International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056
Volume: 06 Issue: 10 | Oct 2019 www.irjet.net p-ISSN: 2395-007

A Literature Review and Classification of Semantic Web Approaches for Web Personalization Research

Sahaya Chithra. E[1], Dr. S. John Peter[2]

[1]Sahaya Chithra.E, Research Scholar, Department of Computer Science, St.Xavier's college, Tirunelveli, India.
[2]Dr. John Peter, Associate Professor and Head, Department of Computer Science, St.Xavier's college, Tirunelveli, India.

Abstract - *Nowadays, the biggest challenge on web is how to retrieve effectively and efficiently what has been actually requested. Users need help for searching the relevant information based on their preferences. Users spend a lot of time without getting satisfactory results. It can be tedious and time-consuming. One approach in satisfying the user is to personalize the information available on the web, called web Personalization. Web Personalization can be achieved by recommending web pages to the users, or by filtering web pages that are of interest to the users. Recommendation systems are software tools, which support Personalization, and its main objective is to help the individuals to select the most promising from the Internet. Semantic web Personalization can solve the information overload problem, and its objective is to provide users with what they really want or need. By considering all the benefits of the web Personalization, this paper presents the Semantic web for Personalization, and the classification of Semantic web approaches used for web Personalization and its related work in every approach.*

Keywords -- **Internet, Recommendation Systems, Semantic web, web Personalization.**

1. INTRODUCTION

As of January 2019, there were over 1.94 billion websites on the Internet. The large volume of information has made many problems that relate to the increasing difficulty of searching in organizing, accessing, and maintaining the relevant information. As the rapid growth of information load on the internet, personalized recommendation systems have become crucial tools for finding users preferences. The most preferable recommendation in all over the world is Spotify, Hulu, Netflix, Pandora, Amazon, and Sigchi. Amazon and Netflix suggest products and movies based on each user's profile, previous purchase history, and online behavior. Facebook and Google news recommend news articles based on user profiles and behavior. Companies are using recommender systems because of increasing sales and attracting customers with a personalized offer as a result of an enhanced customer's positive feedback. Recommendations typically speed up searches and make it easier for individuals to access own content only, and making surprised with offers, the users would have never searched for.

Personalization has been nowadays widely used in e-commerce as valuable tools for personalizing purchase recommendations. Personalization is something that every website or app that provides a reliable interface of a user must have. It prevents users from wasting time in access to what they need and encourages efficient search. Various factors help to acquire the user's preferences when surfing the internet. However, today's searching tools give the irrelevance information of many searches. Therefore, the ultimate need nowadays is that of predicting the user needs in order to improve the usability and user retention of a website. Semantic web can develop languages for expressing information in machine-readable form therefore Semantic web is the most appropriate platform for realizing Personalization [1].

Second section of this paper provides related work in the area of web Personalization and then section 3 will be devoted to survey in the area of web Personalization. This paper continues in Section 4, which describes the various Semantic web techniques and its pros and cons for web Personalization.

2. RELATED WORK

P.Markellou et al., [8] proposed a framework for personalized E-Learning based on collective usage profiles and domain ontology. The authors distinguished two stages online and offline tasks. An Offline task includes ontology creation, data preparation and usage mining and online tasks that include the production of recommendations. Baoyao Zhou et al., [9] proposed a web usage mining approach for semantic web personalization. Provide a semantic web personalization needs to challenge the technical issues on web access activities, convert them into ontology automatically discover hierarchical relationships from web access activities, and presume personalized usage knowledge from the ontology. This approach combined fuzzy logic into Formal Concept Analysis to mine client side web usage data for automatic ontology generation, and then applied fuzzy approximate reasoning to deduce personalized usage knowledge from the ontology.

