

{tag}

{/tag}

International Journal of Computer Applications

© 2015 by IJCA Journal

Volume 122 - Number 9

Year of Publication: 2015

Authors:

Suruchi Chawla

10.5120/21726-4885

{bibtex}pxc3904885.bib{/bibtex}

Abstract

Personalization of web search is used for effective Information Retrieval in order to better satisfy the information need of the user on the web. The web usage mining has been used widely in Personalization of Web Search(PWS). The effectiveness of the Personalization of Web Search based on clustered web usage data depends on the quality of clusters. It is found in research that there exist no clustering algorithms that produce clusters of 100% quality. In this paper the Genetic Algorithm(GA) is used for clusters optimization in order to improve the quality of clusters for effective Personalized web search. Experiment was conducted on the data set of query sessions captured on the web in Academics, Entertainment and Sports Domain. The search results confirm the improvement in the average precision of the PWS(with cluster optimization) in comparison to PWS(without cluster optimization).

Refer

ences

- Akhlaghian, F. , Arzanian, B. , and Moradi, P. 2010. "A Personalized Search Engine Using Ontology-Based Fuzzy Concept Networks, International Conference on Data Storage and Data Engineering, pp. 137 – 141.
- Arzanian, B. , Akhlaghian, F. , and Moradi, P. 2010. A Multi- Agent Based Personalized Meta-Search Engine Using Automatic Fuzzy Concept Networks, Third International Conference

on Knowledge Discovery and Data Mining, pp. 208 – 211.

- Boughanem, M. , Chriment, C. , Mothe, J. , Dupuy, C. S. , and Tamine, L. 2000. Connectionist and genetic approaches for information retrieval, *Soft Computing in Information Retrieval Studies in Fuzziness and Soft Computing*, 50, pp. 173–198.
- Boughanem, M. , Chriment, C. and Tamine, L. 2002. On using genetic algorithms for multimodal relevance optimization in information retrieval, *Journal of the American Society for Information Science and Technology*, 53(11) , pp. 934–942.
- Bremermann, H. J. 1958. The evolution of intelligence. The nervous system as a model of its environment, Technical Report No. 1, Department of Mathematics, University of Washington, Seattle, WA.
- Chawla, S. , and Bedi P. 2007. Personalized Web Search using Information Scent, *International Joint Conferences on Computer, Information and Systems Sciences, and Engineering*, Technically Co-Sponsored by: Institute of Electrical & Electronics Engineers (IEEE), University of Bridgeport, published in LNCS (Springer), pp. 483-488.
- Chawla, S. , and Bedi, P. 2008. Improving information retrieval precision by finding related queries with similar information need using information scent. In *First International Conference on Emerging Trends in Engineering and Technology, ICETET'08*. (pp. 486-491). IEEE.
- Chawla, S. 2012a. Trust in Personalized Web Search based on Clustered Query Sessions. *International Journal of Computer Applications*, 59(7), pp. 36-44.
- Chawla, S. 2012b. Semantic Query Expansion using Cluster Based Domain Ontologies. *International Journal of Information Retrieval Research (IJIRR)*, 2(2), pp. 13-28.
- Chawla, S. 2013. Personalised web search using ACO with information scent. *International Journal of Knowledge and Web Intelligence*, 4(2), pp. 238-259.
- Chawla, S. 2014a. Personalized Web Search using Trust based Hubs and Authorities. *International Journal of Engineering Research and Applications*, 4(7), pp. 157-170.
- Chawla, S. 2014b. Novel Approach to Query Expansion using Genetic Algorithm on Clustered Query Sessions for Effective Personalized Web Search . *International Journal of Advanced Research in Computer Science and Software Engineering*, 4(11), pp. 73-81.
- Chi, E H. , Pirolli, P. , Chen, K. , and Pitkow, J. 2001. Using Information Scent to model User Information Needs and Actions on the Web, *International Conference on Human Factors in Computing Systems*, New York, NY, USA, pp. 490-497.
- Crestani, F. , and Pasi, G. 2000. *Soft Computing in Information Retrieval: Techniques and Application*, 50, Heidelberg, Germany: Physica-Verlag.
- Goldberg, D. E. 1989. *Genetic Algorithms in Search, Optimization and Machine Learning*, Addison-Wesley Longman Publishing Co. , Boston, MA, USA.
- Gordon. M. 1988. Probabilistic and genetic algorithms in document retrieval, *Communications of the ACM* , 31(10) , pp. 1208–1218.
- Heer, J. , and Chi, E. H. 2002. Separating the Swarm: Categorization method for user sessions on the web, *International Conference on Human Factor in Computing Systems*, pp. 243-250.
- Horng, J. T. , and Yeh, C. C. 2000. Applying genetic algorithms to query optimization in document retrieval, *Information Processing & Management*, 36(5) , pp. 737–759.
- Kanungo, T. , Mount, D. M. , Netanyahu, Nathan S. , Silverman, Christine D. Piatko, Ruth and Wu, A. Y. 2002. "An Efficient k- Means Clustering Algorithm: Analysis and Implementation. " *IEEE Transactions on pattern analysis and machine intelligence*,

24(7), 881-892.

