

KHAI THÁC LUẬT KẾT HỢP CÓ ĐỘ HỖ TRỢ TỐI ƯU TRÊN CÁC THUỘC TÍNH SỐ

Hoàng Văn Kiêm , Nguyễn Hoàng Tú Anh
Khoa CNTT, Trường Đại học Khoa học Tự Nhiên - ĐHQG tp.HCM

Tóm tắt :

Trong những năm gần đây, bài toán tìm kiếm luật kết hợp trên các CSDL lớn đã được quan tâm đặc biệt. Luật kết hợp rất hữu hiệu trong việc xác định mối liên hệ giữa các thuộc tính và có nhiều ứng dụng trong lĩnh vực thương mại, tài chính và buôn bán lẻ. Ngoài ra, luật kết hợp tối ưu là cách tiếp cận hiệu quả để tìm những đặc trưng thú vị nhất liên quan đến một số thuộc tính.

Trong bài báo này, chúng tôi tổng quát hóa bài toán tìm kiếm luật kết hợp có độ hỗ trợ tối ưu bằng cách chấp nhận các luật chứa các phép tuyển trên các thuộc tính số chưa được cụ thể hoá. Trong trường hợp luật chứa một thuộc tính số, chúng tôi dùng tiếp cận qui hoạch động để tính luật tối ưu. Trong trường hợp nhiều thuộc tính số, thuật giải di truyền đã được áp dụng kết hợp với tiếp cận qui hoạch động.

MINING OPTIMIZED SUPPORT RULES FOR NUMERIC ATTRIBUTES

Hoang Kiem, Nguyen Hoang Tu Anh
Department of Information Technology, University of Natural Sciences -
VNU.HCM

Abstract:

Mining association rules on large data sets has received considerable attention in recent years. Association rules are useful for determining correlations between attributes of a relation and have applications in marketing, financial and retail sectors. Furthermore, optimized association rules are an effective way to focus on the most interesting characteristics involving certain attributes.

In this paper, we generalize the optimized support association rule problem by permitting rules to contain disjunctions over uninstantiated numeric attributes. For rules containing a single numeric attribute, we use a dynamic programming approach for computing optimized association rules. We also use genetic algorithm with dynamic programming for many numeric attributes.