Charbel Obeid Liris et al.,[10] Proposed recommender system is an assessment tool for students' vocational strengths and weaknesses, interests and capabilities. This proposed ontology-based recommender

system aims to identify the student requirements, interests, preferences, and capabilities to recommend the right major and university for each one. To support personalization, this study proposes a recommender system that can be used to assist students in finding the relevant major and university, by using the semantic web and machine learning techniques. The Semantic web is an extension of the current web, in which information is given in defined meaning, enabling machines and users to cooperate [24]. Moreover, machines will become able to process and "understand" the data. Ontology is the backbone of the Semantic web. Ontology is a recognized and explicit specification of a shared conceptualization. It consists of concepts namely entities, attributes, and properties related to a domain.

Hilal Karaman [11] proposed Hybrid System a ReMovender, which is a content-based and Collaborative movie recommendation system. Here collaborative missing data prediction technique is used to compare the various users' interests with content information of movies. Using ReMovender recommendation system users are requested to rate movies on a scale from one to five. Collaborate Filtering predicts the missing ratings data and similarities from ratings which are given by different users. A Content-based technique is used to correlate the movies by using a set of movie information and provide recommendations for the users.

Gizem Ozturk [12] proposed a hybrid system for personalizing videos to users, which uses both collaborative filtering and content-based approaches. By merging different approaches, it yields more powerful results than using pure methods individually. In this thesis, Adsorption algorithm is enriched by the Content-based approach for providing better suggestions. Besides using rating archives, video and movie content information are also used to suggest new items which help to reinforce recommendations. The system gives better results when the data is not sparse. That is why results with MovieLens data is better than YouTube. Ridhika Malik, Kunjana Vasudev and Udayan Ghose [13] proposed a Human Resource Semantic web management System (HRSWM) that enables the organizations to achieve their goals by transforming disparate, fragmented data into viable information capable of answering key questions. This study defines a system which combines the benefits of Resource Descriptive Framework (RDF) [24] with the high performance of database management systems and thus improves the performance in retrieving information at real time.

Nizar R. Mabroukeh et al., [14] proposed a generic framework called SemAware that integrates Semantic information into web usage mining. Semantic information can be combined into the pattern discovery. A Semantic distance matrix is used in the fixed sequential pattern mining algorithm to trim the search space and partially relieves the algorithm from support counting. A 1st-order Markov model is used to make the mining process and enriched with Semantic data. Yao et al., [15] proposed a framework which acts as an intelligent agent that dynamically gives the recommended websites for users. This can be learning from web usage data and users' behavior and is called as PagePrompter. Like a guide, an agent supports a user in navigating the website. PagePrompter can also be used as a tool for understanding user behaviour, the design of the website, system performance analyzing, website designer for improving websites and generating an adaptive website.

Punam Bedi, Harmeet Kaur and Sudeep Marwaha [16] presented system based on temporal ontologies and trust network, generates the recommendations across multiple domains. The similarity measures are taken care in the form of trust update process. In this Personalization system, Intuitionistic Fuzzy Sets (IFS) are been used to capture uncertainty, inherent in the recommendation process. The tourism domain has a number of subdomains like air travel, geography, food, entertainment, sports etc. is taken as a landmark study. Now the case study uses two domains destination and travel and these can be enhanced easily for other domains. In this proposed model, the recommendations are taken from the trustworthy agents only and the data as well as the methods used to generate the recommendations are with agents, making the recommendation process transparent to the users.

3. SEMANTIC WEB FOR PERSONALIZATION

Personalization is defined as "the ability to provide content and services tailored to individuals based on knowledge about their preferences and behavior" or "the use of technology and customer information to tailor electronic commerce interactions between a business and each individual customer"[2]. The Semantic web is an extension of the current web in which information is given well defined meanings that improved with processable between machines and users. The purpose of Semantic web is to dedicate most of tasks and decisions to machines. Semantic web is used in web Personalization for expressing information in a machine-readable form.

