

An Expertise-enhanced Collaborative Filtering Method for Keywords Recommendation in Searching Engine Marketing

Yuan Zhu

School of Trade and Economics, Taizhou Vocational and Technical College, Zhejiang, China

*Corresponding author e-mail: tzzy246@163.com

Abstract. Nowadays, with the development of information technology, people like to use searching engine (SE) for help. Thus stores and businessmen start to realize the necessity of marketing through searching engine. Searching engine marketing (SEM) becomes a kind of new marketing tools, which can also be called advertisement campaign. SE business users select keywords in the SE system to attract customers searching of these. How to select the right keyword in SEM is a necessary problem needed to be solved recently for SE business users. Traditional keywords recommendation only recommend semantic similar keywords to business users which still makes the result set too large to choose. The new influential factor - expertise value combined with click rate will be integrated into the traditional collaborative filtering to minimize the keywords set waiting to be selected. The final recommendation is based on our expertise-enhanced collaborative filtering combined with semantic similarity.

1. Introduction

With the repaid growth of Web 2.0, customers are used to shopping online and searching the Internet for suggestions instead of going to the offline shopping store in person. Thus the online information has become very important for guiding customers to the right destinations they want. More guidance or references appears online to meet their increasing needs, which has caused information explosion. The exploded data have increased the difficulty for customers to find the right commodity, which made them more and more confused.

Searching Engine (SE) was developed to solve this problem by recommending information or commodities related to the customer searching words. By pressing the enter bottom online of Google Searching Engine, related web links will be provided immediately based on the keywords being typed on the searching bar. As the results of this searching engine can easily affect the shopping decisions of customers, Searching Engine Marketing (SEM) [1] was proposed for businessman to broadcast their products online by paying the SE company on the user clicks searched from the keywords they set. The click-paid method had soon attracted the attention of many retailers. However, as more and more retailers participate in the SEM plan, more keywords appear in the SE database. Which keywords are the most suitable for these retailers start to become a difficult problem. Current SEM system can recommend some keywords to retailers. Yet these keywords are only based on their similar commodities, which lack the consideration of other retailers. Thus we propose a keyword recommendation method, which considers the influence of both retailers and commodities. We classifies the retailer into different levels according to their expertise index extracted and calculated from their own profile to make our recommendation more proper for retailers to succeed in advertise campaign.



2. Related work

Recommendation system has proven to be an effective response to the information overload currently in numerous areas [2], providing suggestions for all kinds of users. Traditional methods of recommendation can be classified into two kinds:

- 1) Content-based recommendation.
- 2) Collaborative filtering (CF).

The former one relies on the completeness of user/item profile. The recommendation process cannot be carry on if the system lack user's content, which was defined as a cold start problem.

The latter one is more used in modern business. For example, Amazon has used collaborative filtering method to recommend books considering other users' choice as references. The predicted recommendation is based on ranking the ratings calculated from its nearest users'/items'. Nevertheless, in the case of rating scarcity, the CF can hardly execute because of rating prediction problem.

For the calculation of CF, usually we use the Pearson correlation coefficient or the cosine measure [3] to acquire the similarities between different users/items, shown as follows [3],

$$\text{Cos}(A, B) = \frac{\sum_{i=1}^n R_{A,i} \times R_{B,i}}{\sqrt{\sum_{i=1}^n (R_{A,i})^2} \times \sqrt{\sum_{i=1}^n (R_{B,i})^2}} \quad (1)$$

where n denotes the number of co-rated items by both user A and user B, $R_{A,i}$ and $R_{B,i}$ are the ratings for the i th item from the n co-rated items rated by user A and user B respectively.

Neither of these two methods can solve the above problem. Recently some methods [4] have been proposed which combine the two methods together to make the recommendation more suitable in different extreme cases. New elements such as trust [5], demographic factors [6], etc. calculated through different algorithms and models have been integrated into traditional recommendation method, which can influence the prediction precision. Parvin, Moradi and Esmaeili [5] have used ant colony optimization to calculate the user trust to enhance the traditional CF method.

