International Journal of Technical Research and Applications e-ISSN: 2320-8163,
www.ijtra.com Volume 6, Issue 1 (JAN-FEB 2018), PP. 27-31

COMPARATIVE STUDY OF EXISTING MODELS OF PERSONALIZED RECOMMENDATIONS

Vandana Srivastava[1], Dr S Q Abbas[2], Dr Mohammad Husain[3],
[1]Research Scholar, SVU, Gajraula,
[2]AIMT, Lucknow,
[3]JIT, Barabanki,
Uttar Pradesh, India

***Abstract*—Recommender systems are widely used these days in e-commerce, for the purpose of personalized recommendation. Based on each user's profile, previous purchase history, and online behavior, they suggest products which they are likely to prefer. For example, Amazon.com is using recommender systems for books. When a user logs-in to the system, it suggests books similar to previously bought ones by the user.**

In this paper we compare some previous work done on personalized recommendations system for web applications, and try to find out what lacked in these previous work.

***Index Terms*— e-commerce, CF-based recommender systems, Personalised, SOA.**

I. INTRODUCTION

World Wide Web has become one of the most extensive information resources in a recent span of time. It mostly covers all the information needed for any user. But, finding data on a large web site is a not an easy task. The users of the web sites mostly suffer from the problem of finding the required data in time. In fact, locating the required dataset on the web has become one of the difficult and time consuming tasks today.

II. BACKGROUND

Wherever Times is specified, Times Roman or Times New Roman may be used. If neither is available on your word processor, please use the font closest in appearance to Times. Avoid using bit-mapped fonts. True Type 1 or Open Type fonts are required. Please embed all fonts, in particular symbol fonts, as well, for math, etc.

III. EASE OF USE

Joonseok Lee, Kisung Lee, Jennifer G. Kim, Georgia Institute of Technology Atlanta, GA 30318, **"Personalized Academic Research Paper Recommendation System", 2010.** In this paper, they have presented a Personalized Academic Research Paper Recommendation System, which recommends related articles for each researcher. Their system makes three contributions. First, they have developed a web crawler to retrieve a huge number of research papers from the web. Second, they define a similarity measure for research papers. Third, they have developed their recommender system using collaboration filtering methods.

These days, many academic papers are coming out from a lot of conferences and journals. Academic researchers should go through all the conferences and journals which are related to their field of research and find out if there is any new articles that may relate to their current works. Sometimes they search the articles from Google scholars with the key words that might show interesting articles to them. However, these two methods require users to commit their time to search articles, which is labor-intensive, and also do not guarantee that they will find the exact articles related to their field of research. In order to reduce their workload, they suggest developing the scholarly paper recommendation system for academic researchers, which will automatically detect their research topics they are interested in and recommend the related articles they may be interested in based on similarity of the works. They believe this system will save the researchers' time to search the articles and increase the accuracy of finding the articles they are interested in. Even though their system showed good performance on recommending relevant topic's paper, they identified two limitations on their content-based recommendation. First, they cannot distinguish the meaning of topics that are narrowed by few specific words. In other words, their system recommends papers based on the words' frequency, so their system will recommend papers that contain many words that the user may be interested in. This cannot discover few words that restrict meaning of its topic, so it causes recommending the paper that has not relevant topic to the user. Also, there are users who have worked previous study on various topics, but are not interested in anymore. Also in this case, because their system does not have any additional information about whether the user is still interested in the paper or not, it is hard to distinguish the papers that recommend to users. They may be able to extend by applying publication year in some way.

To overcome these limitations they need to recommend papers based on not only relevant topics but also other user information. They suggest obtaining a user input about whether they like the recommended papers or not would be helpful information to differentiate the papers that users would be interested in more accurately. Also, through the focus group interview they discovered the interesting fact that even though the topics are not as much as relevant to their research topic, they showed great interest to the papers that their peer researchers, i.e., their former students or the researchers they

have done research together before, wrote. In this way, it will be important to include the information about relevant researchers to users and recommend papers that they found interesting or they have wrote. Also, the subjects replied, if they provide information about which researcher liked this papers, it would also give them great reason and motivation to read that paper.

