

{tag} International Journal of Computer Applications
Foundation of Computer Science (FCS), NY, USA

[Volume 144](#)

-
[Number 12](#)

Year of Publication: 2016

Authors:

Amit Rajeshwarkar, Meghana Nagori

10.5120/ijca2016910498

{bibtex}2016910498.bib{/bibtex}

Abstract

The Triangular Search approach aims at recalculating authenticity of the Search Results provided by the Google API with the help of Semantic similarity provided by Wikipedia API and calculating the cosine similarity between the Document Vectors and query string vectors using enhanced approach of Tf-Idf that reduces calculation involved in it.

The Search Engine Optimization traces anchor texts that are the values between a tag of HTML and body texts of a web page. Using the Vector Space Model, the Term frequency and Inverse document frequency are calculated along with the Page ranking algorithm to get the Search Results. But consideration of anchor texts in search engine optimization techniques leads to some of the non-relevant body texts of a document. Also the top results of a search engine include trending and e-commerce links other than sponsored links but the intent of search is not considered.

This approach proposes and gains user intents behind the search thereby focusing on providing

intent related search results.

References

1. Yajun Du, Wenjun Liu, et al. 2015. "An improved focused crawler based on Semantic Similarity Vector Space Model". The Official Journal of the World Federation on Soft Computing (WFSC) 36, Elsevier, 392–407.
2. What Is Search Engine Optimization / SEO. 2011. Common Craft, Search Engine Land. YouTube. We. 12 Sep.2011.
3. Masumi Shirakawa, Kotaro Nakayama, et al.2015 "Wikipedia based Semantic Similarity Measurements for Noisy Short Texts Using Extended Naive Bayes". IEEE Transactions on Emerging Topics in Computing, DOI 10.1109/TETC.2418716.
4. Z. Yun-tao, et al., 2005. "An improved TF-IDF approach for text classification". Journal of Zhejiang University SCIENCE, 6A(1):49-55.
5. H. A. Haddadene, et al., 2012."On the Pagerank Algorithm for the Articles Ranking". Proceedings of the World Congress on Engineering, Vol I July 4 - 6, 2012.
6. E.Gabrilovich et al., 2007. "Computing semantic relatedness using Wikipedia based Explicit Semantic Analysis". IJCAI, 1606-1611.
7. A.Hliaoutakis, G.Varelas, et.al.,2006 "Information retrieval by semantic similarity", I. J. Semant. WebInf. Syst. 3 (3) 55–73.
8. Budanitsky, Hirst et al. 2006. "Evaluating WordNet based Measures of Lexical Semantic Relatedness". Computational Linguistics: ACM Digital Lib., 32(1), 13-47.

Index Terms

Computer Science

Information Sciences

Keywords

Google API, Wikipedia API, Explicit Semantic Analysis (ESA), Enhanced TF-IDF.