Recently, expertise has been discussed by many others [7-9] in recommendation systems as a key point for rating prediction, which can measure the professionalism of the reference users. In the process of words recommendation, the suggestions of professional ones can be considered more valuable than other reference. Recommended results based on domain experts will be more trustable than those simply based on semantic similarity [7] or rating similarity. Thus we propose an expertise-based personalized keywords recommendation method in SEM to differ from others with the redefinition of experts adapting to the new area.

3. An enhanced expertise-based recommendation method for SEM

Fig.1 below shows the architecture of our expertise-enhanced collaborative filtering recommendation method for SEM, which can be separated into three parts:

- 1) Data extraction. Keywords and user contents that have attended the former advertisement campaign will be extracted from the underlying data forming SE keywords database and SE user log database.
- 2) Rating computation. In this section, we calculate the similarity between keywords and users through the data extracted from the database. Besides, the expertise level of the users is also computed to acquire the expertise-based words rating similarity for the latter recommendation. The process of computation will be elaborated in details later on.
- 3) Recommendation. The top-ranking keywords will be recommended by the calculated expertise-enhanced words rating.

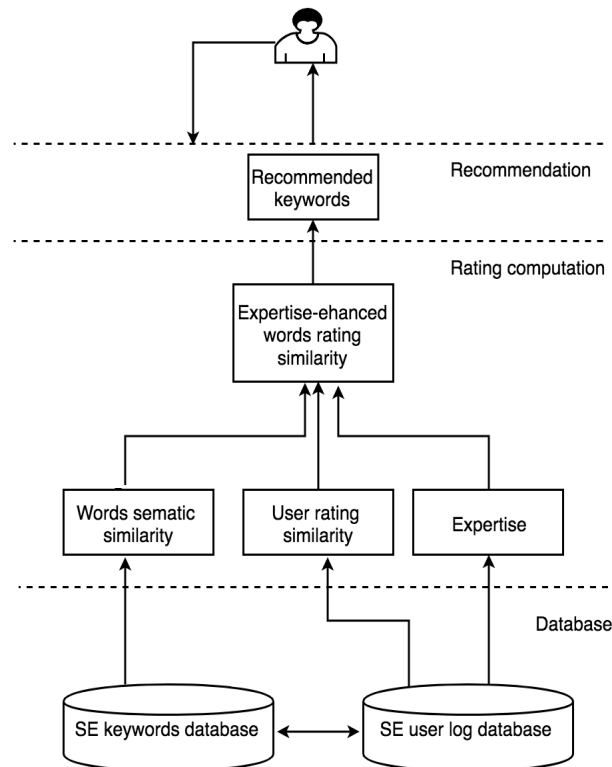


Figure 1. The architecture of our expertise-enhanced CF for SEM

3.1. Words semantic similarity

Ontology-based semantic similarity has long been researched by literature [11-12], which is proved to be accurate and quick responsible. Cai [13] has proposed a semantic similarity method in the area of manufacturing services with his manufacturing ontology. In our work, we would like to take Cai [13]'s method as a reference to compute keyword semantic similarity based on the Abdeljaber [14]'s proposed ontology.

The semantic similarity between keyword KW_i and keyword KW_j within Abdeljaber [14]'s proposed ontology O can be calculated as the equation below [2],

$$Sim(KW_i, KW_j, O) = \alpha \frac{n(C_{super}(KW_i, O) \cap C_{super}(KW_j, O))}{n(C_{super}(KW_i, O) \cup C_{super}(KW_j, O))} + \beta \frac{n(C_{sub}(KW_i, O) \cap C_{sub}(KW_j, O))}{n(C_{sub}(KW_i, O) \cup C_{sub}(KW_j, O))} \quad (2)$$

