Era of Personalized information retrieval system in focus of Ranking & Rating approaches and ontology

Prof. P.Pradeep Kumar [1] **AnjanKumar.M [2]** **Raju G [3]**

1. HOD, Dept. of CSE, Vivekananda Institute of Technology and Science, karimnagar.
2. Asst. Professor, Dept. of IT, Vivekananda Institute of Technology and Science, karimnagar
3. Asst. Professor, [M.TECH(CSE)], Vivekananda Institute of Technology and Science, karimnagar

ABSTRACT

To maximize user experience on your site and ensure they re-visit in a way that is simple but it is not possible with a 'non personalized website'. The major engines are search secrets closely, so one can never be absolutely certain how they are operating. But they are evolving, and personalization seems to be the wave of the future. Recommender systems are becoming increasingly important to individual users and businesses for providing personalized recommendations .We introduce and explore a number of item ranking techniques that can generate recommendations that have substantially higher aggregate diversity across all users while maintaining comparable levels of recommendation accuracy. In addition with recommendation systems personalized ontology model is proposed for knowledge representation and reasoning over user profiles. This model learns ontological user profiles from both a world knowledge base and user local instance repositories. The ontology model is evaluated by comparing it against benchmark models in web information gathering. The results show that this ontology model is successful.

Keywords: personalized information, recommendations, rating, ranking, ontology, knowledge.

I. INTRODUCTION

Is personalization required for every organization? Probably not. If your website does not have enough content to personalize then there is little point in trying to fragment it into profiles or tracked experiences - but if your site is large and you are struggling to ensure users get presented with appropriate content - then it would be one very powerful way to improve the user experience**[2].**

Recommender systems technologies have been introduced to help people deal with these vast amounts of information, and they have been widely used in research as well as e-commerce applications, such as the ones used by Amazon and Netflix. The most common formulation of the recommendation problem relies on the notion of ratings, i.e., recommender systems estimate ratings of items (or products) that are yet to be consumed by users, based on the ratings of items already consumed. Recommender systems typically try to predict the ratings of unknown items for each user, often using other users' ratings, and recommend top *N* items with the highest predicted ratings. Accordingly, there have been many studies on developing new algorithms that can improve the predictive accuracy of recommendations. However, the quality of recommendations can be evaluated along a number of dimensions, and relying on the accuracy of recommendations alone may not be enough to find the most relevant items for each user. In particular, the importance of *diverse* recommendations has been previously emphasized in several studies. **[3][8]**

These studies argue that one of the goals of recommender systems is to provide a user with highly idiosyncratic or personalized items, and more diverse recommendations result in more opportunities for users to get recommended such items. With this motivation, some studies proposed new recommendation methods***[5][8]*** that can increase the diversity of

ISSN Online: 2319 – 9253
Print: 2319 – 9245

recommendation sets for a given *individual* user, often measured by an average dissimilarity between all pairs of recommended items, while maintaining an acceptable level of accuracy. These studies measure recommendation diversity from an individual user's perspective (i.e., *individual diversity*).

As a model for knowledge description and formalization, ontology's are widely used to represent user profiles in personalized web information gathering. However, user profiles, many models have utilized only knowledge from either a global knowledge base or user local information. [8][10]

II. EXISTING MODELS

There is a growing awareness of the importance of aggregate diversity in recommender systems. Furthermore, while, as mentioned earlier, there has been significant amount of work done on improving individual diversity**[7],** the issue of aggregate diversity in recommender systems has been largely untouched. By this it is becoming increasingly harder to find relevant content. This problem is not only widespread but also alarming.by considering the models in ontology are as follows

A. BASELINE MODEL: CATEGORY MODEL

This model demonstrated the noninterviewing user profiles, a user's interests and preferences are described by a set of weighted subjects learned from the user's browsing history. These subjects are specified with the semantic relations of superclass and subclass in an ontology**[8][6].**

B. BASELINE MODEL: WEB MODEL

The web model was the implementation of typical semi interviewing user profiles. It acquired user profiles from the web by employing a web search engine.

III. PROPOSED APPROACHES IN RECOMMENDATION SYSTEM

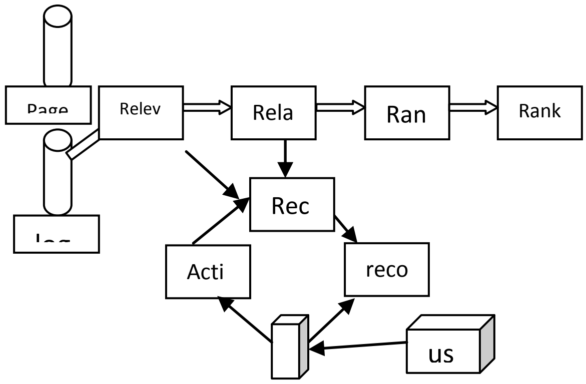

Fig 1. Recommendation systems architecture

In real world settings, recommender systems generally perform the following two tasks in order to provide recommendations to each user. First, the ratings of unrated items are estimated based on the available information (typically using known user ratings and possibly also information about item content or user demographics) using some recommendation algorithm. And second, the system finds items that maximize the user's utility based on the predicted ratings, and recommends them to the user **[3][5][8]**

In particular, these techniques are extremely *efficient*, because they are based on scalable sorting-based heuristics that make decisions based only on the "local" data (i.e., only on the candidate items of each individual user) without having to keep track of the "global" information, such as which items have been recommended across all users and how many times.

