

Abstract

Most search engines today use a system perengkingan the document search results. Thus obtained are so many documents that are not all relevant according to the user. Finally, a user must open one by one document to ascertain whether the document that the user wants. Therefore in this thesis author clustering method on the documents of search results to solve the problem.

Clustering method used is the Generalized Incremental method (GenIc). GenIc using Generalized incremental and evolutionary techniques in the process of grouping documents. The definition is a process of incremental Generalized shifted so that the distance center to center on the cluster of documents is getting closer. Meanwhile, evolutionary engineering is the process of updating center by calculating the probability of each center so that it can be determined where the center is able to "survive" or "murdered". To apply this method, there are several parameters that must be entered by the user ie the number of initial cluster (m), the final number of clusters (k), the number of generations (n), and the value of δ .

Tests conducted to analyzed the influence of four parameters on GenIc time performance and cluster quality based on hypotheses generated byRijbergen, documents relevant to a query that tend to have similarities to each other can be grouped in one same cluster.

Keywords: *clustering, GenIc algorithm, evolutionary, generalized incremental.*