where $C_{super}(KW_i, O)$ and $C_{super}(KW_j, O)$ represent the super classes subsuming KW_i and KW_j within ontology O separately while $C_{sub}(KW_i, O)$ and $C_{sub}(KW_j, O)$ denote the sub classes subsumed by KW_i and KW_j respectively within O . $n(C_{super}(KW_i, O) \cap C_{super}(KW_j, O))$ is the number of super classes subsuming both KW_i and KW_j while $n(C_{super}(KW_i, O) \cup C_{super}(KW_j, O))$ is the number of super classes subsuming either KW_i or KW_j . Besides, $n(C_{sub}(KW_i, O) \cap C_{sub}(KW_j, O))$ is the number of sub classes subsumed both KW_i and KW_j while $n(C_{sub}(KW_i, O) \cup C_{sub}(KW_j, O))$ is the number of sub classes subsumed either KW_i or KW_j . α and β are the weighting factors and we define $\alpha + \beta = 1$ to restrict the value of $Sim(KW_i, KW_j, O)$ to be not too large.

The semantic similarity between user U_i and U_j can be computed as equation [3] likewise,

$$Sim(U_i, U_j, O) = \alpha \frac{n(C_{super}(U_i, O) \cap C_{super}(U_j, O))}{n(C_{super}(U_i, O) \cup C_{super}(U_j, O))} + \beta \frac{n(C_{sub}(U_i, O) \cap C_{sub}(U_j, O))}{n(C_{sub}(U_i, O) \cup C_{sub}(U_j, O))} \quad (3)$$

where $C_{super}(U_i, O)$ and $C_{super}(U_j, O)$ represent the super classes subsuming U_i and U_j within ontology O separately while $C_{sub}(U_i, O)$ and $C_{sub}(U_j, O)$ denote the sub classes subsumed by U_i and U_j respectively within O .

3.2. Expertise computation

In this paper, the expertise user of a searching engine database can be defined as ones whose former keywords being selected to participating in the advertisement campaign have reaped lots of customer user clicks. The more the customer clicks the link related to the keywords, the more professional the user is. Finally it will increase the expertise value of the user.

Suppose a user U_i has selected N keywords to participate in the advertisement campaign and one keyword $KW(U_i, n)$ selected by him has accepted $C_{KW(U_i, n)}$ clicks, thus the total clicks received of user U_i bought by his keywords can be calculated as $\sum_{n=1}^N C_{KW(U_i, n)}$.

Thus the expertise value of a user U_i can be computed through the following equation,

$$EV(U_i) = \frac{\sum_{n=1}^N C_{KW(U_i, n)}}{\sum_{i=1}^M \sum_{n=1}^N C_{KW(U_i, n)}} \quad (4)$$

where $\sum_{i=1}^M EV(U_i) = 1$ and $EV(U_i)$ is the expertise value of user U_i and $\sum_{i=1}^M \sum_{n=1}^N C_{KW(U_i, n)}$ is the total clicks within the search engine database of M total users taking part in the former advertisement campaign.

This way of computation for expertise value is to normalize the result, which make the comparisons between two different users easier.

3.3. Expertise-enhanced collaborative filtering

Thus the expertise-enhanced predicted click rate of a new keyword $KW(U_j)$ by user U_j can be computed through the equation below (6),

$$PC(KW(U_j)) = Sim_U \times Sim_{KW} \times C_{KW(U_i)} \times EV(U_i) \quad (5)$$

where

$$Sim_U = Sim(U_i, U_j, O) \quad (6)$$

$$Sim_{KW} = Sim(KW(U_i), KW(U_j), O) \quad (7)$$

$Sim_U = Sim(U_i, U_j, O)$ means the semantic similarity between user U_i and U_j while $Sim_{KW} = Sim(KW(U_i), KW(U_j), O)$ means the semantic similarity between keywords KW_i and KW_j . $C_{KW(U_i)}$ is the click rate of user U_i and $EV(U_i)$ is the expertise value of U_i .

The predicted click rate $PC(KW(U_j))$ of user U_j 's keyword $KW(U_j)$ will be ranked according to the final results and the keywords with the top highest predicted click rate will be recommended to the user for them to select as the advertisement campaign reference.

