

## **ABSTRACT**

*A graph database is a database that stores data in the form of a graph. In a graph database, the data is stored in the form of nodes, and their relations are stored in the form of edges. This way, a search query through the database can be done in a traversal way, hence the time needed is relatively faster compared to relational database. But the way of storing data in nodes and edges can be expensive storage wise. This paper will explore a method to compress a graph database.*

*In this paper, Power Graph Analysis is used to compress graph databases that contain collaboration data between international journal authors. The method converts a graph into a power graph, which does not store data in nodes and edges. Rather, it stores data in Power Nodes and Power Edges that can form bicliques, cliques, and stars. Power nodes are clusters of nodes that have the same set of neighbors. They are created using Complete Linkage Hierarchical Clustering with Jaccard Index as its distance measure. Power nodes are linked together using power edges. Power edges are determined by greedy search to find power edges that abstract most edges first.*

*Result shows that changing a graph into a power graph can reduce the number of edges up to 82%, conversion rate up to 24,92; and database compression rate up to 74,17%.*

**Keywords:** *graph database, compression, power graph analysis, complete linkage hierarchical clustering, jaccard index, scientific journal authors*