

# Benchmarking Personalize Recommendation Algorithms for Web Based Applications

Neha Goel

Assistant professor, VIPS  
Delhi, India

Dr. C.K. Jha

Associate Professor, Banasthali Vidyapith  
Rajasthan, India

**Abstract**— World Wide Web (WWW) is generating huge amount of data every day. This massive data makes it difficult for the users to access the relevant information. Thus the concept of Personalization was introduced. The main aim of Personalization is to understand the user's behaviour and suggest him pages which he can likely visit next time depending upon his current searches. This paper brings out the comparison of Personalization Recommendation techniques being used in various Web Based Applications with their pros and cons. We present a thorough evaluation/comparison of Personalization techniques used in various web-based applications. The comparison is carried out based on their domain knowledge required, degree of personalization offered, access time consideration and user profiling. Through the comparison reported here, we conclude that the amalgamation of two or more techniques will yield better results and offer higher degree of personalisation.

**Keywords**— Personalization, Recommendation, Personalization services, Collaborative Filtering, WWW

## I. INTRODUCTION

People are relying more and more on internet for finding out any piece of information. But not everyone is able to reach to the desired location from where he can access what he wants. Thus the concept of Personalization came into practice. Personalization services aim at studying the user's behaviour and suggest him some web pages which he can likely visit next. The concept of Personalization was majorly used for E-commerce applications which helped a lot in increasing business. But now it has extended to almost all areas. Web Personalization can be considered as a solution to information overload problem. Personalization offers several advantages which include reducing access time, increased satisfaction of users due to effective search responses and improving the usability of a website [1].

The main aim of this paper is to briefly analyse the personalization services being offered in various areas of WWW and bring out the advantages and disadvantages of each. Section II gives us a brief about Web Mining and its

relationship with Personalization. Section III discusses the major areas offering Personalization Services. It elaborates on which technology is being used to design Personalization Services in a respective domain along with the advantages and disadvantages of each. Section IV brings out the comparative analysis. Conclusion of the overall work has been given finally in section V.

## II. WEB MINING AND PERSONALIZATION SERVICES

Web mining refers to the application of data mining techniques on web data. It is divided majorly into Web Content Mining (discovering useful information from web content), Structure (studying the hyperlink structure) and Usage Mining (studying the navigational behaviour of users from web access logs). The demand for web mining is increasing everyday due to the increase in number of users' everyday on the web. Reaching the desired location on the web is a major problem that almost every user faces. In order to cope up with this problem the Concept of Web Personalization came into practice. Web Personalization refers to studying the users' behaviour and delivering him custom content which could be helpful to him. It extracts knowledge from users behaviour recorded in web logs (Web Usage Mining) and can be easily combined with content (Web Content Mining), structure (Web Structure Mining) as well as user interests and preferences [1]. These days almost in every area Personalization is gaining importance. It is indeed helpful for the users as it saves a lot of time. The technologies used in Personalization include Collaborative Filtering (CF), Markov Models, Clustering, Association Rule Mining etc. These technologies when collaborated with web data help in generating interesting recommendations to the user.

## III. APPLICATION AREAS AND PERSONALIZATION SERVICES

A study of Personalization being offered in various areas of WWW was done and some of the major contributions are given below:

### A. Personalization Services for E-Commerce

Personalization recommendation has been considered to be the key technology for E-commerce and affects its performance significantly [2]. It has been seen that recommendation can convert potential requirement to actual requirement. Interests and preference of users' must surely be taken into account while designing personalization services for E-commerce sites. A lot of work has been done on designing Personalization Services in this area. The major technologies that E-commerce uses while designing the Personalization Services is Collaborative Filtering, Association Rule Mining, Content Mining, clustering, Usage Mining, Classification etc. A combination of any two techniques or more than two can also be utilized while designing the Personalization service. It was observed that CF if applied alone suffered from problems like cold start, sparsity, new users, scalability problems which lead to poor recommendations [2][3]. The advantages that CF offers include self adaptability and no requirement of domain knowledge [2]. Even in case of Association rules and Content Based, domain knowledge is not required but at the same time it is really difficult to generate the association rules. The personalization being offered is also not much. Just like CF even the content based recommendation suffers from sparsity problem. Rumeil in [4] by utilizing Collaborative Clustering and focusing on Page weight in user's clusters and Users average evaluation on pages designed a recommendation system. It was observed that this recommendation system increased real time response speed of recommendation algorithm and the degree of personalization offered was also quite good. Another work was published by Yoon et. al. [3] wherein two technologies viz. Web Usage Mining and Decision Tree Induction were used. Here customer preference and product association are learned from click stream and prospective customers are selected through decision tree induction. The authors have stated that this technology increased effectiveness and better quality of recommendations were generated. Although it was found that it considers only the recommendation problem of helping selective customers find which products they would like to purchase by suggesting a list of top n recommender products at a given point of time. Gopinath et. al.[5] used Rank correlation and pair-t-test for designing the recommender system. Its advantages are it is extensible to other domains like health care, world tourism etc. whereas its disadvantage is that it is not highly sensitive to slight changes.

