅 许可 (1109)
基于联邦学习的跨源数据错误检测方法 .....	陈璐 郭宇翔 葛丛丛 郑白桦 高云君 (1126)
基于贡献度证明共识机制的去中心化联邦学习框架 .....	韩楠 杨国平 李贺 袁冠 毛睿 元昌安 Louis Alberto GUTIERREZ (1148)
联邦学习贡献评估综述 .....	王勇 李国良 李开宇 (1168)
跳跃滤波: 一种面向大数据治理的动态数据摘要设计 .....	符鹏涛 罗来龙 郭得科 赵翔 李尚森 王怀民 (1193)
面向开放大数据环境的动态数据保护系统 .....	屠要峰 牛家浩 王德政 高洪 徐进 洪科 阳方 (1213)
面向大数据分析的分布式矩阵计算系统研究进展 .....	陈梓浩 徐辰 钱卫宁 周傲英 (1236)
基于多粒度注意力网络的知识超图链接预测 .....	庞俊 刘小琪 谷峪 王鑫 赵宇海 张晓龙 于戈 (1259)
属性公平的异质信息网络上的社区搜索算法 .....	乔连鹏 侯会文 王国仁 (1277)
基于宽容训练和隐私保护的快速监控视频检索模型 .....	覃浩 王平辉 张若非 覃遵颖 (1292)

### 系统软件与软件工程

基于双重信息检索的 Bash 代码注释生成方法 .....	陈翔 于池 杨光 濮雪莲 崔展齐 (1310)
开源软件供应链安全研究综述 .....	纪守领 王琴应 陈安莹 赵彬彬 叶童 张旭鸿 吴敬征 李昀 尹建伟 武延军 (1330)

### 模式识别与人工智能

深度知识追踪模型综述和性能比较 .....	王宇 朱梦霞 杨尚辉 陆雪松 周傲英 (1365)
-----------------------	---------------------------

### 数据库技术

数据定价与交易研究综述 .....	江东 袁野 张小伟 王国仁 (1396)
高速流环境下近似连续 $k$ 代表轮廓查询算法 .....	朱睿 宋楸尧 王斌 杨晓春 张安珍 夏秀峰 (1425)
多模态医疗数据中海量小文件存储优化方法 .....	曾梦 邹北骥 张文生 杨雪冰 朱承璋 (1451)

### 计算机网络与信息安全

基于区块链的物联网认证机制综述 .....	程冠杰 邓水光 温盈盈 严学强 赵明宇 (1470)
基于智能合约的工业互联网数据公开审计方案 .....	李涛 杨安家 翁健 郭梓繁 (1491)
地区网络边界发现方法 .....	朱金玉 张宇 曾良伟 张宏莉 方滨兴 (1512)

《软件学报》投稿指南 ..... (封三)

