大数据环境下网络数据隐私保护算法研究

潘明波

（云南工商学院 信息工程学院，云南 昆明 651701）

摘　要:针对传统的数据隐私保护算法存在执行时间较长,隐匿率较高,适应性较差等问题,提出基于密度聚类的网络数据隐私保护算法，在大数据环境下，随机生成一个满足给定的大数据环境下网络隐私数据特征的变换函数,然后利用这个变换函数对原始的网络隐私数据的数值进行变换，并将变换后的数值作为随机化回答的结果；然后根据网络的特征对网络节点进行密度聚类分析，生成符合大小的任意形状的簇，通过量化网络隐私数据信息丢失量和网络隐私数据结构信息丢失量对隐私数据进行有效性分析;最后在成功生成的簇内插入真实网络节点，通过增加边等技术，完成对网络数据信息的隐私保护。仿真实验结果证明，所提算法具有较好的隐私保护效果，在对多种不同背景知识攻击的情况下，具有较好的适应效果.

关键词：　〖HTF〗大数据环境下；基于密度聚类；网络数据隐私保护算法

Big Data Environment Network Data Privacy Protection Algorithm Research

PAN Ming-bo

(Information Engineering Institute,Yunnan Technology and Bussiness University,Kunming 651701,China)

Abstract：There are many problems in the traditional data privacy protection algorithm, such as long execution time, high concealment and poor adaptability. The proposed network data density clustering algorithm based on privacy protection, in the big data environment, the transformation function generates a random given the big data environment of network privacy data, then use the numerical transform function of network privacy data are changed, and the numerical transform as the result of random response then according to the characteristics of the network; the network node density clustering analysis, generated in accordance with any size and shape of the clusters, through the data network privacy and network privacy quantitative information loss data structure information for effective analysis of data privacy loss amount; and finally into the real network node in the successful generation of cluster, through increasing etc. technology, complete privacy protection of network information. The simulation results show that the proposed algorithm has a better privacy protection effect, and it has a good adaptability to the attack of a variety of different background knowledge.

Key words：big data environment; based on density clustering ; network data privacy protection algorithm

作者简介：

潘明波　男，(1984-)，讲师.研究方向为计算机科学技术、计算机应用.

E-mail：panmingbo6431@163.com.