For the perspective of machine learning, they may need to consider about scalability. Although their current system runs within a few minutes, it may take more time when they craw more data. First, they can improve accuracy of similarity measure by allowing counting the frequency of each word in a document, instead of bit vector model. TF-IDF model can be a great candidate to implement. In this model, they give more weight for frequently used words in a specific document, but not in other ones. Also, they may need to speed up the calculation. For this, dimension reduction will be helpful. Specifically, it would be better to add more stemming logic because this can deal with more words as same ones, so they can successfully reduce dimension. They may use L-Distance algorithm for calculating similarity of each word pair, and decide whether they are same or not.

Yoon Ho Choa,1, Jae Kyeong Kimb, Dongyang Technical College, 62-160 Kochuk-dong, Kuro-gu, Seoul 152-714, South Korea **"Application of Web usage mining and product taxonomy to collaborative recommendations in e-commerce". 2004**. The rapid growth of e-commerce has caused product overload where customers on the Web are no longer able to effectively choose the products they are exposed to. To overcome the product overload of online shoppers, a variety of recommendation methods have been developed. Collaborative filtering (CF) is the most successful recommendation method, but its widespread use has exposed some well-known limitations, such as sparsity and scalability, which can lead to poor recommendations.

However, as the number of customers and that of products managed in an e-commerce site grow rapidly, its application to ecommerce has exposed two major issues that must be addressed. The first issue is related to sparsity. In a large ecommerce site such as Amazon.com, there are millions of products and so customers may rate only a very small portion of those products. Most similarity measures used in CF work properly only when there exists an acceptable level of ratings across customers in common. Such sparsity in ratings makes the formation of neighborhood inaccurate, thereby resulting in poor recommendation. Many approaches have been proposed to overcome the sparsity problem. These approaches can be classified into three categories: implicit ratings, hybrid filtering and product-to product correlation. The implicit ratings approaches attempt to increase the number of ratings through observing customers' behavior. The hybrid filtering approaches combine content-based filtering and CF for augmenting sparse preference ratings. These approaches learn to predict which products a given customer will like by matching properties associated with each product to those associated with products that he/she has liked in the past, and then use such a content-based prediction to convert a sparse customer profile into a dense one. Instead of identifying the neighborhood of similar customers, the product-to-product correlation approach analyzes the customer profile to identify relationships between different products and then uses these relations to compute the prediction score for a given customer–product pair.

The second issue is related to scalability. Recommender systems for large e-commerce sites have to deal with millions of customers and products. Because these systems usually handle very high-dimensional profiles to form the neighborhood, the nearest neighbor algorithm is often very time-consuming and scales poorly in practice. To address the scalability problems in CF-based recommender systems, a variety of approaches have been developed. These approaches can be classified into two main categories: dimensionality reduction techniques and model-based approaches. Latent Semantic Index (LSI) is a widely used dimensionality reduction technique. It uses singular value decomposition (SVD) to factor the original rating space into three matrices and performs the dimensionality reduction by reducing the singular matrix. In model-based approaches, a model is first built based on the rating matrix and then the model is used in making recommendations. Usually, the model is expensive to build, but rapid to execute. Several data mining techniques such as Bayesian network, clustering and association rule mining have been applied to building the model.

