



Search bar with magnifying glass icon and settings gear icon

Home Contact Us

Look Inside Get Access

Find out how to access preview-only content

Chapter

Advances in Knowledge Discovery and Data Mining

Volume 5476 of the series Lecture Notes in Computer Science pp 254-265

# Quantifying Asymmetric Semantic Relations from Query Logs by Resource Allocation

Zhiyuan Liu, Yabin Zheng, Maosong Sun



Buy chapter

Buy this eBook

Get Access

\* Final gross prices may vary according to local VAT.

## Abstract

In this paper we present a bipartite-network-based resource allocation(BNRA) method to extract and quantify semantic relations from large scale query logs of search engine. Firstly, we construct a query-URL bipartite network from query logs of search engine. By BNRA, we extract asymmetric semantic relations between queries from the bipartite network. Asymmetric relation indicates that two related queries could be assigned different semantic relevance strength against each other, which is more conforming to reality. We verify the validity of the method with query logs from Chinese search engine Sogou. It demonstrates BNRA could effectively quantify semantic relations from We further construct query semantic networks, and introduce several measures to analyze the networks. BNRA is not only 'language oblivious' and 'content oblivious', but could also be easily implemented in a paralleled manner, which provides commercial search

Other

- » Ex
» Ab
» Re
» Pe
» Ad

Zhiyuan Liu, Yabin Zheng, and Maosong Sun
Department of Computer Science and Technology,
State Key Lab on Intelligent Technology and Systems,
National Lab for Information Science and Technology,
Tsinghua University, Beijing, China, 100084
(liuzhiyuan, zhenyabin, sunms@sempringtan.edu.cn
http://dx.doi.org/10.1007/978-3-642-03429-9\_15

Abstract. In this paper we present a bipartite-network-based resource allocation(BNRA) method to extract and quantify semantic relations from large scale query logs of search engine. Firstly, we construct a query-URL bipartite network from query logs of search engine. By BNRA, we extract asymmetric semantic relations between queries from the bipartite network. Asymmetric relation indicates that two related queries could be assigned different semantic relevance strength against each other, which is more conforming to reality. We verify the validity of the method with query logs from Chinese search engine Sogou. It demonstrates BNRA could effectively quantify semantic relations from We further construct query semantic networks, and introduce several measures to analyze the networks. BNRA is not only 'language oblivious' and 'content oblivious', but could also be easily implemented in a paralleled manner, which provides commercial search engine

Keywords: Semantic relations, asymmetric.

### 1 Introduction

With the development of Internet, the most important tool to get information is not a perfect method to find the relevance using keywords in documents and queries. The most important factors that affect the performance of search engines.

However, with pages returned from search engines are not always relevant to search intentions of users. An independent survey of 40,000 web users found that after a failed search, 70% of them will try to rephrase their queries on the same search engine instead of resorting to a different one [1]. Therefore, it is a non-trivial task for search engines to find better query representation of user search intentions in order to enhance search performance.

Y. Theodoropoulos et al. (Eds.), PAKDD 2009, LNCS 5476, pp. 254–265, 2009. © Springer-Verlag Berlin Heidelberg 2009

Look Inside

## Keywords

Semantic relations query log resource allocation asymmetric

### Supplementary Material (0)

### References (24)

## References

1. Beeferman, D., Berger, A.: *Agglomerative clustering of a search engine query log*. In: *Proceedings of the sixth ACM SIGKDD international conference on knowledge discovery and data mining (2000)*
2. Baeza-Yates, R., Tiberi, A.: *Extracting semantic relations from query logs*. In: *Proceedings of the 13th ACM SIGKDD international conference on knowledge discovery and data mining (2007)*
3. Wen, J.R., Jian-Yun, N., Hong-Jiang, Z.: *Query clustering using user logs*. *ACM Transactions on Information Systems* 20(1) (2002)
4. Shen, D., Pan, R., Sun, J.T., Pan, J.J., Wu, K., Yin, J., Yang, Q.: *Query enrichment for web-query classification*. *ACM Transactions on Information Systems* 24(3), 320–352 (2006)  
[CrossRef](#)
5. Beitzel, S.M., Jensen, E.C., Lewis, D.D., Chowdhury, A., Frieder, O.: *Automatic classification of web queries using very large unlabeled query logs*. *ACM Transactions on Information Systems* 25(2), 9 (2007)  
[CrossRef](#)
6. Baeza-Yates, R., Hurtado, C., Mendoza, M.: *Query recommendation using query logs in search engines*. In: *Workshops on current trends in database technology of 9th international conference on extending database technology (2004)*
7. Chirita, P.A., Firan, C.S., Nejdl, W.: *Personalized query expansion for the web*. In: *Proceedings of the 30th annual international ACM SIGIR conference on research and development in information retrieval*, pp. 7–14 (2007)
8. He, X.F., Yan, J., Ma, J.W., Liu, N., Chen, Z.: *Query topic detection for reformulation*. In: *Proceedings of the 16th international conference on World Wide Web*, pp. 1187–1188 (2007)
9. Zhou, T., Ren, J., Medo, M., Zhang, Y.C.: *Bipartite network projection and personal recommendation*. *Physical Review E* 76(4) (2007)
10. Deshpande, M., Karypis, G.: *Item-based top-n recommendation algorithms*. *ACM Transactions on Information Systems* 22(1) (2004)
11. Liu, Z.Y., Sun, M.S.: *Asymmetrical query recommendation method based on bipartite network resource allocation*. In: *Proceedings of the 17th international conference on World Wide Web, Beijing (2008)*
12. Raghavan, V.V., Sever, H.: *On the reuse of past optimal queries*. In: *Proceedings of the 18th annual international ACM SIGIR conference on research and development in information retrieval*, pp. 344–350 (1995)
13. Fitzpatrick, L., Dent, M.: *Automatic feedback using past queries: social searching?* In: *Proceedings of the 20th annual international ACM SIGIR conference on research and development in information retrieval*, pp. 306–313 (1997)
14. Baeza-Yates, R., Hurtado, C., Mendoza, M.: *Query clustering for boosting web page ranking*. In: *Advances in Web Intelligence*, pp. 164–175 (2004)
15. Sahami, M., Heilman, T.D.: *A web-based kernel function for measuring the similarity of short text snippets*. In: *Proceedings of the 15th international conference on World Wide Web*, pp. 377–386 (2006)