A. POSTING THE OPINION

We get the opinions from various people about business, e-commerce and products through online. The opinions may be of two types. Direct opinion and comparative opinion. Direct opinion is to post a comment about the components and attributes of products directly. Comparative opinion is to post a comment based on comparison of two or more products. The comments may be positive or negative. [3]

ISSN Online: 2319 – 9253
Print: 2319 – 9245

B. RECOMMENDATION TECHNIQUE

However, the quality of recommendations can be evaluated along a number of dimensions, and relying on the accuracy of recommendations alone may not be enough to find the most relevant items for each User, these studies argue that one of the goals of recommender systems is to provide a user with highly personalized items, and more diverse recommendations result in more opportunities for users to get recommended such items. With this motivation, some studies proposed new recommendation methods that can increase the diversity of recommendation sets for a given *individual* user.[5][1] They can give the feedback of such items.

C. RECOMMENDATION ALGORITHM

There exist multiple variations of neighborhood-based CF techniques. In this paper, to estimate $R^*(u, i)$, i.e., the rating that user u would give to item i, we first compute the similarity between user u and other users u' using a cosine similarity metric. Where $I(u, u')$ represents the set of all items rated by both user u and user u'. Based on the similarity calculation, set $N(u)$ of nearest neighbors of user u is obtained. The size of set $N(u)$ can range anywhere from 1 to $|U|$-1, i.e., all other users in the dataset[7][4].

Then, $R^*(u, i)$ is calculated as the adjusted weighted sum of all known ratings $R(u', i)$ Here $R(u)$ represents the average rating of user u. A neighborhood-based CF technique can be user-based or item-based, depending on whether the similarity is calculated between users or items, the user-based approach, but they can be straightforwardly rewritten for the item-based approach because of the symmetry between users and items in all neighborhood-based CF calculations. In our experiments we used both user-based and item-based approaches for rating estimation.[8][6].

D. RATING PREDICTION

First, the ratings of unrated items are estimated based on the available information (typically using known user ratings and possibly also information about item content) using some recommendation algorithm. Heuristic techniques typically calculate recommendations based directly on the previous user activities (e.g., transactional data or rating values). For each user, ranks all the predicted items according to the predicted rating value ranking the candidate (highly predicted) items based on their predicted rating value, from lowest to highest (as a result choosing less popular items.[3][10]

E. RANKING APPROACH

Ranking items according to the rating variance of neighbors of a particular user for a particular item. There exist a number of different ranking approaches that can improve recommendation diversity by recommending items other than the ones with topmost predicted rating values to a user [3]. A comprehensive set of experiments was performed using every rating prediction technique in conjunction with every recommendation ranking function on every dataset for different number of top-*N* recommendations.

IV. PROPOSED APPROACHES IN ONTOLOGY

The world knowledge and a user's local ins1tance repository (LIR) are used in the proposed model. World knowledge is commonsense knowledge acquired by people from experience and education. An LIR is a user's personal collection of information items. From a world knowledge base, we construct personalized ontologies by adopting user feedback on interesting knowledge. A multidimensional ontology mining method, Specificity and Exhaustivity, is also introduced in the proposed model for analyzing concepts specified in ontologies**[2]**. The users' LIRs are then used to discover background knowledge and to populate the personalized ontologies.

Compared with the TREC model, the Ontology model had better recall but relatively weaker precision performance. The Ontology model discovered user background knowledge from user local instance repositories, rather than documents read and judged by users. Thus, the Ontology user profiles were not as precise as the TREC user profiles.**[8], [5]**

The Ontology profiles had broad topic coverage. The substantial coverage of possibly-related topics was gained from the use of the WKB and the large number of training documents. Compared to the web data used by the web model, the LIRs used by the Ontology model were controlled and contained less uncertainties. Additionally, a large number of uncertainties were eliminated when user background knowledge was discovered. As a result, the user profiles acquired by the Ontology model performed better than the web model**[2],[3]**.

V. WORLD KNOWLEDGE BASE

The world knowledge base must cover an exhaustive range of topics, since users may come from different backgrounds. The structure of the world knowledge base used in this research is encoded from the LCSH references**[5]**.

