



Mining Web Graphs for Recommendations

Shaik Moulali

M.Tech Scholar ,
Dept.of CSE, Quba

College Of Engineering & Technology
Nellore, JNTUA, India

Syed Abdul Haq

Associate Professor & HOD,
Dept.of CSE, Quba

College Of Engineering & Technology
Nellore, JNTUA, India

Abstract : *By increasing various contents in web, recommendation by user also be increased as well as techniques for those maintenance also be very necessary. Various recommendations on web are movies, news, music, books and images, etc. Number of recommendations is search by all around world. These data recommendation modeled in data resources using different types of graphs. But these methods are less structured and more diverse. So here we implemented a framework for recommendation mining in web graphs. This framework developed by using heat diffusion techniques in both directed and undirected graphs. It is applied for large scalable data also.*

Index Terms: *Recommendation, heat diffusion, query suggestion, sub query.*

1. Introduction:

Now a day, we found rapid growth of web information as well as search engines also. But a small and critical problem faced by the users is utilization of information effectively. Mainly in Web 2.0 based application faced this problem. The user related recommendations are free style manner and less structured. So that mining of those related recommendations is very critical. Then may be unrelated and unexpected data also displays in web for particular recommendation. This data not satisfied by user fully. So for giving satisfaction to users we need to improve web user's experience in many Web applications, Recommender Systems.

Mostly collaborative filtering technique used in many recommender systems. This technique predicts data automatically when the user starts to enter recommendation in web search. Here web displays related data to recommendation, among those users may selects which they want particularly. And one more main thing done by collaborative filtering is mostly used recommendation is displayed in top of layer. So this collaborative technique is very simple as well as very effective method for mining recommendations and used by many commercial institutions.

1.1. Collaborative Filtering:

Typical problem occurred in this type of collaborative filtering is, this filtering algorithm needs rating matrix for each user-item. The user-item specifies raring preferences of user-specific recommendation and user's characteristics. But rating matrix for each user-item is not available in search engines, because web recommendation is less structured. In web search engines any type of recommendation may be searched. Those recommendations are modeled by web graphs, which are maybe directed or undirected graphs. Designing of graphs for recommendation is compulsory in mining concept. But up to now we are facing many challenges in designing of web graphs. So in this paper we proposed a model for to face challenges of graphs.

1.2. Challenges:

- Rating matrix for each user-item.
- Web recommendation is less structured.

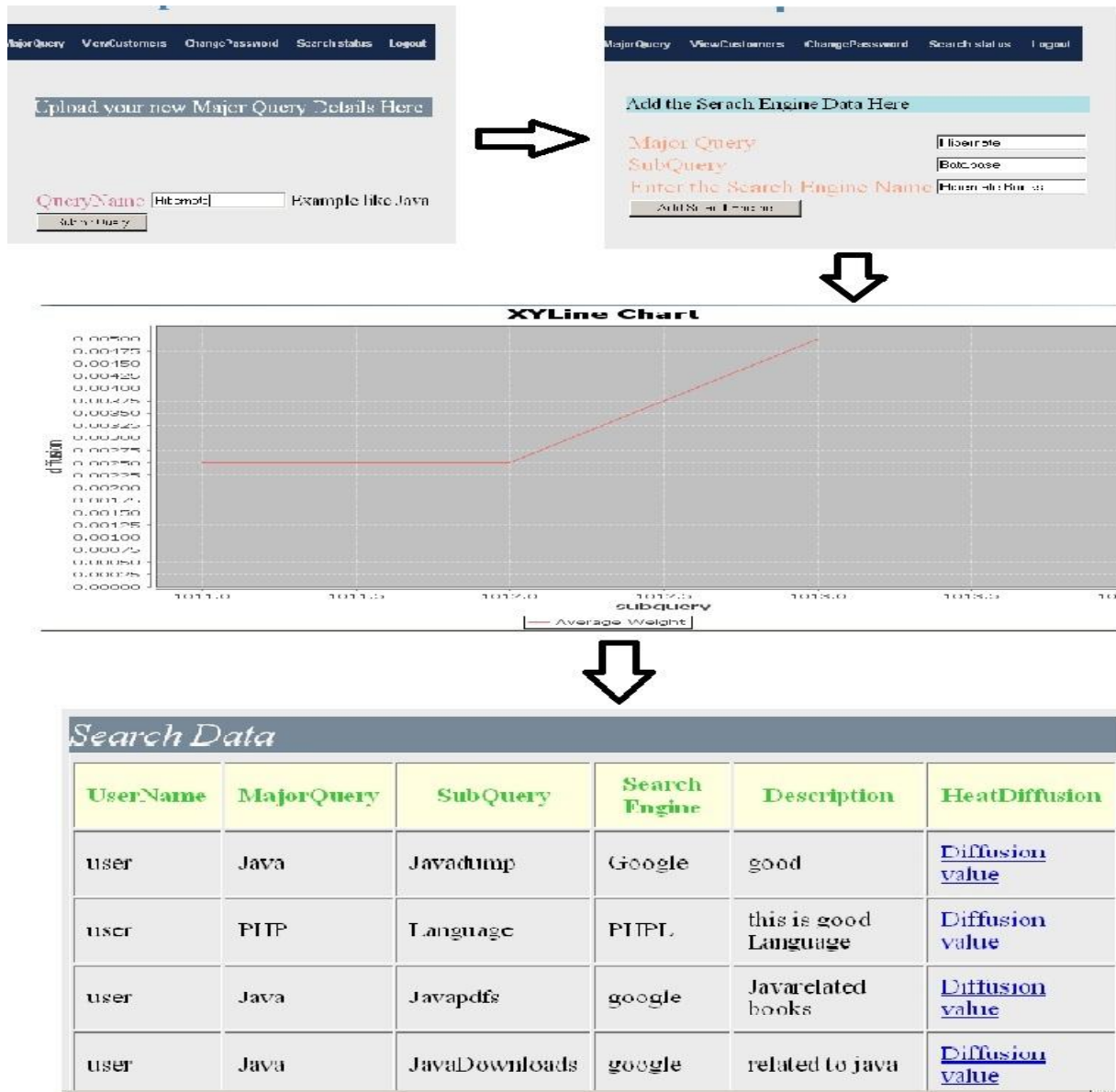
2. Framework for the Recommendations:

