3. PR Methodologies

3.1 Community web directories

Web directories which contains groups of users known as user communities of similar interest. The construction of user community models, i.e., usage patterns representing the browsing preferences of the community members. The members of a community can use the community directory as a starting point for navigating the web, based on the topics that they are interested in without the requirement of accessing vast web directories. Instead of using web directory for personalization, it personalizes directory itself [1].

3.2 Temporal Web Access Pattern

3.2.1 PR through Knowledge base

Knowledge base was constructed by using temporal web access pattern as input. The system that generates personalized recommendations in a timely manner by mining periodic access patterns, which occurs frequently in a particular period. Such periodic access patterns are very useful for understanding user's web access habits and behaviours. With such knowledge, it is possible to deduce and prepare the web resources that the user is most probably interested in during a specific time period without the need for user's current web access information for effective personalized web content recommendation services [2].

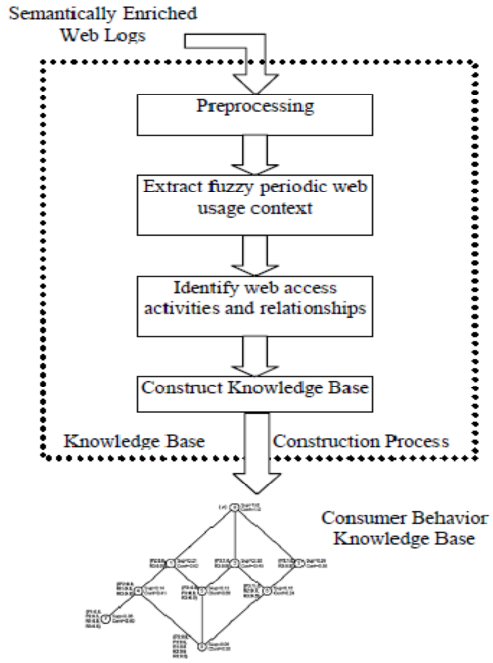

Figure 4: Knowledge base construction process

3.2.2 Personalized ontology based on users emotional and behaviour analysis

Semantic web usage mining approach was used for automatic generation of periodic pattern based web usage ontology for the semantic web. Periodic access patterns were directly mined from web usage logs that have been semantically enriched with information on topics and emotional influence. Second is personalized web usage lattice which consists of web access activities was constructed. Finally periodic pattern-based usage ontology was generated from the web usage lattice. Web ontology language (OWL) was used for ontology specification [18].

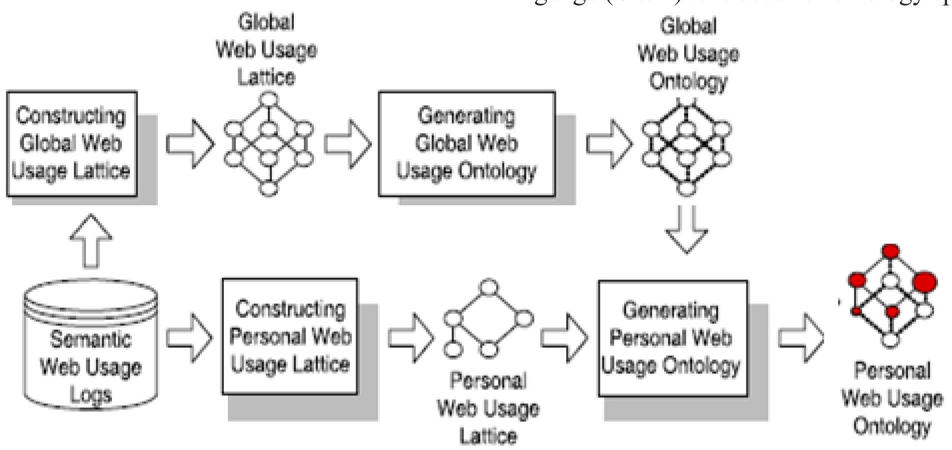

Figure 5: Web usage ontology generation

3.2.3 Tuning and Binding approaches

Tuning and binding clustering approaches was used to cluster a user's based on timely access behaviour and page preference. These two approaches define clusters with users that show similar visiting behaviour at the same time period, by varying the priority given to page or time visiting. Tuning will approach initiates with the page preferences and then proceeds to the time aspect while the second follows the reverse logic. Vector space model was used for user's representation [16].

