A New Data Mining Method Based on Relational Graphs of Products

Shyi-Ming Chen

National Taiwan University

of Science and Technology

R.O.C.

Chun-Te Lee

National Taiwan University

of Science and Technology

R.O.C.

Abstract

Data mining is a technique to extract implicit knowledge from raw data. In recent years, many methods have been proposed to deal with data mining. In this paper, we propose a new data mining method based on relational graphs of products. The proposed method statistically analyzes the buying relationships or associations between customers and products. It constructs the relational graphs of products, and generates the corresponding association rules and serial rules from the constructed relational graphs. The proposed method is an incremental data mining method. It uses the relational graphs as a visualization tool to describe the relationships among product pairs in the transactional database. It obtains the occurrence frequency of each itemset from the transactional database statistically and constructs the relational graph. Then, it finds out the complete β-polygon sets (i.e., frequent β-itemsets) from the constructed relational graph whose associated values of edges are larger than the frequency threshold value α by using a graph search method, where $\alpha \geq 0$, $\beta \geq 2$, and α and β are user-defined positive integers. We choose the minimum value for α to prune the least number of edges from the relational graph and

choose the maximum value for β to find the maximum complete β-polygon set. Based on the derived maximum complete β-polygon sets, it generates association rules and serial rules of length γ, where $\gamma \geq 3$. Finally, it calculates the minimum support and minimum confidence values of the generated rules, respectively. The proposed method can generate association rules and serial rules from raw data in a more flexible and more intelligent manner than the existing methods.