

Abstract of thesis presented to the Senate of Universiti Malaysia Terengganu in fulfilment of the requirements for the degree of Doctor of Philosophy

**AN ENHANCEMENT OF INCREMENTAL ECLAT ALGORITHM FOR  
FREQUENT ITEMSET MINING**

**MAHADI MAN**

**2022**

**Main Supervisor : Associate Professor Masita @ Masila Abdul Jalil, Ph.D**

**Faculty : Faculty of Ocean Engineering Technology and  
Informatics**

Frequent itemset mining is an important subfield of association rule mining, which aims to discover interesting and useful patterns in a transaction database. The traditional task of frequent itemset mining is to find groups of items (itemsets) that frequently appear together in transactions made by customers. Although frequent itemset mining is designed for market basket analysis, it can be seen more generally in a variety of domains such as manufacturing, pharmaceuticals, and product recommendations. Due to its many applications, frequent itemset mining has become a popular research area and is widely studied among databases and the data mining community. Although existing algorithms can mine frequent patterns in datasets, long processing times and large memory usage are still two major problems faced by existing algorithms, especially when dealing with large databases. An algorithm called the Fast Incremental Equivalence Class Transformation or Fi-Eclat has been developed to perform frequent itemset mining more efficiently by reducing processing time and large memory usage. This thesis describes the frequent itemset mining algorithm implemented that produced promising results. Motivated by the fast intersection in the Incremental Equivalence Class Transformation (i-Eclat), this algorithm has been adopted through the *My Structured Query Language* (MySQL) database management system (DBMS) as its platform. It serves as a frequent itemset mining database engine

in testing Frequent Item Set Mining (FIMI) datasets from online repositories. The experimental results showed that the Fi-Eclat algorithm outperformed the i-Eclat algorithm in processing time performance by 42% in the tidset, 23% in the diffset, 14% in the sortdiffset, and 54% in the postdiffset. Thus, among the four (4) variants, the average processing time performance of the Fi-Eclat algorithm was concluded to be 33% better than the i-Eclat algorithm.

Abstrak tesis yang dikemukakan kepada Senat Universiti Malaysia Terengganu sebagai memenuhi keperluan untuk ijazah Doktor Falsafah Sains Komputer

**PENAMBAHBAIKAN ALGORITMA ECLAT TAMBAHAN BAGI  
PERLOMBONGAN SET ITEM YANG KERAP**

**MAHADI MAN**

**2022**

**Penyelia Utama : Profesor Madya Masita @ Masila Abdul Jalil, Ph.D**

**Fakulti : Fakulti Teknologi Kejuruteraan Kelautan dan Informatik**

Perlombongan set item yang kerap ialah subbidang penting dalam perlombongan peraturan persatuan yang bertujuan untuk mencari paten yang menarik dan berguna dalam transaksi pangkalan data. Tujuan asal perlombongan set item yang kerap adalah untuk mencari kumpulan item (set item) yang kerap muncul bersama dalam transaksi yang dibuat oleh pelanggan. Walaupun perlombongan set item yang kerap direka untuk analisis bakul pasaran, ia boleh diaplikasikan secara lebih umum dalam pelbagai domain seperti pembuatan, farmaseutikal dan pengesyoran produk. Disebabkan oleh aplikasinya yang berbagai, perlombongan set item yang kerap telah menjadi bidang penyelidikan yang popular dan dikaji secara meluas dalam kalangan komuniti pangkalan data dan komuniti perlombongan data. Walaupun algoritma sedia ada boleh melombong paten yang kerap dalam set data, namun masa pemprosesan yang panjang dan penggunaan memori yang besar masih merupakan dua kekangan utama yang dihadapi oleh algoritma tersebut terutamanya apabila berurusan dengan pangkalan data yang besar. Algoritma yang dipanggil Transformasi Kelas Kesetaraan Bertahap Pantas atau Fi-Eclat telah dibangunkan untuk melakukan perlombongan set item yang kerap dengan cara yang lebih cekap dengan mengurangkan masa pemprosesan dan penggunaan memori yang tinggi. Tesis ini menerangkan bagaimana algoritma perlombongan set item yang kerap dilaksanakan dan menghasilkan keputusan yang memberansangkan. Didorong oleh persilangan pantas dalam Transformasi Kelas

Kesetaraan Bertahap (i-Eclat), algoritma ini telah diadaptasikan melalui sistem pengurusan pangkalan data (DBMS) *My Structured Query Language* (MySQL) sebagai pelantarnya. Ia berfungsi sebagai enjin pangkalan data perlombongan set item kerap dalam menguji set data Perlombongan Set Item Kerap (FIMI) daripada repositori dalam talian. Keputusan eksperimen menunjukkan bahawa algoritma Fi-Eclat mengatasi prestasi algoritma i-Eclat sebanyak 42% dalam tidset, 23% dalam diffset, 14% dalam sortdiffset, dan 54% dalam postdiffset dari segi masa pemprosesan yang diambil. Oleh itu, antara empat (4) varian, purata prestasi algoritma Fi-Eclat dari sudut masa pemprosesan dapat disimpulkan 33% lebih baik daripada algoritma i-Eclat.