3.3 Web page prediction

Web prediction is a classification problem in which we attempt to predict next set of web pages that a user may visit based on the knowledge of previously visited pages. Data are extracted from the logs of web servers, and they contain sequences of pages that users have visited along with the visit date and duration. The web prediction problem (WPP) can be generalized and applied in many essential industrial applications such as search engines, caching systems, recommendation systems, and wireless applications.

Licensed Under Creative Commons Attribution CC BY

Improving prediction process can reduce the user's access times while browsing, and it can ease network traffic by avoiding visiting unnecessary pages. For accurate prediction, first simple probabilistic models such as association rule mining (ARM), Markov model and all Kth model in the WPP utilizing different N-grams. Second is two-tier framework to reduce prediction time in case of consulting many classifiers to resolve prediction. Third is Modified Markov model that handles the excess memory requirements in case of large data sets by reducing the number of paths during the training and test phases [3].

3.4 Multiple criteria decision making approach (MCDM)

Raising customer loyalty in electronic commerce requires an emphasis on one-one to marketing and personalized services. It is essential to understand individual customer preferences for products. The identification and recommendation of such products are all based on the use of customer's real time web usage behaviour, including activities such as viewing, basket placement, and purchasing of products. Information on web usage behaviour for the products determines the ordinal relationship among the products, which express that certain products is preferred to other products across the multiple aspects. The ordinal relationships among the products and the multiple aspects of products lead to the consideration of multiple-criteria decision making approach (MCDM).

MCDM can be effectively utilized in the evaluation of alternatives (i.e., products) since it is a method of evaluating alternate decisions under a finite number of criteria (e.g., specifications). Thus, if a user directly states that he or she prefers one product over another, or if such information is indirectly obtained from the web usage behaviour, it is reasonable to assume that the value of preferred alternative is at least as great as that of the less preferred alternative. A well-known method for the valuation of alternative over multiple criteria is to use an additive form in which the values (scores or performance) of an alternative, measured with respect to each of the criteria, are added to obtain the overall value of the alternative across the multiple criteria. This definition can be expressed by the function

$$V(a_l) = \sum W_i V_i(a_l)$$

Where a_l alternative under consideration, $V_i(a_l)$ is the known value of a_l with respect to the ith criterion, and w_i is the unknown weight attached to to the ith criterion [4].

3.5 Formal concept analysis (FCA) in social tagging

3.5.1 Formal context

It is a mathematical theory used for conceptual data analysis and unsupervised machine learning. FCA models the world of data through the use of objects and attributes. The relations between objects and attributes in a data set form the formal context. This is represented by the triple (G, M, I), where G is a set of objects, M refers to a set of attributes, and I belongs to G*M specifies the binary relation between G and M.

3.5.2 Lattice of Formal Concepts

Concept lattice created for the formal context. The lattice is known as Galois lattice. The lattice is partial set in which two any elements have both a least upper bound and a greatest lower bound. The lattice maps will order from the most general to the most specific concept, from top to bottom. The topmost concept (the largest sub concept) is called the supremum, and the concept at the very bottom (the smallest sub concept) is called the infimum. However, the visualization of a Galois lattice can become quite complex as the number of users and/or the number of tags become large. Another drawback is creation of generic lattice is not possible by aggregate user interest from multiple sources [5].

