

Abstract

The abundant amount of digital information in the form of text does not mean the searching and the identification of the needed document can be done easily. One way to identify a document is by keywords. However, not all documents are provided with keywords.

Keyword extraction is a system to automatically extract a set of keywords from text. Keyword extraction built for this final work implements TextRank, a graph-based algorithm. TextRank for keyword extraction has 3 input parameters, which are graph type, word filter, and window size.

Testing was done by changing the setting of TextRank parameters and comparing system's keywords with human's keyword. The experiment was also done using various damping factor and threshold values. Testing parameter used in this final work are recall. From testing, the using of *undirected graph* and verb filter results in better testing parameters' values in almost sizes of window than other using of graphs and word filters. In general, recall increases when damping value is bigger although it's not significant and recall is not really influenced by threshold value.

Keywords: keyword, keyword extraction, textrank