- Kim, S. , and Zhang, B. T. 2000. Web document retrieval by genetic learning of importance factors for html tags, International Workshop on Text and Web Mining, Melbourne, Australia, pp. 13–23.
- Kim, H. , Lee, S. , Lee, B. , and Kang, S. 2010. Building Concept Network-Based User Profile for Personalized Web Search, 9th International Conference on Computer and Information Science ,pp. 567 – 572.
- Leung, K. W. -T. , Ng, W. , and Lee, D. L. 2008. Personalized Concept-Based Clustering of Search Engine Queries, Journal IEEE Transactions on Knowledge and Data Engineering, 20(11), pp. 1505 – 1518.
- Liu, F. , Yu,, C. , and Meng, W. 2004. Personalized Web search for improving retrieval effectiveness, Journal IEEE Transactions on Knowledge and Data Engineering, 16(1), pp. 28 – 40.
- Loia , V. , and Luongo, P. 2001. An evolutionary approach to automatic web page categorization and updating, Conference on Web Intelligence: Research and Development, Springer-Verlag, pp. 292–302.
- Nasraoui, O. , and Petenes, C. 2003. Combining Web Usage Mining and Fuzzy Inference for Website Personalization, International Conference on Knowledge Discovery and Data Mining, pp. 37-46.
- Navrat, P. , Kovacik, M. , Ezzeddine, A. B. , Rozinajova, V. 2008. Web search engine working as a bee hive, Journal Web Intelligence and Agent Systems, 6(4), pp. 441–452.
- Pal, S. K. , Talwar, V , & Mitra, P. 2002. Web Mining in Soft Computing Framework: Relevance, State of the Art and Future Directions, IEEE Transactions on Neural Networks, 13(5), pp. 1163-1177.
- Peng, Wen-Chih, and Lin, Yu-Chin. 2006. Ranking Web Search Results from Personalized Perspective, The 8th IEEE International Conference on E-Commerce Technology and The 3rd IEEE International Conference on Enterprise Computing, E-Commerce, and E-Services, pp. 12.
- Pirolli. P. 1997. Computational models of information scent-following in a very large browsable text collection, Conference on Human Factors in Computing Systems, pp. 3-10.
- Pirolli, P. 2004. The use of proximal information scent to forage for distal content on the world wide web, Working with Technology, Mind: Brunswikian. Resources for Cognitive Science and Engineering, Oxford University Press.
- Rekha, C. , Sujatha, N. , and Iyakutti, K. . 2011 ,Algorithm to Improve the Cluster Quality using Genetic Algorithm, Research Journal of Computer Systems Engineering-RJCSE, 2(4).
- Smith, M. P, and, Smith, M. 1997. The use of genetic programming to build Boolean queries for text retrieval through relevance feedback, Journal of Information Science, 23 (6), pp. 423–431.
- Tamine, L. , Chrisment, C. , and Boughanem, M. 2003. Multiple query evaluation based on an enhanced genetic algorithm, Information Processing and Management , 39(2), pp. 215–231.
- Vrajitoru, D. 1998. Crossover improvement for the genetic algorithm in information retrieval, Information Processing & Management, 34(4), pp. 405–415.
- Wen, J. R. , Nie, J. Y, and Zhang, H. J. 2002. Query Clustering Using User

Logan, Journal ACM Transactions on Information Systems, 20(1), pp. 59-81.

- Wu, X. , Kumar, V. , Quinlan, J. R. , Ghosh, J. , Yang, Q. , Motoda, H. , McLachlan, G. J. , Ng, A. , Liu, B. , Yu, P. S. , Zhou, Z. -H. , Steinbach, M. , Hand D. J. ,and Steinberg, D. 2007 "Top 10 Algorithms in Data Mining", Knowledge Information Systems, 14(1) pp. 1-37.

- Yang, J. , Korfhage , R. R. , and Rasmussen, E. 1992. Query improvement in information retrieval using genetic algorithms—a report on the experiments of the TREC project, 1st text retrieval conference (TREC-1), pp. 31–58.

- Zhao, Y and Karypis, G. 2001. Criterion functions for document clustering: Experiments and Analysis. Technical report, University of Minnesota, Minneapolis, MN.

- Zhao, Y. and Karypis, G. 2002. Comparison of agglomerative and partitional document clustering algorithms, SIAM Workshop on Clustering High-dimensional Data and its Applications.

- Zhu, Z , Xu, J. , Ren, X. , Tian, Y. and Li, L. 2007. Query Expansion Based on a Personalized Web Search Model, Third International Conference on Semantics, Knowledge and Grid, pp. 128 – 133.

Index Terms

Computer Science

Software Engineering

Keywords

Web Information Retrieval Personalized Web Search Genetic Algorithms
Clustering

Optimization

Information Science