They propose a recommendation methodology, called Web usage mining driven CF recommendation methodology using Product Taxonomy (WebCF-PT), to address the sparsity and scalability problems of current CF based recommender systems. Web usage mining is employed as an implicit ratings approach to address the sparsity problem. Web usage mining analyzes customers' shopping behaviors on the Web and collects their implicit ratings. This increases the number of ratings rather than only collecting explicit ratings, thereby reducing the sparsity. E-commerce data are rich and detailed compared to off-line commerce data. One type of collected e-commerce data is a clickstream that tracks visitors' path through a Web site. The clickstream in Web retailers provides information essential to understand the shopping patterns or prepurchase behaviors of customers such as what products they see, what products they add to the shopping cart, and what products they buy. By analyzing such information via Web usage mining, it is possible not only to make a more accurate analysis of the customer's interest or preference across all products (than analyzing the purchase records only), but also to increase the number of ratings (when compared to collecting explicit ratings only).

Nevertheless, the existing research in recommender systems has not offered a formal way for capturing implicit ratings of individual customer through Web usage mining. In this paper, they suggest a formal scheme to capture implicit ratings by analyzing customers' online shopping behaviors and to build the customer profiles. To solve the scalability problem, they use a nearest neighbor CF algorithm. But, before applying the algorithm, they reduce the dimensionality of the customer

profiles. As a dimensionality reduction technique, they employ a product taxonomy that represents hierarchical relationships between products as domain specific knowledge provided by marketing managers or domain experts. Similar products are identified and they are grouped together using the product taxonomy so as to build the customer profiles and to search for the neighbors in the reduced dimensional space.

This paper proposes a recommendation methodology based on Web usage mining, and product taxonomy to enhance the recommendation quality and the system performance of current CF-based recommender systems. Web usage mining populates the rating database by tracking customers' shopping behaviors on the Web, thereby leading to better quality recommendations. The product taxonomy is used to improve the performance of searching for nearest neighbors through dimensionality reduction of the rating database. Several experiments on real e-commerce data show that the proposed methodology provides higher quality recommendations and better performance than other CF methodologies. The research work presented in this paper makes several contributions to the recommender systems related research. First, they applied the product taxonomy both to reducing the sparsity in the rating database and to improving the scalability of searching for neighbors. Second, they developed a Web usage mining technique to capture implicit ratings by tracking customers' shopping behaviors on the Web and applied it to reducing the sparsity. Third, they developed a Web usage mining technique to choose proper products to recommend from the neighborhood. While their experimental results suggest that the proposed methodology is effective and efficient for product recommendations in the Internet business environment, these results are based on data sets limited to the particular e-commerce site that has a small number of customers, products, and transactions. Therefore, it is required to evaluate their methodology in more detail using data sets from a variety of large e-commerce sites. Furthermore, it will be an interesting research area to conduct a real marketing campaign to target customers using our methodology and then to evaluate its performance.

Choon-oh Lee, Minkyu Lee, Dongsoo Han, School of Engineering Information and Communications University (ICU) Daejeon, Korea, **"A Framework for Personalized Healthcare Service Recommendation"**.

The Internet, which brought the most innovative improvement on information society, has also brought many remarkable changes of healthcare services. Via the Internet, accessing information about healthcare services became relatively easier for service consumers who need adequate medical treatments. Moreover, consumers can communicate with doctors to get medical advices or to make appointment by e-mail or instant messengers, which are more convenient communication channels than by phone. Because of these benefits, much more healthcare service providers started publishing web sites for their service on the Internet competitively; as a consequence, consumers can obtain wide choice of services and better service quality.