To solve above challenges, in this paper we proposed a novel solution and framework for the recommendations by using heat diffusion techniques for both directed graphs and undirected graphs. This is general framework and used all types of recommendations on web. This framework predicts exact data which relevent to recommendations. And it is designed for very large recommendations also in natural treatment manner.

2.1. Definitions:

- Heat diffusion is a physical method. Always in a medium, heat flows from high temperature position to low temperature position. Heat diffusion techniques are mostly applied in various domains such as classification and dimensionality reduction problems.
- In Diffusion on Undirected Graphs, we Consider
graph $G = (V, E)$
where V is the vertex set, and $V = (v_1, v_2, \dots, v_n)$.
 $E = (v_i, v_j)$ E is an edge between v_i to v_j and is considered as pipe which connected in between vertexes.
And $f_i(t)$ specifies heat at node v_i at time t . t is zero in beginning.

- Directed Graphs also used in heat diffusion model. These graphs mainly used in online recommender systems and social networking sites. Because in this type of recommended data contains trust levels for each user like from whom to whom relation exists.
- Random Jump defines how heat moved between nodes and which value it taken. Randomly jumps are occurred in between links. Here it does not a matter whether those are connected or unconnected or undirected.



2.2. Procedure:

Every web search engine has heat values internally for each and every URL. After entering suggestion by user, for that recommendation also we generate heat value automatically. Now our framework match the recommended heat value with URL's heat value, and generates diffusion value which specifies how many search engine exist particular recommended data. Now the mined recommendations only predict and displayed in web. In this proposed system we also enter sub query for main query and selects particular search engine manually. Finally we also see ran king of various search engines.

3. Implementation:

Above entire diagram shows a procedure how the mining of recommendation worked. The first sub diagram shows that how a main query is uploaded. That query is mentioned as main query. After entering main query we submit query to data source. If we need particular search engine to search our topic then we a chance to add particular search engine name, that will shown in second sub diagram. Here one more facility is we will also add sub query along with main query. Then we submit query for specified search engine. Now search engine searches the main and sub queries and display results, which is shown in last sub diagram. Here not only particular search engine results display, but also other search engines data which related to query displayed. And finally diffusion graph which was in third sub diagram shows the diffusion value which was determined based on the number of search engines.

4. Conclusion:

We proposed a solution for recommendation mining in web which is very large scale based on heat diffusion framework. Collaborative filtering was successfully modified by our technique and gives efficient solution in both directed and undirected graphs manner. As part of future enhancement we are planning to investigate comparing methods of rankings in between previous web search results and present web search results. Because in every search of same recommendation in different search engines in a single search engines may varies the data, by the rapid growth and updating in web.

