

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

[Volume 144](#)

-
[Number 12](#)

Year of Publication: 2016

Authors:

Sunil D. Jejurkar, Vivek P. Kshirsagar

10.5120/ijca2016910503

{bibtex}2016910503.bib{/bibtex}

Abstract

In Today's world, with the increased use of internet the large volume of data is stored on World Wide Web. To use this large data the different search engines are provided. But the accuracy of the data is again based on the appropriate search query submitted by the user to search engine. Depending on the search query the search engine retrieves the massive amount of relevant data by using different algorithms such as page rank algorithm or relevancy algorithm. Further, the returned results decide the performance as well as the efficiency of the search engine. Search result clustering problem means clustering the search results returned by the search engine.

In this paper a comparative analysis of Suffix Tree Clustering algorithms is done to decide the how accurately it clusters the search results i.e. an empirical analysis which is done by using standard datasets.

References

1. Oren Zamir and Oren Etzioni. Document Clustering: A Feasibility Demonstration. Proceedings of the 19th International ACM SIGIR Conference on Research and Development of Information Retrieval, 1998, pp 46-54.
2. Oren Zamir and Oren Etzioni. Grouper: A Dynamic Clustering Interface to Web Search Results. WWW8/Computer Networks, Amsterdam, Netherlands, 1999.
3. Oren E. Zamir. Clustering Web Documents: A Phrase-Based Method for Grouping Search Engine Results. Doctoral Dissertation, University of Washington, 1999.
4. Scatter/gather a cluster based approach to browsing large document collections. Douglassr cutting,David R.Karger ,Jan O Pederson,15 annual International SIGIR 92,ACM 0-89791-542-0912/0006/0318.
5. Antonio Di Marco and Roberto Navigli, Clustering Web Search Results with Maximum Spanning Trees other publication details.
6. Ke,W., Sugimoto, C.R., Mostafa, J.: Dynamicity vs. effectiveness: studying online clustering for scatter/gather. In: Proc. of SIGIR 2009, MA, USA, 2009, pp. 19–26.
7. Carpineto, C., Osinski, S., Romano, G.,Weiss, D.: A survey of web clustering engines. ACM Computing Surveys 41(3), 2009, pp. 1–38.
8. Kamvar, M., Baluja, S.: A large scale study of wireless search behavior: Google mobile search. In: Proc. of CHI 2006, New York, NY, USA, 2006, pp. 701–709.
9. Osinski, S., Weiss, D.: A concept-driven algorithm for clustering search results. IEEE Intelligent Systems 20(3), 2005, 48–54.
10. Sanderson, M.: Ambiguous queries: test collections need more sense. In: Proc. of SIGIR 2008, Singapore, 2008, pp. 499–506.
11. Chen, J., Zaïane, O.R., Goebel, R.: An unsupervised approach to cluster web search results based on word sense communities. In: Proc. Of WI-IAT 2008, Sydney, Australia, (2008), pp. 725–729.
12. Zhang, X., Hu, X., Zhou, X.: A comparative evaluation of different link types on enhancing document clustering. In: Proc. of SIGIR 2008, Singapore, 2008, pp. 555–562.
13. iBoogie – meta search engine with automatic document clustering. <http://www.iBoogie.tv/>.
14. Robert Navigli and Giuseppe Crisafulli department of Informatics,Rome,Proceedings of the 2012 Conference on EMpherical Methods in Natural Language Processing,Pg 116-126 MIT,USA OCT9-11 2010 @)ACL.

Index Terms

Computer Science

Algorithms

Keywords

Suffix Tree Clustering, Search Results.