3.6 Collaborative filtering (CF), Implicit and Explicit attributes

CF recommends items that are liked by other users with similar interests. Preference tree was constructed by using implicit and explicit attributes. Explicit attribute-based module using server usage logs of learners and ratings. Implicit attribute-based module using weights of implicit attributes for each user using a genetic algorithm. Issue is many systems need to react immediately to online requirements and make recommendations for all users regardless of rating history on visited resources, which demands high scalability on CF system [6][7].

3.7 Inferring user search goals with feedback sessions in search engine

An approach used to infer user search goals by analysing search engine query logs. Feedback sessions are constructed from user click-through logs.

Search results	Click sequence
www.thesun.co.uk/	0
www.nineplanets.org/sol.html	1
www.solarviews.com/eng/sun.htm	2
en.wikipedia.org/wiki/Sun	0
www.thesunmagazine.org/	0
www.space.com/sun/	0
en.wikipedia.org/wiki/The_Sun_(newspaper)	3
imagine.gsfc.nasa.gov/docs/science/know_l1/sun.html	0
www.nasa.gov/worldbook/sun_worldbook.html	0
www.enchantedlearning.com/subjects/astronomy/sun/	0

Figure 6: Search engine query log. '0' in click sequence means unclicked

Second is to use an optimization method to combine an enriched URL to generate pseudo documents, which can effectively reflect the information need of user. Based on pseudo documents, user search goals then be discovered and depicted with some keywords [8].

3.8 Heat diffusion on web graphs

Different kind of recommendations are made on the web every day like movies, music, images, books, recommendations, query suggestions, tags recommendations, etc. These data sources are modelled in form of various types of graphs. Heat diffusion is a model which is used for mining web graphs for recommendations like query suggestions, image recommendations. Heat diffusion is a physical phenomenon in which heat flows from high to lower temperature. Same was applied on web

Licensed Under Creative Commons Attribution CC BY

graphs in which information propagation happens. Diffusion was applied on both directed and undirected graphs for query suggestions and image recommendations. However, applying heat diffusion on social recommendation will be needed in social based web applications [9].

3.9 NPRF approach in image retrieval

Content-based image retrieval (CBIR) is used in the image retrieval systems. However no of iterative feedbacks was needed for CBIR to refine and produce the results. Navigation-pattern-based relevance feedback (NPRF) approach will reduce the no of feedbacks by using the navigation patterns discovered from the user query log and query refinement strategies. The three strategies are query point movement (QPM), Query reweighting (QR) and Query expansion (QE). Using of above two methods with CBIR will increase the accuracy of image retrieving and produce personalized results to users. However scalability and incorporating with users profiles will needed to enhance this approach [11].

3.10 MSN in social recommender system

Online sharing systems gather data that reflects user's collective behaviour and their shared activities. This data can be used to extract different kinds of relationships which can be grouped into layers and which are the basic components of multidimensional social network (MSN). Layers are formed based on direct or object based relationships. Direct relationships are user's direct communications. Indirect relationships are user's communication with other user's through multimedia objects (MO) like comments, tags etc. Strength of each layer is analysed and used in recommender system to suggest personalized recommendations to user in social networks [13].

3.11Item recommendation in collaborative tagging systems

Item recommendation in collaborative tagging system is based on only two way correlations between user and item. Traditional recommender system for collaborative tagging was using this for recommendations. However this will not satisfy the requirements of web 2.0. Using of higher order singular value decomposition (HOSVD) algorithm will consider the three way correlations between user, item and tags for personalized item recommendation. The above approach will increase the recommendation accuracy by understanding user's needs more accurately in collaborative tagging system [14].

3.12 Combining domain and web usage knowledge

Domain knowledge can be extracted by using semantic information of web page. Semantic information of page will give the short and useful description about that the page. By combining that knowledge with the web usage knowledge effective page recommendation for a user can be achieved. However it is possible only when semantic knowledge of web page is proper and accurate [19].