The LCSH system contains three types of references:

Broader term- The BT references are for two subjects describing the same topic, but at different levels of abstraction (or specificity). In our model, they are encoded as the is-a relations in the world knowledge base.

Used-for- The UF references in the LCSH are used for many semantic situations, including broadening the semantic extent of a subject and describing compound subjects and subjects subdivided by other topics. When object A is used for an action, becomes a part of that action (e.g., "a fork is used for dining"); when A is used for another object, B, A becomes a part of B (e.g., "a wheel is used for a car"). These cases can be encoded as the part-of relations.

Related term- The RT references are for two subjects related in some manner other than by hierarchy. They are encoded as the related-to relations in our world knowledge base. **[6],[5]**

VI. CONCLUSION

We proposed a number of recommendation techniques that can provide significant improvements in recommendation diversity with only a small amount of accuracy loss. In addition, these ranking techniques offer flexibility to system designers, since they are parameterizable and can be used in conjunction with different rating prediction algorithms. They are also based on scalable sorting based heuristics and, thus, are extremely efficient.

The investigation will extend the applicability of the ontology model to the majority of the existing web documents and increase the contribution and significance of the present work.

The present work assumes that all user local instance repositories have content-based descriptors referring to the subjects, however large volume of documents existing on the web may not have such content-based descriptors. For this problem, strategies like ontology mapping and text classification/clustering were suggested. These strategies will be investigated in future work to solve this problem.

In our future work, we will investigate the methods that generate user local instance repositories to match the representation of a global knowledge base.

REFFERENCES

[1] Contemporary model and initiatives for information evalution. Systems: M.Anjankumar, R.kamalakar. Proceedings of the 63rd Internationalconference AMCM.2010

[2] Evolutionary approaches of personalized information retrieval systems ,"by M.ANJANKUMAR , R.KAMALAKAR A.HUSSAIN in Proceedings of the 3rdNationl Conference inf VITS(n6),knr Through ISTE

ISSN Online: 2319 – 9253
Print: 2319 – 9245

[3] Improving Aggregate Recommendation DiversityUsing Ranking-Based Techniques Gediminas Adomavicius, Member, IEEE, and YoungOk Kwon, IEEE TRANSACTIONS ON KNOWLEDGE AND DATA ENGINEERING,2011

[4] M¨akel¨a, E., Hyv¨onen, E., Saarela, S., “Ontogator - a semantic view based search engine service for web applications,” in International Semantic Web Conference, pp. 847-860, 2006.

[5] TRANSACTIONS ON KNOWLEDGE AND DATA ENGINEERING, VOL. 23, NO. 4, APRIL 2011A Personalized Ontology Model for Web Information Gathering Xiaohui Tao, Yuefeng Li, and Ning Zhong, Senior Member, IEEE.

[6] Bennett, and S. Lanning, “The Netflix Prize,” Proc. of KDD-Cup and Workshopat the 13th ACM SIGKDD Int’l Conf. on Knowledge and Data Mining, 2007.

[7] D. Billsus and M. Pazzani, “Learning Collaborative Information Filters,” Proc.Int’l Conf. Machine Learning, 1998.

[8] K. Bradley and B. Smyth, “Improving Recommendation Diversity,” Proc. of the12th Irish Conf. on Artificial Intelligence and Cognitive Science, 2001.

[9] S. Breese, D. Heckerman, and C. Kadie, “Empirical Analysis of PredictiveAlgorithms for Collaborative Filtering,” Proc. of the 14th Conf. on Uncertainty inArtificial Intelligence, 1998.

[10] Sacco, G.M., “Research results in dynamic taxonomy and faceted searchsystems,” in DEXA

AUTHORS PROFILES

Prof Pradeep Kumar

B.E M.Tech(Ph.D) having 18+ years of experience inAcademic & Industry. Currently he is theProfessor & Head of Department Of Computer Science &EngineeringAtVivekananda Institute of Technology & Science, Karimnagar, he has guidedmany UG & PG students. His research areas ofinterest are Software Engineering, Data mining &Data Warehousing, Information Security, Web Technologies

M.ANJAN KUAMR.

Has received MCA In kakatiya university and pursuing M.Tech in JNTUH. Working in VITS(N6) as assistant prof. & web Master Department of IT/MCA Karimnagar.He has working under the domains of DataEngineering, DataMining & warehousing, Information retrieval systems and AI With machine learning tools.And also working with Content management tools like joomla , drupal and word press.He has 5 international publications and 4 national publications with good impact ratios

GANTA RAJU

pursuing M.Tech CS fromVivekanandaInstitute of Technology &Science, Karimnagar . and awarded the degree of Masterof Computer Applications from JNTU Hyderabad. His Research areas include Programming In JAVA,Security, Cryptography & Web Technologies currentlyfocusing on Information estimation on Wireless Networks.