系统工程与电子技术 2012, 34(7) 1505-1510 DOI: 10.3969/j.issn.1001-506X.2012.07.36 ISSN: 1001-506X CN: 11-2422/TN

软件、算法与仿真

基于SMC的隐私保护聚类模型

方炜炜[1,2], 杨炳儒[2], 夏红科[1,2]

1. 北京信息科技大学计算中心，北京 100192；
2. 北京科技大学信息工程学院，北京100083

摘要：

隐私保护数据挖掘指在实现准确挖掘知识的同时确保敏感数据不泄露。针对垂直分布式数据存储结构的聚类隐私保护问题，提出基于全同态加密协议和数据扰乱方法的隐私保护聚类模型。该模型通过采用安全比较协议解决了垂直分布式聚类的两个隐私保护关键步骤：求解最近簇和判断质心变化，从而实现了数据的有效保护。理论证明了该模型的安全性并分析了其时间复杂度和通信耗量，实验结果表明该隐私保护聚类模型是安全有效的。

关键词： 安全多方计算 同态加密 聚类 隐私保护数据挖掘

Privacy-preserving clustering modeling based on SMC

FANG Wei-wei[1,2], YANG Bing-ru[2], XIA Hong-ke[1,2]

1. Computer Center, Beijing Information Science and Technology University, Beijing 100192, China;
2. School of Information Engineering, Beijing University of Science and Technology, Beijing 100083, China

Abstract:

Privacy-preserving data mining aims to accurately mine knowledge while unrevealing sensitive data. For solving the privacy-preserving clustering problem in vertical distribution, a privacy-preserving clustering model based on full homomorphous encryption protocols and data perturbation technology is proposed. The model protects original data effectively by using secure comparison protocols to compute the nearest cluster and estimate the updating of the cluster center, which are two key steps in clustering process. Theory argument demonstrates the security of the privacy-preserving clustering model and analyzes computation complexity and communication costs. Experiment results prove that the privacy-preserving clustering model is secure and effective.

Keywords: secure multi-party computation（SMC） homomorphous encryption clustering privacy preserving data mining

收稿日期 修回日期 网络版发布日期

DOI: 10.3969/j.issn.1001-506X.2012.07.36

基金项目:

通讯作者:

作者简介:

作者Email:

参考文献：

本刊中的类似文章

1．周洪娟,刘帅,金铭,乔晓林.基于DOA参数的雷达信号预分选[J]. 系统工程与电子技术, 2009,31(11): 2575-2577
2．申晓勇1, 雷英杰1,李 进1, 蔡 茹1,2.基于目标函数的直觉模糊集合数据的聚类方法[J]. 系统工程与电子技术, 2009,31(11): 2732-2735
3．吴静，吴晓燕，高忠长.基于模糊聚类和粗糙集的仿真可信性模糊综合评估[J]. 系统工程与电子技术, 2010,32(4): 770-773
4．阳春, 张向荣, 焦李成.结合Nystr m逼近的图半监督纹理图像分割[J]. 系统工程与电子技术, 2009,31(12): 2820-2825
5．宋晓宇，刘锋，孙焕良.基于粗糙集的聚类算法中阈值自动选取[J]. 系统工程与电子技术, 2010,32(1): 192-

194
6. 徐金华， 刘光斌， 余志勇.基于Vague聚类方法的战场电磁目标分选[J]. 系统工程与电子技术, 2010,32(05): 1011-1013
7. 李序,张葛祥,荣海娜.基于加权K-近邻法和SVC的雷达辐射源信号识别[J]. 系统工程与电子技术, 2010,32(6): 1215-1219
8. 刘松涛 1,2 ， 王维 2 ， 殷福亮 1.基于动态广义直方图均衡的红外图像增强方法[J]. 系统工程与电子技术, 2010,32(7): 1411-1414
9. 曾华，吴耀华，黄顺亮.非均匀类簇密度聚类的多粒度自学习算法[J]. 系统工程与电子技术, 2010,32(8): 1760-1765
10. 钟燕飞，张良培.遥感影像K均值聚类中的初始化方法[J]. 系统工程与电子技术, 2010,32(9): 2009-2014
11. 刘福才，任丽娜，路平立. 基于T-S模型的自适应多变量模糊预测控制[J]. 系统工程与电子技术, 2010,32(12): 2660-2663
12. 刘刚, 梁晓庚, 张京国.基于轮廓波变换和改进模糊c均值聚类的红外图像分割[J]. 系统工程与电子技术, 2011,33(2): 443-448
13. 谈璐璐, 张涛, 杨汝良.基于模糊聚类的PolInSAR数据非监督分类方法[J]. 系统工程与电子技术, 2011,33(2): 305-309
14. 徐选华, 范永峰.改进的蚁群聚类算法及在多属性大群体决策中的应用[J]. 系统工程与电子技术, 2011,33(2): 346-349
15. 胡利平，殷红成，陈渤，周平.改进的核子类判决分析[J]. 系统工程与电子技术, 2011,33(05): 1176-