3.13 Endorsements for domain analysis

Domain analysis is a labour-intensive task in which related software systems are analysed to discover their common and variable parts. Domain analysis for software projects depend upon the human effort for identifying potential features and also feature identification is limited to the knowledge of domain analyst. Incremental diffusive algorithm (IDC) is used to extract raw features from online project repositories and group together extracted product feature into an aggregate representation of candidate feature. Association rule mining (ARM) is used to recommend based on the feedback of user profile and K-nearest neighbour approach to make additional recommendations. Recommendation based on analysing more than project repositories will give more common features recommendations for domain analysis produce quality products [10].

4. Conclusion

This paper surveyed various personalized recommendation methodologies through web usage mining. In every works researchers mainly depend on web usage data from server log files of each user. Usage data will be extracted and refined to remove unnecessary noisy data from it. And those data were combined with some approaches like association rule mining, k-nearest neighbour to provide recommendations. The no of areas where improvement needed is scalability problem in community web directories, getting temporal web access patterns, recommendations in collaborative tagging, domain analysis, scalability and profile enrichment in image retrieval. Researches can be carried out on above areas will improve personalized recommendations in web usage mining.

References

[1] Dimitrios Pierrakos, Member, Georgios Paliouras, "Personalizing web directories with the aid of web usage data", IEEE TRANSACTIONS Or KNOWLEDGE AND DATA ENGINEERING, VOL. 22,NO. 9, SEPTEMBER 2010.

[2] A.C.M. Fong, Senior Member, Baoyao Zhou, S.C.Hui, Senior Member, Guan Y. Hong, and The Anh Do,"Web content recommender system based on consumer behaviour modeling", IEEE TRANSACTIONS ON CONSUMER ELECTRONICS, VOL. 57, NO. 2, MAY 2011.

[3] MamounA.Awad and Issa Khalil,"Prediction of User's Web-Browsing behaviour: Application of markov model",IEEE TRANSACTIONS ON SYSTEMS, MAN, AND CYBERNETICS-PART B: CYBERNETICS, VOL.42.NO.4,AUGUST 2012.

[4] Duke Hyun Choi and ByeongSeokAhn, "Eliciting customer preferences for products from navigation behaviour on the web: A Multicriteria decision approach with implicit feedback", IEEE TRANSACTIONS ON SYSTEMS, MAN, AND CYBERNETICS-PART A: SYSTEMS AND HUMANS, VOL.39.NO.4, JULY 2009.

[5] Hak-Lae Kim, John G.Breslin, Stefan Decker, Hong-Gee Kim,"Mining and Representing User Interests: The Case of Tagging Practices", IEEE

Licensed Under Creative Commons Attribution CC BY

TRANSACTIONS ON SYSTEMS, MAN, AND CYBERNETICS-PART A: SYSTEMS AND HUMANS, VOL.41.NO.4, JULY 2011.

[6] Fabian Abel, Iglbert Bittencourt, Evandro Costa, Nicola Henze, Daniel Krause, Julita Vassileva, ”Recommendations in online discussion forums for E-Learning systems”, IEEE TRANSACTIONS ON LEARNING TECHNOLOGIES, VOL. 3, NO. 2, APRIL-JUNE 2010.

[7] Mojtaba Salehi, Isa Nakhai Kamalabadi, Mohammed B. Ghaznavi Ghoushi, Member,“An Effective recommendation framework for personal learning environments using a learner preference tree and a GA”, IEEE TRANSACTIONS ON LEARNING TECHNOLOGIES, VOL. 6, NO. 4, OCTOBER-DECEMBER 2013.

[8] Zheng Lu, Student Member, Hongyuan Zha, Xiaokang Yang, Senior Member, Weiyao Lin, Member, ZhaohuiZheng, “A new algorithm for inferring user search goals with feedback sessions”, IEEE TRANSACTIONS ON KNOWLEDGE AND DATA ENGINEERING, VOL. 25, NO. 3, MARCH 2013.

