

# Sử dụng Bitmap trong khai thác tập có thể xóa trên dữ liệu tăng cường

NGUYỄN THỊ HOÀI LINH \*

Trường Đại học Kinh tế - Tài chính TP.HCM  
 Nhận bài: 04/08/2022 - Duyệt đăng: 15/10/2022

(\*) Liên hệ: [linhnth@uef.edu.vn](mailto:linhnth@uef.edu.vn); ĐT: 0945110590

**Tóm tắt:**

**K**hai thác tập có thể xóa là một trong những lĩnh vực thú vị trong khai thác tập phổ biến. Phân tích dữ liệu tăng cường trở nên quan trọng hơn vì dữ liệu thú vị liên tục được tích lũy trong các lĩnh vực ứng dụng khác nhau bao gồm cả các lĩnh vực công nghiệp. Bài viết này cung cấp một phương pháp tiếp cận sử dụng Bitmap để khai thác các tập có thể xóa trên dữ liệu tăng cường. Không giống như ý nghĩa Bitmap truyền thống cho các tập phổ biến, một Bitmap cho tập các tập có thể xóa được khai thác ở đây biểu thị mối quan hệ mà một sản phẩm bao gồm ít nhất một nguyên liệu (item) trong một vật phẩm. Sử dụng biểu diễn Bitmap có thể dễ dàng tìm thấy các sản phẩm mong muốn để kiểm tra, do đó giảm số lần quét cơ sở dữ liệu.

**Từ khóa:** Khai thác dữ liệu, Tập có thể xóa, Khai thác tăng dần, Biểu diễn Bitmap

**Abstract:**

Erasable itemset mining is one of the interesting areas in frequent set mining. Incremental data analytics becomes more important as interesting data is continuously accumulated in various application areas including industrial sectors. This article provides an approach using Bitmap to mine erasable itemset on incremental data. Unlike the traditional Bitmap meaning for frequent itemsets, a Bitmap for erasable itemset mining here denotes the relationship that a product includes at least one material (item) in a specified itemset. Using the Bitmap representation can easily find the desired products to check, thus reducing the number of database scans.

**Keywords:** Data mining, Erasable itemsets, Incremental mining, Bitmap representation.