Information retrieval is the most common objective of internet users. For a given query, existing search engines will return a long list of search results ranked by their relevance to the given query. In most of the existing search engines, Page Rank algorithm is used to find the relevance of the documents to the given query. The problem with this method is that it considers only the incoming and outgoing links and the popularity of the document. The proposed method eliminates this problem by considering occurrence of the given term in domain name, URL, TITLE tag, META tag and inside the document in addition to the popularity of the document. It makes the relevant documents to have higher ranking.

Reference

- Mei Q, Shen X and Zhai C, 2007, “Automatic Labeling of multinomial topic models”. In Proceeding of the Knowledge Discovery and Data Mining Conference (KDD’07).

Index Terms

Computer Science Web Mining

Key words

Page Rank

Term frequency

Domain name

URL