[9] Hao Ma, Irwin King, Senior Member, Michael Rung-Tsong Lyu, Fellow, ”Mining web graphs for recommendations”, IEEE TRANSACTIONS ON KNOWLEDGE AND DATA ENGINEERING, VOL. 24, NO. 6, JUNE 2012.

[10] Negar Hariri, Carios Castro-Herrera, Member, Mehdi Mirakhorli, Student Member, Jane Cleland-Huang, Member, Bamshad Mobasher, Member, “Supporting domain analysis through mining and recommending features from online product listings”, IEEE TRANSACTIONS ON SOFTWARE ENGINEERING, VOL. 39, NO. 12, DECEMBER 2013.

[11] Ja-Hwung Su, Wei-Jyun Huang, Philip S. Yu, Fellow, Vincent S. Tseng, Member,“Efficient relevance feedback for content-based image retrieval by mining user navigation patterns”, IEEE TRANSACTIONS ON KNOWLEDGE AND DATA ENGINEERING, VOL. 23, NO. 3, MARCH 2011.

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

[13] Przemyslaw Kazienko, Katarzyna Musial, TomazKajdanowicz,“Multidimensional social network in the social recommender system”, IEEE TRANSACTIONS ON SYSTEMS, MAN, AND CYBERNETICS-PART A: SYSTEMS AND HUMANS, VOL.41.NO.4,JULY 2011.

[14] Alexandros Nanopoulos, “Item recommendation in collaborative tagging systems”, IEEE TRANSACTIONS ON SYSTEMS, MAN, AND CYBERNETICS-PART A: SYSTEMS AND HUMANS, VOL.41.NO.4, JULY 2011.

[15] Olfa Nasraoui, Member, Maha Soliman, Member, EsinSaka, Member, Antonio Badia, Member, Richard Germain, “A Web usage mining framework for Mining evolving user profiles in dynamic web sites” IEEE TRANSACTIONS ON KNOWLEDGE AND DATA ENGINEERING, VOL. 20, NO. 2, FEBRUARY 2008.

[16] Sophia G.Petridou, Vassikili A. Koutsonikola, Athena I. Vakali, Georgios I. Papadimitriou, Senior Member, “Time Aware Web Users Clustering” IEEE TRANSACTIONS ON KNOWLEDGE AND DATA ENGINEERING.

[17] Sharon Paradesi, Oshani Seneviratne, Lalana Kagal, Decentralized Information Group CSAIL, MIT, “Policy Aware Social Miner” IEEE CS Security and Privacy Workshops.

[18] A.C.M. Fong, Senior Member, Baoyao Zhou, Siu C. Hui,Jie Tang, Member, Guan Y. Hong, Member, “Generalization of Personalized Ontology Based on Consumer Emotion and Behaviour Analysis”, IEEE TRANSACTIONS ON AFFECTIVE COMPUTING, VOL. 3, NO. 2, APRIL-JUNE 2012.

[19] ThiThanh Sang Nguyen, Hai Yan Lu, Jie Lu, “Web-page Recommendation based on Web Usage and Domain Knowledge”, IEEE JOURNALS.

[20] Mehdi Adda and Petko Valtchev University of Montreal RokiaMissaouiUniversity of Quebecn Outaouais Chabane Djeraba University of Lille, France, “Toward Recommendation Based on Ontology- Powered Web-Usage Mining”, International Conference on Internet Computing held on July. August 2007, Published by IEEE COMPUTER SOCIETY.

Author Profile

M. N. Shrigowtham received the B.E degree in Computer Science and Engineering from Hindustan College of Engineering, Chennai in 2011. He is currently doing his M.E in Nandha Engineering College, Erode, India.

S. Kavitha received the M.E degree in Computer Science and Engineering from Nandha Engineering College. She is currently working as Assistant Professor in Nandha Engineering College, Erode, India.

Licensed Under Creative Commons Attribution CC BY