#### **B. Personalization Services for Research Papers**

Academicians whether they are students, professors or researchers need to study a lot of research papers whenever they are working on research projects. Finding out useful and content specific research papers is highly essential for them. And hence the need for Personalization services in the areas of Research Papers was felt. The concept of Research Paper Tagging was used by Worasit[6] for research paper

recommendation generation. He calculated cosine similarity score between user profile and research paper index and if it reached a particular threshold level the paper was recommended. He has stated that the system he designed is extensible to online photo and video sharing. Although it was found that his study is based on certain assumptions which may or may not hold in practical scenario. Another work in the similar area was done by Chenguang et. al in [7]. He used Topic Analysis and thematic similarity and alleviated the cold start problem which is generally caused by CF technique. Much work has not been done if a new item has been registered in the system and apart from that the level of recommendation generation is also satisfactory.

#### **C. Personalization Services for E-learning**

E-learning refers to using the internet in the field of education for delivering lectures or any study material. As the students are now a day's getting attracted more and more towards internet E- Learning can help them a lot in studies too. But due to abundant data on WWW getting the desired information is indeed very difficult. Jinhua et. al [8] designed an adaptable personalized system for E-learning and demonstrated how it could be used and effectively applied. He utilized Association rules as data mining algorithm for the same. He proposed an architecture wherein the major work was done by recommendation engine module. It lacks the feature of learners past study pattern of user. Nava et. al. [9] identified learning styles using Felder Silverman Learning Style Model (FSLSM) and grouped learners of similar learning styles to adapt learning service accordingly. In their contribution they have considered the issue of clustering similar group of users if user profile data set is there. Its advantage is that it targets smaller data set and provides solution to learners learning styles and expertise level. Another model was proposed by Vincenza et. al. in [10] which offers courses personalization for both teachers and students.

#### **D. Personalization Services for Web News**

Personalization of Web News and blog posts is another important area which requires attention of research community. Ida [11] incorporated Web Usage mining for designing such a system. The author has designed a new strategy called Early Adopter Graph (EAG) which would generate interesting recommendations to the user. Its advantage is that it leverages not only on user similarity but on latent temporal patterns of users visits to web pages also.

#### **E. Personalization Services for Curriculum Resources**

Qing et. al [12] made use of Semantic web technology for designing Personalization services. By utilizing Domain Ontology and Nearest Neighbor Technique it was observed that service quality of teaching was significantly improved. It requires domain specific knowledge and has not been tested for extensibility to other application areas.

**F. Personalization Services for Smart Phones**

Fu et. al. [13] used a model called Client Interest Model for generating recommendations. For designing this service record of applications from logs of mobile phones was taken. Interest degree and weight vector were calculated and accordingly sorting of applications was done. Problem of information overload and resource disorientation were resolved and true Client Interest was exploited. Its disadvantage is that client’s privacy has not been taken into consideration.

**G. Personalization Services for Generalized Areas**

It has been stated earlier also that Personalization services are required in almost all domains. Some of the researchers have hence given their contribution in a very generic way i.e. not specifically for a web application. Qingyan et. al.[14] used Collaborative Filtering and Topic Aware Markov Model for recommending web pages. They stated that by using this technology lesser response time was achieved. They also consider topical information of web pages while recommendation which is ignored by majority of the other recommender systems. In another work by Shiva et. al [15] a combination of content and collaborative was used. Although it generated interesting recommendations but users’ feedback was completely ignored in this. Yechun et. al. [16] integrated personalized user and item based CF in Personalized Hybrid CF. They have shown that the hybrid approach used here leads to significant improvement in performance. Sule et. al [17] based on Sequence of visiting web pages and time spent on each page designed a recommender system which offered very high personalization. The advantages seen were search space was reduced, recommendation time was reduced, and this model could be used on websites with complex structures. However it considered only single visit to website.