References:

- [1] E. Agichtein, E. Brill, and S. Dumais, "Improving Web Search Ranking by Incorporating User Behavior Information," SIGIR '07: Proc. 29th Ann. Int'l ACM SIGIR Conf. Research and Development in Information Retrieval, pp. 19-26, 2006.
- [2] E. Auchard, "Flickr to Map the World's Latest Photo Hotspots," Proc. Reuters, 2007.
- [3] R. TiberiBaeza-Yates and A. Tiberi, "Extracting Semantic Relations from Query Logs," KDD '07: Proc. 13th ACM SIGKDD Int'l Conf. Knowledge Discovery and Data Mining, pp. 76-85, 2007.
- [4] R.A. Baeza-Yates, C.A. Hurtado, and M. Mendoza, "Query Recommendation Using Query Logs in Search Engines," Proc. Current Trends in Database Technology (EDBT) Workshops, pp. 588- 596, 2004.
- [5] D. Beeferman and A. Berger, "Agglomerative Clustering of a Search Engine Query Log," KDD '00: Proc. Sixth ACM SIGKDD Int'l Conf. Knowledge Discovery and Data Mining, pp. 407-416, 2000.
- [6] M. Belkin and P. Niyogi, "Laplacian Eigenmaps for Dimensionality Reduction and Data Representation," Neural Computation, vol. 15, no. 6, pp. 1373-1396, 2003.
- [7] J.S. Breese, D. Heckerman, and C. Kadie, "Empirical Analysis of Predictive Algorithms for Collaborative Filtering," Proc. 14th Conf. Uncertainty in Artificial Intelligence (UAI), 1998.
- [8] S. Brin and L. Page, "The Anatomy of a Large-Scale Hypertextual Web Search Engine," Computer Networks and ISDN Systems, vol. 30, nos. 1-7, pp. 107-117, 1998.
- [9] J. Canny, "Collaborative Filtering with Privacy via Factor Analysis," SIGIR '07: Proc. 25th Ann. Int'l ACM SIGIR Conf. Research and Development in Information Retrieval, pp. 238-245, 2002.
- [10] H. Cao, D. Jiang, J. Pei, Q. He, Z. Liao, E. Chen, and H. Li, "Context-Aware Query Suggestion by Mining Click-Through and Session Data," KDD '08: Proc. 14th ACM SIGKDD Int'l Conf. Knowledge Discovery and Data Mining, pp. 875-883, 2008.
- [11] P.A. Chirita, C.S. Firan, and W. Nejdl, "Personalized Query Expansion for the Web," SIGIR '07: Proc. 30th Ann. Int'l ACM SIGIR Conf. Research and Development in Information Retrieval, pp. 7- 14, 2007.
- [12] N. Craswell and M. Szummer, "Random Walks on the Click Graph," SIGIR '07: Proc. 30th Ann. Int'l ACM SIGIR Conf. Research and Development in Information Retrieval, pp. 239-246, 2007.
- [13] H. Cui, J.-R. Wen, J.-Y. Nie, and W.-Y. Ma, "Query Expansion by Mining User Logs," IEEE Trans. Knowledge Data Eng., vol. 15, no. 4, pp. 829-839, July/Aug. 2003.
- [14] A.S. Das, M. Datar, A. Garg, and S. Rajaram, "Google News Personalization: Scalable Online Collaborative Filtering," WWW '07: Proc. 16th Int'l Conf. World Wide Web, pp. 271-280, 2007.
- [15] M. Deshpande and G. Karypis, "Item-Based Top-n Recommendation," ACM Trans. Information Systems, vol. 22, no. 1, pp. 143-177, 2004.
- [16] G. Dupret and M. Mendoza, "Automatic Query Recommendation Using Click-Through Data," Proc. Int'l Federation for Information Processing, Professional Practice in Artificial Intelligence (IFIP PPAI), pp. 303-312, 2006.
- [17] N. Eiron, K.S. McCurley, and J.A. Tomlin, "Ranking the Web Frontier," WWW '04: Proc. 13th Int'l Conf. World Wide Web, pp. 309- 318, 2004.
- [18] W. Gao, C. Niu, J.-Y. Nie, M. Zhou, J. Hu, K.-F. Wong, and H.-W. Hon, "Cross-Lingual Query Suggestion Using Query Logs of Different Languages," SIGIR '07: Proc. 30th Ann. Int'l ACM SIGIR Conf. Research and Development in Information Retrieval, pp. 463-470, 2007.