

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

[Volume 170](#)

-  
[Number 2](#)

Year of Publication: 2017

Authors:

Lobo L. M. R. J, R. S. Bichkar

10.5120/ijca2017914698

{bibtex}2017914698.bib{/bibtex}

### **Abstract**

A large amount of information is available on the web. Generating relevant information from the web for a user has become a question of great concern. This information is made available to the user by search engines as per the query given by a user. Search engines return pages depending on a ranking algorithm based on links, to and from the page and on how popular a page is, with respect to the hits received by users. In most cases the pages returned are too many and irrelevant. It is therefore necessary to use a technique that caters to relevance of the page returned with respect to the query fired by a user. A system associated with use of synonyms of terms in the query has proved to be useful. A combination of such synonyms fired to a search engine has returned relevant information pages. In most cases it has also generated a better page than an individual search engine.

The present paper reflects on how synonyms of terms generated from the query are crossed over and fired to the search engine to generate more relevant semantic associated pages. These pages are then tested against the pages returned by the individual search engine with

use of the original query using a standard page ranking tool. The pages are also examined for relevance to specific users and usefulness of content to a specific domain. The pages are examined for their positions using ranking tools, trustworthiness tools and intent drifting. It is found that the pages returned using the method of combining synonyms of terms of the user query are placed at better ranking positions. An analysis of the pages returned also indicates relevance to user, usefulness of content to specific domains and possibility of intent drift.

### References

1. S. Brin and L. Page, "The anatomy of a large-scale hyper-textual web search engine," in *Computer Networks and ISDN Systems*. Elsevier Science Publishers B. V., 1998, pp. 107–117.
2. J. Hou and Y. Zhang, "Effectively finding relevant web pages from linkage information," in *IEEE Transactions On Knowledge And Data Engineering*, Vol. 15, No. 4, 2003.
3. M. J. Martin-Bautista and M.-A. Vila, "A survey of genetic feature selection in mining issues," in *Proceedings of the 1999 Congress on Evolutionary Computation, CEC 99, Vol 3*, IEEE, 1999, pp. 1314– 1321.
4. M. J. Martin-Bautista, M.-A. Vila, and H. L. Larsen, "Building adaptive user profiles by a genetic fuzzy classifier with feature selection." *IEEE*, 2000, pp. 308–312.
5. H. Chela, Y.-M. Chung, M. Lamsey, C. C. Yang, P.-C. Ma, and J. Yen, "Intelligent spider for internet searching," in *Thirtieth IEEE Hawaii International Conference on System Sciences*, 1997.
6. N. Tomca, "A flexible tool for jaccard score evaluation," in *B. Sc. Thesis, University of Belgrade, Belgrade, Serbia, Yugoslavia*, 1997.
7. Mahbub, "Genetic algorithm in adaptive web search," in *Filed under Research*, 2007.
8. M. Richardson, A. Prakash, and E. Brill, "Beyond page rank: Machine learning for static ranking," in *Proceedings of the International Conference on World Wide Web*, 2006.
9. M. Tsukada, T. Washio, and H. Motoda, "Automatic web-page classification by using machine learning methods," in *Web Intelligence: Research and Development, LNAI*. Springer-Verlag, 2001, pp. 303–313.
10. F. Ciravegna, A. Lavelli, D. Petrelli, and F. Pianesi, "The geppetto development environment - version 2.1 - user manual," *Tech. Rep.*, 1997.
11. M. B. Jadhav and B. M. Patil, "File annotation and sharing on low end devices in pan," in *IJCA Journal Volume 83 - Number 13*, 2013.
12. J. B. Filho and J. Gensel, "A contextual annotation-based access control model for pervasive environments," in *Second International Workshop on Security and Privacy in Spontaneous Interaction and Mobile Phone Use (IWSSI/SPMU)*, 2010.
13. C. A. N. Soules and G. R. Ganger, "A contextual annotation-based access control model for pervasive environment : why can't i find my files? New methods for automating attribute assignments," in *Proceedings of the 9th conference on Hot Topics in Operating Systems HOTOS'03 - Volume 9 Pages 20-20*, 2003.
14. M. J. Carman, M. Baillie, and F. Crestani, "Tag data and personalized information retrieval," in *In Proceedings of the CIKM workshop on Search in social media. ACM*, 2008, pp. 27–34.
15. R. Jsckke, R. Marinho, A. Hotho, L. Schmidt-thieme, and G. Stumme, "Tag recommendations in folks anomies," in *PKDD, LNAI 4702, Springer*, 2007, pp. 506–514.
16. G. Bharathi and D. Venkatesan, "Improving information retrieval using document