**IV. COMPARATIVE ANALYSIS**

This section brings out the comparison of various techniques used in leading web-based application areas such as E-commerce, Internet Shopping Mall, E-learning, Research Papers etc. The comparison is carried out on the basis of parameters viz. domain knowledge, degree of personalization, access time and user profiling. The parameters chosen are essential but not necessarily sufficient for any web-based application. But for an efficient and effective web-based application, the application must adhere to these parameters. Efficiency aspect is related to how many new customers it can attract. Effectiveness deals with ability of the system to retain its existing customers. Depending upon the type of Web Application being designed apart from the parameters suggested a few other things can be considered while designing Personalization Services like feedback of user, past study pattern of the learner, whether the user has previously visited the website, privacy of the customers ( thereby

increasing customers loyalty). As we are talking about personalization in web-based applications, the very obvious parameter is the degree of personalization being offered. It refers to extent the designed system can help in finding relevant information and the [3] quality of recommendation provided to the user. Higher the quality of recommendation, higher is the degree of personalisation. For providing the user with relevant information and recommendation, domain knowledge is the key parameter which must be taken care of while designing the system. Better domain knowledge helps in attaining the system scalability and increasing its prediction accuracy [18]. Another vital parameter is User profiling, which takes care of the user’s interest and preferences. User profiling can be judged in a better manner by knowing the user’s hobbies, interest areas [19]. Better the user profiling, better is the degree of personalization. Next factor is the access time. This tells us about the time spent by the user on each web page. This factor is essential to record, because it will tell the interest of the user in which particular web page he is concerned more, which will ultimately help in refining the quality of recommendation and in turn increasing the degree of personalization.

Henceforth the parameters decided will help in ultimately deciding the degree of personalization being offered by the application. The parameters selected have been rated. Domain Knowledge, Access Time Consideration and User Profiling have been rated in terms of Yes or No. However the parameter Degree of Personalization offered has been given the values Satisfactory, High and Very High.

TABLE I : COMPARATIVE ANALYSIS OF VARIOUS TECHNOLOGIES BASED ON PARAMETERS

Name of Technology	Application area	Domain Knowledge Required	Degree of Personalization Offered	Access Time Consideration	User Profiling
Collaborative Clustering	E-commerce	No	High	No	Yes
Web Usage Mining and Decision Tree Induction	Internet shopping Mall	No	High	No	Yes
Rank correlation	Web based	No	High	No	Yes

Technique and Pair-test	Commerce Applications				
Research Paper Tagging	Research Papers	No	High	No	Yes
Topic Analysis		No	Satisfactory	No	Yes
Association Rules	E-learning	No	High	No	Yes
FSLSM		Yes	High	No	Yes
Web Usage Mining	Web News and Blog Posts	No	High	No	Yes
Semantic Web	Curriculum Resources	Yes	High	No	Yes
Client Interest Model	Smart Phones	No	High	No	No
Collaborative Filtering and Topic Aware Markov Model	Generic	No	High	No	Yes
Hybrid		No	High	No	Yes
Personalized Hybrid Collaborative Filtering (PHCF)		No	High	No	Yes
Click Stream Tree representation of Users' clusters		No	Very high	Yes	Yes

The comparison shows that User Profiling has been implemented majorly whereas access time has not been considered much till date while designing personalization services. Domain Knowledge has also been ignored during implementation. However degree of Personalization being offered is high or more than that every time and has been given due importance in almost all the cases.

### V. CONCLUSIONS

Personalization Services are gaining popularity in almost all sections of WWW. Major areas where personalization is offered were studied and it was found that extensive work has been done in e-commerce domain. Although a variety of technologies are there which can be used for designing personalization service but majorly CF is being used. CF when used alone can lead to various disadvantages which have been stated earlier. Not only CF but almost all technologies if used alone can lead to several disadvantages. So a blend of two or hybrid technology can lead to better results and offer higher degree of personalization. User's feedback, access time criticality, user profiling, domain knowledge are some of the factors which should be taken into consideration while designing Personalization services. Apart from these, few other parameters have been suggested which can be considered depending upon the type of web application for which Personalization Service is being designed.

### REFERENCES

- [1] K. Sridevi and Dr. R. Umarani , "Web Personalization approaches : A survey", International Journal of Advanced Research in Computer and Communication Engineering, Vol. 2, Issue 3, pp. 1533-1539, March 2013.
- [2] Luo Ya, "The Comparison of Personalization Recommendation for E-Commerce", in Proc.2012 ICSSDMS, Elsevier, 2012, p. 475.
- [3] Yoon ho cho, Jae Kyeong Kim and Soung Hie Kim, " A personalized recommender system based on web usage mining and decision tree induction" , Expert Systems with Applications: 2002 Elsevier Science Limited, vol. 23, pp. 329-342, 2002.
- [4] Rumeil Lian , "The construction of Personalized Web Page recommendation system in E-commerce", IEEE, pp. 2687-2690, 2011.
- [5] Gopinath Ganapathy and K. Arunesh , " Models for recommender systems in web usage mining based on user ratings", in Proc. WCE 2011, 2011.
- [6] Worasit Choochaiwattana , " Use of tagging for research paper recommendation", in Proc. 2010 ICACTE ,IEEE , 2010, p. 439.
- [7] Chenguang Pan and Wenxin Li, "Research Paper Recommendation with Topic Analysis", in Proc. 2010 ICCDA, IEEE, 2010, p. 264.
- [8] Jinhua Sun and Yanqi Xie, " A Recoomender System Based on Web Data Mining for Personalized E-learning" , IEEE, 2009.