4. Conclusion

In this paper, an expertise-enhanced collaborative filtering method has been proposed to address the keywords selection problem in searching engine market (SEM) for advertisement campaign. We use ontology-based semantic method to compute the user similarity and keywords similarity to firstly filter the dissimilar keywords with less semantic similarity. Secondly, we give a new expert definition for

the users in SE system and calculate the expertise value of the users through keywords click rates. The final predicted click rate for a new coming keyword will be computed considering the influence of both user/keywords semantic similarity and user expertise. However, there are still some restrictions in the process of rating prediction. The method in our work does not consider the various domains of different keywords. Besides, experts in different fields may have different expertise value. In this case, more studies should be conducted for further improvement and perfection.

Acknowledgments

This work was supported by Normal Program of Provincial Education Department of Zhejiang (No.Y201636388).

References

- [1] Drèze, Xavier, and François-Xavier Hussherr. 2003. "Internet Advertising: Is Anybody Watching?" *Journal of Interactive Marketing*. Vol. 17, no. 4, pp. 8–23.
- [2] M. I. Matin-Vicente, A. Gil-Solla, M. Ramos-Cabrer, Y. Blanco-Fernandez, M. Lopez-Nores,"Semantic inference of user's reputation and expertise to improve collaborative recommendations", *Expert System with Applications*, vol. 39, pp. 8248-5258, 2012.
- [3] L. Mekouar, Y. Iraqi, R. Boutaba, "An analysis of peer similarity for recommendations in P2P systems", *Multimedia Tools and Applications*, vol. 60, no. 2, pp. 277-303, 2012.
- [4] Christou, I. T., Amolochitis, E., & Tan, Z.-H.. AMORE: design and implementation of a commercial-strength parallel hybrid movie recommendation engine. *Knowledge and Information Systems*, Vol. 47, No. 3, pp. 671–696 , 2015.
- [5] Parvin, H ; Moradi, P; Esmaeili, S . TCFACO: Trust-aware collaborative filtering method based on ant colony optimization, *EXPERT SYSTEMS WITH APPLICATIONS*. Vol.118, pp. 152-168, 2019.
- [6] B. Yapriady, A. L. Uitdenbogerd, "Combining Demographic Data with Collaborative Filtering for Automatic Music Recommendation", *Knowledge-Based Intelligent Information and Engineering Systems*, vol. 3684, pp. 201-207, 2005.
- [7] Y. H. Xu, X. T. Guo, J. X. Hao, J. Ma, R. Y. K. Lau, W. Xu, "Combining social network and semantic concept analysis for personalized academic researcher recommendation", *Decision Support Systems*, vol. 54, no. 1, pp. 564-573, 2012.
- [8] K. Kwon, J. Cho, Y. Park, "Multidimensional credibility model for neighbor selection in collaborative recommendation", *Expert Systems with Application*, vol. 36, no. 3, pp. 7114-7122, 2009.
- [9] Z. X. Huang, X. D. Lu, H. L. Duan, C. H. Zhao, "Collaboration-based medical knowledge recommendation", *Artificial Intelligence in Medicine*, vol. 55, no. 1, pp. 13-24, 2012.
- [10] Chen Y, Lin SP. Technology of semantic search based on ontology. *Comput Eng Appl*. 2006;S1:78-80. ^[1]_[SEP]
- [11] Bhogal J, MacFarlane A, Smith P. A review of ontology based query expansion. *Inf Process Manag*. Vol. 43, no. 4, pp. 866-886, 2007. ^[1]_[SEP]
- [12] Yong LI, Zhang ZG. Semantic retrieval research based on ontology. *Comput Eng Sci*. vol. 30, no. 4, pp. 17-18, 2008. ^[1]_[SEP]
- [13] M. Cai, W. Y. Zhang, K. Zhang, "ManuHub: a Semantic Web system for ontology-based service management in distributed manufacturing environments", *IEEE Transactions on Systems, Man, and Cybernetics Part A: Systems and Humans*, vol. 41, no. 3, pp. 574-582, 2011.
- [14] Hikmat A. M. Abdeljaber, "Profile-Based Semantic Method using Heuristics for Web Search Personalization" *International Journal of Advanced Computer Science and Applications(ijacsa)*, vol. 9, no. 9, pp. 191-198, 2018

Reproduced with permission of copyright owner. Further reproduction prohibited without permission.