clusters and semantic synonym extraction,” in *Journal of Theoretical and Applied Information Technology* February 2012. Vol. 36 No.2 ISSN: 1992-8645), 2012.

17. P. Choudhary, “A comparative analysis of various web search engines,” in *International Journal of Computing and Business Research (IJCBR)* ISSN (Online): 2229-6166 Vol. 3 Issue 2, 2012.

18. G. Madhu, “Intelligent semantic web search engines: A brief survey,” in *International journal of Web & Semantic Technology (IJWest)* Vol.2, No.1, 2011.

19. R. K. Roul and S. K. Sahay, “An effective information retrieval for ambiguous query,” in *arXiv : 1204.1406v1 [cs.IR]*, 2012.

20. J. Beel and E. Wilde, “Academic search engine optimization (aseo): Optimizing scholarly literature for google scholar & co.” in *Journal of Scholarly Publishing*, 41 (2):, 2010, pp. 176–190.

21. H. Ishkewy and H. Harb, “Iswse: Islamic semantic web search engine,” in *International Journal of Computer Applications (0975 8887)* Volume 112 No. 5, 2015.

22. J. Beel and B. Gipp, “Google scholars ranking algorithm: An introductory overview,” in *12th International Conference on Scientometrics and Informatics (ISSI09)*, volume 1, Rio de Janeiro (Brazil), 2009, pp. 230–241.

23. X. Wei, F. Peng, H. Tseng, Y. Lu, X. Wang, and B. Dumoulin, “Search with synonyms: Problems and solutions,” in *coling 2010: Poster Volume*, Beijing, 2010, pp. 1318–1326.

24. A. Hliaoutakis, G. Varelas, E. Voutsakis, E. G. M. Petrakis, and E. Milios, “Information retrieval by semantic similarity,” in *Intern. Journal on Semantic Web and Information Systems (IJSWIS)*, 3(3):5573, July/Sept. 2006. Special Issue of Multimedia Semantics, 2006.

25. Y. Li, “Search with synonyms: Problems and solutions,” in *IEEE Internet Computing* 1089-7801/98, 1998.

26. G. P. Sudhakar and R. Kumar, “Content based ranking for search engines,” in *International Multi-Conference of Engineers and Computer Scientists Vol I*, 2012.

27. H. Cui, J.-R. Wen, J.-Y. Nie, and W.-Y. Ma, “Query expansion by mining user logs,” in *IEEE TRANSACTION ON KNOWLEDGE AND DATA ENGINEERING*, vol. 15, no. 4, pp. 829–839, 2003.

28. K. Chakrabarti, M. Ortega, K. Porkaew, and S. Mehrotra, “Query refinement in similarity retrieval systems,” in *Bulletin of the Technical Committee on Data Engineering Vol. 24 No. 3 IEEE Computer Society*, 2001.

29. X. Tang, K. Liu, J. Cui, S. Member, F. Wen, and X. Wang, “Intent search: capturing user intention for one-click internet image search,” in *IEEE Transactions On Pattern Analysis And Machine Intelligence*, Vol. 34, No. 7, 2012.

30. L. Hasan and E. Abuelrub, “Assessing the quality of web sites,” in *INFOCOMP Vol 7 DOI: 10.1016/j.aci.2009.03.001*, 2008.

31. L. M. R. J Lobo and R. S Bichkar, “Finding the best page using synonyms,” in *International Journal of Computer Applications*, Volume 65 No.8, 2013.

### Index Terms

Computer Science

Information Sciences

## Keywords

Best Page, relevance, users' interest, synonyms, ranking.