- [9] Nava Jyothi, Kaveri Bhan, Uday Mothukuri, Sandesh Jain and Dhanander Jain, "A Recommender system assisting instructor in building learning path for Personalized Learning system", in Proc. 2010 ICICCI, IEEE, 2010, p.214.
- [10] Vincenza Carchiolo, Alessandro Longheu, Michele Malgeri and Giuseppe Mangioni, "Courses Personalization in an E-learning environment", in Proc. 3rd IEEE ICALT, 2003.
- [11] Ida Mele, "Web Usage Mining for Enhancing Search Result Delivery and Helping Users to find interesting web content", in Proc. WSDM'13, ACM, 2013, p.765.
- [12] Qing Yang, Junli Sun, Jinqiao Wang and Zhiyong Jin, "Semantic web based personalized recommendation system of courses knowledge search", in Proc. IEEE ICWS 2011, 2011, p.211.
- [13] Fu Changhong, Yang fan, Wu Shunxiang and Zhu Shunzi, "Research of Personalized Recommendation System based on Client Interest Model in the smart Mobile Phone", in Proc. 5th ICCSE, IEEE, 2010, p. 371.
- [14] Qingyan Yang, Ju Fan, Jianyong Wang, and Lizhu Zhou, "Personalizing Web Page Recommendation via Collaborative Filtering and Topic Aware Markov Model", in Proc. 2010 IEEE ICDM, 2010, p. 1145.
- [15] Shiva Nadi, Mohammad Hossein Sarraee and Ayoub Bagheri, "A Hybrid recommender system for dynamic web users", International Journal Multimedia and Image Processing, vol. 1, Issue 1, pp. 3-8, March 2011.
- [16] Yechun Jiang, Jianxun Liu, Mingdong Tang, and Xiaoqing Liu, "An effective Web Service Recommendation Method based on Personalized Collaborative Filtering" in Proc. 2011 IEEE ICWS, 2011, p. 211.
- [17] Sule Gunduz and M. Tamer Ozsu, "A web page Prediction model based on click stream tree representation of user behavior" in Proc. SIGKDD '03, ACM, 2003, p.535.
- [18] Li, Tao and Anand, Sarabjot S, Exploiting Domain Knowledge by Automated Taxonomy Generation in Recommender Systems, ser. Lecture Notes in Computer Science: E-Commerce and Web Technologies. Di Noia, Tommaso, E Buccafurri, Francesco: Springer, 2009, vol. 5692.
- [19] Mu Zhang, "Research of Personalization Services in E-commerce Site based on Web Data Mining", in Proc. ICCIC 2011 IEEE, paper ICCIS.2011.215, p.438

### Authors Profile



**Ms. Neha Goel** received her M.C.A. degree from GGSIP University, Delhi. She is currently pursuing her Ph.D. from Banasthali University, Rajasthan and also working as Assistant Professor in an institute affiliated to GGSIPU, Delhi. She is NET qualified in Computer Science and

Applications. She has published several research papers in various National and International Journals of repute. She has also presented research papers in various Conferences at National level. She has guided more than fifty research projects for M.C.A. students. She has an experience of more than 6 years in both industry and academia. She has been actively involved in organizing conferences at both national and international level.



**Dr. C.K. Jha** has done his Ph.D. from Banasthali University, Rajasthan. He has done his Masters in Computer Applications (M.C.A.) from BITS, Ranchi and is also a CCNA Certified instructor. He is a member of renowned bodies like IEEE, CSI and Indian Science National Congress. He has been

awarded as best instructor by CISCO Networking Academy and has an extensive experience of more than 24 years in both teaching and academia. Currently he is working as Head and Associate Professor of Computer Science, Engineering and IT department, Banasthali University (Rajasthan). He has more than 41 research papers to his contribution which have been published in various National and International Journals of repute. He has also written a book on Data Communications and Networking. He has been actively involved in projects sponsored by UGC, MCIT and DST, Govt. of India. He is amongst one of the paper reviewers for reputed journals like IEEE, Springer, IJCA, IJETT, IJCSIT etc. He has been guiding several M.Tech and Ph.D. students.