However, there are also negative effects caused by exponential growth of the healthcare web sites. Because of too much information available, consumers cannot easily choose proper healthcare service among them. Some of them might not be able to judge what healthcare services are helpful because evaluating those services usually requires medical expertise. Moreover, there might be over-advertising web sites that show off exaggerated information about services. In this case, healthcare services on the Internet may confuse service consumers and make them more questionable. To help users to choose a proper service among the available services on the Internet, many brokering web sites for healthcare services such as healthcare web portals and search engines have been developed. The users can use the brokering web sites as starting points and find appropriate healthcare services using them. This improvement allows the users to access information about the services much easier than before, and the healthcare providers to save more lives. The brokering web sites, however, showed their limitation that more sophisticated mechanism is required in the domain of healthcare. The most of the users who does not have any knowledge about healthcare or any idea what is wrong with their bodies cannot find out proper healthcare services. What they need is not organized information about services, but a professional guideline to the most appropriate services for a specific user. Therefore, recommendation systems for better healthcare are proposed. Recommendation system for the healthcare is a web site that recommends healthcare services or provides useful information to the users considering. Healthcare Provider Recommendation System is an example of well proposed healthcare recommendation system. User can search the healthcare providers using location, providers' specialty, and reputation. However, what this system could not solve yet is that novice users and patients still may not be able to find out proper treatment for them when they do not know their exact health status. Because the most of people lack of medical knowledge, the system may not be effective in real life. To recommend appropriate healthcare services to novice or nonprofessional users more effectively, recommendation system must be aware of not only users' essential contexts such as location, but users' health status. In this paper, they suggest the Healthcare Service Recommendation Framework (HSRF) that can recommend healthcare services to each service consumer considering their health status. A main functionality of HSRF is that the framework automatically selects suitable healthcare services for a specific user among enormous services in the service repository. As a result, healthcare service recommendation of the framework can be personalized and very helpful for novice users. Moreover, to enhance extensibility of the framework, HSRF supports a convenient registration for new healthcare services or recommendation logics. Furthermore, they implemented HSRF; they evaluated the framework's functionality and feasibility successfully.

In this paper, they suggested a personalized healthcare service recommendation framework that considers consumers' health status to find adequate services for them. Their

framework gathers information about service consumer's health status and calculates medical similarities between consumer and healthcare services automatically. Based on these similarities of each consumer, the framework arranges and recommends proper healthcare services. Also, they implemented HSRF and evaluated its functionality and feasibility. Although the evaluation was not fully certain to prove all approaches of this paper, they concluded that their framework is quite enough to provide better healthcare service recommendation to novice users and patients.

IV. LACKS IN PREVIOUS RESEARCH

A. Personalized Academic Research Paper Recommendation System, 2010

In this paper, they cannot distinguish the meaning of topics that are narrowed by few specific words. In other words, their system recommends papers based on the words' frequency, so their system will recommend papers that contain many words that the user may be interested in. This cannot discover few words that restrict meaning of its topic, so it causes recommending the paper that has not relevant topic to the user. Also, there are users who have worked previous study on various topics, but are not interested in anymore. Also in this case, because their system does not have any additional information about whether the user is still interested in the paper or not, it is hard to distinguish the papers that recommend to users.

B. Application of Web usage mining and product taxonomy to collaborative recommendations in e-commerce, 2004

In this research recommender systems has not offered a formal way for capturing implicit ratings of individual customer through Web usage mining. They suggest a formal scheme to capture implicit ratings by analyzing customers' online shopping behaviors and to build the customer profiles. While their experimental results suggest that the proposed methodology is effective and efficient for product recommendations in the Internet business environment, these results are based on data sets limited to the particular e-commerce site that has a small number of customers, products, and transactions. Therefore, it is required to evaluate their methodology in more detail using data sets from a variety of large e-commerce sites.

C. A Framework for Personalized Healthcare Service Recommendation

Because of too much information available, consumers cannot easily choose proper healthcare service among them. Some of them might not be able to judge what healthcare services are helpful because evaluating those services usually requires medical expertise. Moreover, there might be over-advertising web sites that show off exaggerated information about services. In this case, healthcare services on the Internet may confuse service consumers and make them more questionable.

V. PROPOSED SUGGESTIONS

After studying previous models we found that a system is required for web applications in e-commerce area for user which not only search the requirement but also suggest best out of suggested list. For this we proposed a System Service Recommendation Framework (SSRF), it is a computerized system that recommends suitable services to service consumers based on their various interest. In other words, the framework acts as a mediator for business or non-business interactions between system service providers and consumers. A service provider is a Network node that provides a service interface for a software asset that manages a specific set of tasks. A service provider node can represent the services of a business entity or it can simply represent the service interface for a reusable subsystem. Therefore, it is an essential functionality for career web software such as e-career portals or search engines for system services.

