Abstract

Information Retrieval (IR) is part of computer science related to the retrieval of information from documents that are based on the content and context of the documents themselves. Processes in Information Retrieval can be described as a process to obtain relevant documents from the collection of documents through search queries entered by users. Various approaches to improve the performance of Information Retrieval (IR) has been performed. One way to improve performance is to compress the index (index compression). Two types of compression techniques on Information Retrieval System is lossy compression and lossless compression. In lossless compression of all information will be maintained, for example, posting the file compression.

Index is the most important part in the Information Retrieval System. The aim is to save the index to optimize the speed and performance in finding relevant documents for the search query. Without indexes, search engines will scan every document, which will require much time and computing power. However, a growing number of incoming documents also increases the capacity of the index. Index Compression is a technique used to minimize the index, both of capacity and performance of Information Retrieval System. By compressing the index, can reduce the capacity used up to 75%. Index Compression can also increase transfer speed from disk to memory.

In this thesis, will be implemented using Variable Compression Index Byte Code. Variable byte code is one of the techniques applied in the compression index on Information Retrieval in order to reduce the unused disk capacity and usage of a faster search time. Therefore it is expected that after using the Index Compression using Variable Byte Code, the capacity will be reduced and the performance index of the Information Retrieval System to increase.

.

Keywords**:** *Information Retrieval, Information Retrieval System, Index Compression, lossless compression, Variable Byte Code.*