期刊基本参数: CN11-2560/TP\*1990\*m\*16\*518\*zh+en\*P\*¥70\*2023\*25\*2023-03

## Contents

**SPECIAL TOPIC ON BIG DATA GOVERNANCE: THEORY AND TECHNIQUES**

- 1007 Preface  
*DU Xiao-Yong, YANG Xiao-Chun, TONG Yong-Xin*
- 1010 Construction and Optimization of Co-occurrence-attribute-interaction Model for Column Semantic Recognition  
*GAO Shan, YUAN Wan-Zhu, LU Wei, WANG Lan, ZHANG Jing, DU Xiao-Yong*
- 1027 Physical Metadata Management in Apache IoTDB for Aggregate Queries  
*ZHAO Dong-Ming, QIU Yuan-Hui, KANG Rui, SONG Shao-Xu, HUANG Xiang-Dong, WANG Jian-Min*
- 1049 Comprehensive Error Detection Method for Multiple Types Errors Based on Multiple Views  
*PENG Jin-Feng, SHEN De-Rong, KOU Yue, NIE Tie-Zheng*
- 1065 Time Series Data Quality Rules Discovery with Both Row and Column Dependencies  
*DING Xiao-Ou, LI Ying-Ze, WANG Chen, WANG Hong-Zhi, LI Hao-Xuan*
- 1087 Interpretability of Entity Matching Based on Pre-trained Language Model  
*LIANG Zheng, WANG Hong-Zhi, DAI Jia-Jia, SHAO Xin-Yue, DING Xiao-Ou, MU Tian-Yu*
- 1109 Secure Multi-party  $\theta$ -join Algorithm Toward Data Federation  
*ZHANG Yuan-Yuan, LI Shu-Yuan, SHI Ye-Xuan, ZHOU Nan, XU Yi, XU Ke*
- 1126 Cross-source Data Error Detection Approach Based on Federated Learning  
*CHEN Lu, GUO Yu-Xiang, GE Cong-Cong, ZHENG Bai-Hua, GAO Yun-Jun*
- 1148 Decentralized Federated Learning Framework Based on Proof-of-contribution Consensus Mechanism  
*QIAO Shao-Jie, LIN Yu-Feng, HAN Nan, YANG Guo-Ping, LI He, YUAN Guan, MAO Rui, YUAN Chang-An, Louis Alberto GUTIERREZ*
- 1168 Survey on Contribution Evaluation for Federated Learning  
*WANG Yong, LI Guo-Liang, LI Kai-Yu*
- 1193 Jump Filter: Dynamic Sketch Design for Big Data Governance  
*FU Peng-Tao, LUO Lai-Long, GUO De-Ke, ZHAO Xiang, LI Shang-Sen, WANG Huai-Min*
- 1213 Dynamic Data Protection System for Open Big Data Environment  
*TU Yao-Feng, NIU Jia-Hao, WANG De-Zheng, GAO Hong, XU Jin, HONG Ke, YANG Fang*
- 1236 Research Progress on Distributed Matrix Computation Systems for Big Data Analysis  
*CHEN Zi-Hao, XU Chen, QIAN Wei-Ning, ZHOU Ao-Ying*
- 1259 Knowledge Hypergraph Link Prediction Based on Multi-granular Attention Network  
*PANG Jun, LIU Xiao-Qi, GU Yu, WANG Xin, ZHAO Yu-Hai, ZHANG Xiao-Long, YU Ge*
- 1277 Community Search Algorithm on Heterogeneous Information Networks Based on Attribute Fairness  
*QIAO Lian-Peng, HOU Hui-Wen, WANG Guo-Ren*
- 1292 Fast Surveillance Video Retrieval Model Based on Tolerant Training and Privacy Protection  
*QIN Hao, WANG Ping-Hui, ZHANG Ruo-Fei, QIN Zun-Ying*

**SYSTEM SOFTWARE AND SOFTWARE ENGINEERING**

- 1310 Bash Code Comment Generation Method Based on Dual Information Retrieval  
*CHEN Xiang, YU Chi, YANG Guang, PU Xue-Lian, CUI Zhan-Qi*
- 1330 Survey on Open-source Software Supply Chain Security  
*JI Shou-Ling, WANG Qin-Ying, CHEN An-Ying, ZHAO Bin-Bin, YE Tong, ZHANG Xu-Hong, WU Jing-Zheng, LI Yun, YIN Jian-Wei, WU Yan-Jun*

**PATTERN RECOGNITION AND ARTIFICIAL INTELLIGENCE**

- 1365 Review and Performance Comparison of Deep Knowledge Tracing Models  
*WANG Yu, ZHU Meng-Xia, YANG Shang-Hui, LU Xue-Song, ZHOU Ao-Ying*

**DATABASE TECHNOLOGY**

- 1396 Survey on Data Pricing and Trading Research  
*JIANG Dong, YUAN Ye, ZHANG Xiao-Wei, WANG Guo-Ren*
- 1425 Approximate Continuous  $k$  Representative Skyline Query Algorithm over High-Speed Streaming Data Environment  
*ZHU Rui, SONG Fu-Yao, WANG Bin, YANG Xiao-Chun, ZHANG An-Zhen, XIA Xiu-Feng*
- 1451 Optimization Method for Storing Massive Small Files in Multi-modal Medical Data  
*ZENG Meng, ZOU Bei-Ji, ZHANG Wen-Sheng, YANG Xue-Bing, ZHU Cheng-Zhang*

**COMPUTER NETWORKS AND INFORMATION SECURITY**

- 1470 Survey on Blockchain-based Internet of Things Authentication Mechanisms  
*CHENG Guan-Jie, DENG Shui-Guang, WEN Ying-Ying, YAN Xue-Qiang, ZHAO Ming-Yu*
- 1491 Public Auditing Scheme for Industrial Internet Data Based on Smart Contracts  
*LI Tao, YANG An-Jia, WENG Jian, GUO Zi-Fan*
- 1512 Method of Discovering Regional Network Border  
*ZHU Jin-Yu, ZHANG Yu, ZENG Liang-Wei, ZHANG Hong-Li, FANG Bin-Xing*