For more personalized recommendation of the services, SSRF applies user's interest to its recommendation process. Interest is the information about user's current states or conditions, and it is the most important key to determine what specific services are suitable for the user. However, to use interest without any technical obstacle, the interest must be measurable and standardized. As a mediator, SSRF manages complex interactions between system service providers, consumers, and system administrators. The system service providers such as Professor, Director, Engineers, Technical advisor etc. can describe and register their own services on SSRF. Even if service not registered in SSRF, but SSRF provided service on-demand to incorporate new sources of data is required. Then, multiple recommendation mechanisms that are developed by the system administrators eventually search and recommend those registered or unregistered services for the users. Users can retrieve information about recommended services and evaluate them. The web portal system provides various services and information about system and also acts as an interface of SSRF for service consumers. The web portal actually triggers a recommendation process automatically delivering users' interest to SSRF; SSRF performs recommendation process with given users' contexts, analysis of data in runtime done by runtime analysis tools and sends recommended results back to the portal.

For this we suggested a flexible architecture of SSRF considering extensibility and scalability of the framework. Because, a brand-new type of service and interest can emerge at any time after the system is published, SSRF should require less effort to adapt those changes. Also, SSRF must be able to handle large amount of services and consumers. A consumer should be able to receive recommended results with high quality and low delay even if there are many services or requests from other consumers. To meet requirements above, we adopted SOA (Service Oriented Architecture) design paradigm to SSRF. System services can be implemented using the Web Services technology and registered easily at runtime. Also, core logics for the recommendation can be realized to

web services. For instance, we can imagine that there are number of web services available and a recommendation web service that gathers and arranges the services is deployed on the system. Likewise, there are recommendation web services that are in charge of their own categories and all the results from them are reorganized for users.

VI. CONCLUSION

After studying above papers we found that a system is required for web applications in e-commerce area for user to search their need in less time and get accurate result. Not only get results of their related search but also analyzed result for seared options, which option is best for their requirement, so that user select the best option without getting confused. For this we developed an effective framework for generating recommendations.

REFERENCES

[1] G. Adomavicius and A. Tuzhilin. Toward the next generation of recommender systems: A survey of the state-of-the-art and possible extensions. IEEE Transactions on Knowledge and Data Enginerring, 717(6):734–749, 2005.

[2] J. S. Breese, D. Heckerman, and C. Kadie. Empirical analysis of predictive algorithms for collaborative filtering. pages 43–52. Morgan Kaufmann, 1998.

[3] L. D. D. Algorithms for nonnegative matrix factorization. Advances in Neural Information Processing Systems, 13:556–562, 2001.

[4] D. Goldberg, D. Nichols, B. M. Oki, and D. Terry. Using collaborative filtering to weave an information tapestry. Commun. ACM, 35:61–70, December 1992.

[5] Q. He, J. Pei, D. Kifer, P. Mitra, and L. Giles. Context-aware citation recommendation. In Proceedings of the 19th international conference on World wide web, WWW '10, pages 421–430, New York, NY, USA, 2010. ACM.

[6] W. Hill, L. Stead, M. Rosenstein, and G. Furnas. Recommending and evaluating choices in a virtual community of use. In Proceedings of the SIGCHI conference on Human factors in computing systems, CHI '95, pages 194–201, New York, NY, USA, 1995. ACM Press/Addison-Wesley Publishing Co.

[7] D. Jannach, M. Zanker, A. Felfernig, and G. Friedrich. Recommender Systems – An Introduction. Cambridge, 2011.

