Abstract

Today online news is something that very common among people of Indonesian. Currently online news data that has been stored in a storage reach billions data. To easier the search process, manipulation, and processing data the online news data needs to be modeled into graph model. And to facilitate online news readers then the online news data need to grouped by news content association. One method which can used to grouping news is graph clustering.

Before performing graph clustering, the online news data must converted into graph model. First, the online news data will perform preprocessing, then calculate the news content association using cosine similarity, after that result from cosine similarity normalized to be the edge that connects the online news documents. After the online news data become graph, next step is performing graph clustering. In this final project used Chinese Whispers algorithm for grouping news, because Chinese Whispers algorithm is able to cluster very large graph in comparatively short time. So its suitable for clustering online news.

In this final project has been tested the performance of Chinese Whispers algorithm in quality and accuracy of clustering result. After testing performed was shows that quality of cluster from Chinese Whispers algorithm have good quality because almost 95% vertex already have higher intra connectivity then inter connectivity, and the average accuracy of cluster result is 80%.

Keywords : Graph, Graph Database, Clustering, Graph Clustering, Chinese Whispers