16. Fonseca, B.M., Golgher, P.B., de Moura, E.S., Ziviani, N.: Using association rules to discover search engines related queries. In: *Proceedings of the first conference on Latin American Web Congress*, pp. 66–71 (2003)
17. Broder, A.: A taxonomy of web search. *ACM SIGIR Forum* 36(2), 3–10 (2002)  
[CrossRef](#)
18. Kang, I.H., Kim, G.C.: Query type classification for web document retrieval. In: *Proceedings of the 26th annual international ACM SIGIR conference on research and development in informaion retrieval*, pp. 64–71 (2003)
19. Gravano, L., Hatzivassiloglou, V., Lichtenstein, R.: Categorizing web queries according to geographical locality. In: *Proceedings of the 12th international conference on information and knowledge management*, pp. 325–333 (2003)
20. Baeza-Yates, R.: Graphs from search engine queries. In: van Leeuwen, J., Italiano, G.F., van der Hoek, W., Meinel, C., Sack, H., Plá šl, F. (eds.) *SOFSEM 2007. LNCS*, vol. 4362, pp. 1–8. Springer, Heidelberg (2007)  
[CrossRef](#)
21. Zhou, T., Jiang, L.L., Su, R.Q., Zhang, Y.C.: Effect of initial configuration on network-based recommendation. *Europhysics Letters* 81(5), 58004 (2008)  
[CrossRef](#)
22. Ross, S.M.: *Introduction to Probability Models*, 9th edn. Academic Press, Inc., Orlando (2006)  
[MATH](#)
23. Kapp, A.V., Tibshirani, R.: Are clusters found in one dataset present in another dataset? *Biostatistics* 8(1), 9–31 (2007)  
[MATH](#) [CrossRef](#)
24. Newman, M.E.J.: The structure and function of complex networks. *SIAM Review* 45(2), 167–256 (2003)  
[MATH](#) [CrossRef](#) [MathSciNet](#)

## About this Chapter

### Title

Quantifying Asymmetric  
Semantic Relations from Query  
Logs by Resource Allocation

### Book Title

[Advances in Knowledge  
Discovery and Data Mining](#)

### Book Subtitle

13th Pacific-Asia Conference,  
PAKDD 2009 Bangkok, Thailand,  
April 27-30, 2009 Proceedings

### Pages

pp 254-265

### Copyright

2009

### DOI

10.1007/978-3-642-01307-2\_25

### Print ISBN

### Topics

[Artificial Intelligence \(incl.  
Robotics\)](#)  
[Data Mining and Knowledge  
Discovery](#)  
[Information Storage and Retrieval](#)  
[Probability and Statistics in  
Computer Science](#)  
[Multimedia Information Systems](#)  
[Computer Appl. in Administrative  
Data Processing](#)

### Keywords



Semantic relations  
query log  
resource allocation  
asymmetric

### Industry Sectors

[Electronics](#)  
[Telecommunications](#)  
[IT & Software](#)

### eBook Packages

### Editors

[Thanaruk Theeramunkong](#)  <sup>(19)</sup>  
[Boonserm Kijsirikul](#)  <sup>(20)</sup>  
[Nick Cercone](#)  <sup>(21)</sup>  
[Tu-Bao Ho](#)  <sup>(22)</sup>

### Editor Affiliations

19. Sirindhorn International  
Institute of Technology,  
Thammasat University  
20. Dept. of Computer  
Engineering, Faculty of  
Engineering, Chulalongkorn  
University  
21. Faculty of Science &  
Engineering, York University  
22. School of Knowledge  
Science, Japan Advanced  
Institute of Science and  
Technology

### Authors

[Zhiyuan Liu](#)  <sup>(23)</sup>  
[Yabin Zheng](#)  <sup>(23)</sup>  
[Maosong Sun](#)  <sup>(23)</sup>

978-3-642-01306-5

Computer Science  
eBook Package english full  
Collection

### Author Affiliations

23. Department of Computer  
Science and Technology, State  
Key Lab on Intelligent Technology  
and Systems, National Lab for  
Information Science and  
Technology, Tsinghua University,  
Beijing, China, 100084

---

### Online ISBN

978-3-642-01307-2

---

### Series Title

Lecture Notes in Computer  
Science

---

### Series Volume

5476

---

### Series ISSN

0302-9743

---

### Publisher

Springer Berlin Heidelberg

---

### Copyright Holder

Springer Berlin Heidelberg

---

### Additional Links

[About this Book](#)

Over 9 million scientific documents at your fingertips

Browse by Discipline 

### Our Content

Journals

Books

Book Series

Protocols

Reference Works

### Other Sites

Springer.com

SpringerProtocols

SpringerMaterials

### Help & Contacts

Contact Us

Feedback Community

Impressum

© Springer International Publishing AG, Part of Springer Science+Business Media | [Privacy Policy](#), [Disclaimer](#), [General Terms & Conditions](#)

Media

沪 ICP 备 13017623 号

Not logged in Unaffiliated 122.70.132.162