Abstract

Data mining services require accurate input data for their results to be meaningful, but privacy concerns may influence users to provide spurious information. To preserve client privacy in the data mining process, a variety of techniques based on random perturbation of data records have been proposed recently. One known fact which is very important in data mining is discovering the association rules from database of transactions where each transaction consists of set of items. Two important terms support and confidence are associated with each of the association rule. Actually any rule is called as sensitive if its disclosure risk is above a certain privacy threshold. Sometimes we do not want to disclose a sensitive rule to the public because
of confidentiality purposes. This paper is extension of work done in [1]. In [1] a reduction of 3-SAT problem from optimal sanitization in association rule hiding is presented. This paper proves that optimal sanitization in association rule hiding is NP-Complete. The proofs are based on reduction from 3-SAT.

Reference


Index Terms

Computer Science Communications
Key words

3-SAT
Association Rule

NP-Complete
Data mining
Data Sanitization