[8] Y. Koren. Factorization meets the neighborhood: a multifaceted collaborative filtering model. In Proceeding of the 14th ACM SIGKDD international conference on Knowledge discovery and data mining, KDD '08, pages 426–434, New York, NY, USA, 2008. ACM.

[9] V. I. Levenshtein. Binary codes capable of correcting deletions, insertions, and reversals. Soviet Physics Doklady, 10(8):707–710, 1966.

[10] C. D. Manning, P. Raghavan, and H. Schutze.

[11] N. F. Matsatsinis, K. Lakiotaki, and P. Delias. A system based on multiple criteria analysis for scientific paper recommendation. 2007.

[12] S. M. McNee, I. Albert, D. Cosley, P. Gopalkrishnan, S. K. Lam, A. M. Rashid, J. A. Konstan, and J. Riedl. On the recommending of citations for research papers. In Proceedings of the 2002 ACM conference on Computer supported cooperative work, CSCW '02, pages 116–125, New York, NY, USA, 2002. ACM.

[13] T. M. Mitchell. Machine Learning. McGraw-Hill, 1997. Paterek. Improving regularized singular value decomposition for collaborative filtering. Statistics, 2007:2–5.

[14] M. Pazzani and D. Billsus. Content-based recommendation systems. In P. Brusilovsky, A. Kobsa, and W. Nejdl, editors, The Adaptive Web, volume 4321 of Lecture Notes in Computer Science, pages 325–341. Springer Berlin, Heidelberg, 2007.

[15] J. D. M. Rennie and N. Srebro. Fast maximum margin matrix factorization for collaborative prediction. In Proceedings of the 22nd international conference on Machine learning, ICML '05, pages 713–719, New York, NY, USA, 2005. ACM.

[16] P. Resnick, N. Iacovou, M. Suchak, P. Bergstrom, and J. Riedl. Grouplens: an open architecture for collaborative filtering of netnews. In Proceedings of the 1994 ACM conference on Computer supported cooperative work, CSCW '94, pages 175–186, New York, NY, USA, 1994. ACM.

[17] R. Salakhutdinov and A. Mnih. Probabilistic matrix factorization. 2008.

[18] B. Sarwar, G. Karypis, J. Konstan, and J. Riedl. Item-based collaborative filtering recommendation algorithms. pages 285–295, 2001.

[19] T. Strohman, W. B. Croft, and D. Jensen. Recommending citations for academic papers. In Proceedings of the 30th annual international ACM SIGIR conference on Research and development in information retrieval, SIGIR '07, pages 705–706, New York, NY, USA, 2007. ACM.

[20] X. Su and T. M. Khoshgoftaar. A survey of collaborative filtering techniques. Adv. in Artif. Intell., 2009:4:2–4:2, January 2009.

[21] J. Tang and J. Zhang. A discriminative approach to topic-based citation recommendation. In T. Theeramunkong, B. Kijsirikul, N. Cercone, and T.-B. Ho, editors, Advances in Knowledge Discovery and Data Mining, volume 5476 of Lecture Notes in Computer Science, pages 572–579. Springer Berlin / Heidelberg, 2009.

[22] Chishe Wang, Qi Shen, LinjunZou(2010), "Research and Application of Web Recommendation System Based on Cluster Mode", International Conference on E-Business and E-Government (ICEE), Pp. 1445 – 1447.

[23] Anitha, A., Krishnan, N.(2010), "A web usage mining based recommendation model for learning management systems", IEEE International Conference on Computational Intelligence and Computing Research (ICCIC), Pp. 1 – 4.

[24] Zhang Junyan, Shao Peiji(2010), "Personalized recommendation based on WAR", International Conference on Computer Application and System Modeling (ICCASM), Vol. 2, Pp. 108 - 111.

[25] MinutoEspil, Mauricio, Ale, Juan M.(2010), "An induction-based compaction of sets of association rules among Web concepts", Fourth International Conference on Research Challenges in Information Science (RCIS), Pp. 311 – 318.