

ABSTRACT

Graph is an abstract concept that is fundamentally has been used, which allows to model the real world system. So did in data, any type of data can be modelled using the relationship of graph. Graph database was adopted to facilitate and assist in understanding, modeling, and analyzing a process. Graph database very suitable to use in unstructured and semi-structured data, compare to relational database which have a disadvantage that when increasing data that involving increasing possibility of joining between table. As well, the need to analyze and visualize large database. In application, the amount of data in graph database are rapidly growing to millions and event billions of nodes and edges so the computation cost to perform analysis and visualization of the graph can be very hard for current system capabilities.

To resolve these problem above, we need a method that can reduce the size of the graph without losing it important information, by implementing the Principle of Minimum Description Length Rissaenen's (MDL) with greedy to merge and combine it with the representation graph G which consists of Graph Summary and a set of Correction, then the resulting graph database is highly compressed.

Keywords: *graph database, graph summarization, graph representation, MDL principle, lossles, lossy, compression, greedy, Rissaenen's Minimum Description Length*