

doi:10.1631/FITEE.1601479

题目：优化算法在所有权保留数据挖掘中的应用

概要：从敏感数据中提取知识往往需要协同工作。统计数据库根据这些敏感数据生成，并由各利益相关方共享。在此情况下，共享数据的所有权保护变得尤为重要。水印技术正逐渐成为一种推行数字数据格式所有权的有效工具，但该技术也可能导致数据失真。因此，从具有水印的数据中提取的知识可能不准确。数据失真程度由可用性约束条件来控制，这反过来又限制了可用于添加水印的带宽。尽管大带宽能保证鲁棒性，但可能降低数据质量。该问题可以通过在可用性约束条件下优化可用带宽来解决。如今，优化技术——尤其是生物启发式技术——已成为解决该类问题的首选。本文分析了多种优化方案及其可行性，用于优化添加水印的最大可用带宽，并期望达到以下两个目标：（1）保持数据中存储的知识不变；（2）在可用性约束条件下使可用带宽最大化，以取得最佳鲁棒性。第一个目标利用一个可用性约束模型实现，该模型能确保知识不会因嵌入水印而受到损害。第二个目标通过找到满足第一个目标的可用性约束条件下最大带宽实现。采用不同指标对多种优化方案性能进行了评估。

关键词：信息安全；优化技术；数字版权；水印技术