

## Abstract

In the grouping (clustering) results, there are still obstacles often faced is determining the proper formulation of clustering. This is due to relationships between objects are highly variable. Determination of clustering is wrong can lead to poor performance of web search. Therefore we need a clustering method which will produce an optimal cluster. Optimal is viewed not only have similar documents in one cluster, but also has a data structure similar to the 'true cluster'.

Process optimization in clustering is a process that aims to find the global minimum value, where the global minimum is a similarity value of documents in each cluster. It needs a clustering algorithm based on cost function optimization. One algorithm that can solve the problem of finding the global minimum is Simulated Annealing.

In this final project will be implemented clustering optimization using Simulated Annealing Algorithm. In the process of optimization using Simulated Annealing uses K-Means algorithm as an initial cluster solution. Determination of K-Means Algorithm Simulated Annealing due to better implemented in partitionial clustering.

Tests on the clustering process in the optimization process is carried out to improve the quality of clusters generated by Algorithm Simulated Annealing when compared with the initial cluster (K-Means algorithm). It also tested the influence of each input parameter on the results of cluster quality. The result of cluster quality indicated by the value of Jaccard Criterion.

The result of clustering optimization with Simulated Annealing Algorithm is shown that the value of Jaccard Criterion cluster after process optimization using Simulated Annealing Algorithm, better than the cluster before undergoing a process of optimization. The number of documents on the test data also influences the quality of the cluster. In addition, input parameters ( $T_f$ ,  $T_0$ ,  $MaxIt$ ,  $\alpha$ ) is also great influence on the quality of clustering.

**Keywords:** Document clustering, *Simulated Annealing Algorithm, K-Means Algorithm, Jaccard Criterion*