Wu et al., [3] proposed a general learning services personalized framework. Semantic technology, ontology to represent the learning contents and user profile are used in this framework. Social relationships with trusted friends are to improve the collaborating filtering algorithm to recommend personalized learning resources for users. Experiments showed that the proposed framework is effective. The experiment system includes 10 courses in computer field and a total of 3806 learning resources. This study used Recall and Precision to determine the efficiency of the proposed framework. This study selected 10 users and built the user domain profile and the user social relationship profile for each user using Generating User Domain Knowledge algorithm and ReRanking algorithm respectively. Mahendra Thakur and Geetika S. Pandey[4]

proposed a Personalization framework using the Semantic web. In this study, the visitors' navigational behavior is updated with this semantic knowledge to create an enhanced version of weblogs, Clogs, as well as semantic document clusters. Clogs are in turn mined to generate both a set of URI and category-based association rules. Finally, the recommendation engine uses these rules, along with the semantic document clusters in order to provide the final, semantically enhanced set of recommendations to the end user.

Kruti Jani and Dr.V.M Chavda [5] build and evaluate Semantics-aware collaborative filtering movie recommendation system using all the notions and elements of Linked Open Data. This method collects the user's preferences for items, search for a set of similar preferences by individuals, and infers rating on a particular item based on the information gathered from neighbors. Collaborative filtering methods first find the most similar users to the active user and consider them as a set of the neighbors of active user and finally predict and provide suggestions. The goal of the prediction of the target item's score is for active user. The members who voted for target item only have been studied in order to create a set of neighbors. Standard Pearson correlation criteria are used to evaluate the dependencies between scoring patterns of users and active user. When the similarity of all users with an active user is found then set of neighbors are chosen for further process. The scores offered for target item by a set of the neighbors can be predicted. This knowledge can be applied to design an advanced recommendation system with the ability to generate more individual and accurate prediction results.

Honghua Dai and Mobasher [6] proposed a study that provides an overview of semantic knowledge approaches on web usage mining and personalization practices. Semantic knowledge is the only solution to personalize systems, which is complex objects based on their underlying properties and attributes. Successful integration of Semantic knowledge from different sources, such as the content and the structure of web sites for Personalization are discussed in this study. Ahu Sieg, Bamshad Mobasher and Robin Burke [7] proposed a system which utilizes the user context to personalize search results by re-ranking the results returned from a search engine for a given query. This study has presented a framework for contextual information access using ontologies and demonstrated that the Semantic knowledge embedded in an ontology combined with long-term user profiles can be used to effectively trim the search results based on users' choices and likes.

4. METHODS FOR WEB PERSONALIZATION

4.1 Collaborative Filtering System

Collaborative filtering is one of the most successful and popular methodologies in recommendation systems. CF technique is used to compute Semantic similarities between users, if two users are having high Semantic similarity, the Personalization recommendations has been made. Collaborative filtering systems can deal bulky of users and multiple different items. However, traditional collaborative filtering has some limitations such as the item sparsity and cold start problem. Lee Jae-Won, Nam Kwang-Hyun, and Lee Sang-Goo [22] used a Semantic match with the domain ontology, while the traditional collaborative filtering uses an exact match to find similar users. Therefore, this methodology solved the item sparsity problem by mapping users and items to a set of concepts, which are the nodes of domain ontology. When recommending items, the effects of number of similar users are analysed. By mapping them with concepts, the item sparsity can be reduced. The precious of this study is mapping items and users to a set of concepts that are the nodes of ontology.

4.2. Content-based filtering System

When there is no or only little amount of user data is available, Content-based filtering is used. Content-based filtering recommends on basis of the content similarity between items. The contents may be structured or unstructured. Therefore, to make similarities between objects easier the contents should be converted into structured. The first step in Content-based filtering is to keep a private profile for each item. Then CB finds the similarities between the attributes or features of items, and finally, it recommends the most similar item for an item. Finding similarities between structured content is straightforward whereas deciding the features or attributes of the unstructured item is very difficult. Content-based Personalization systems recommend items similar to those a given user has liked in the past. Here user profile is matched up with preferences and interests, which are stored in past web surfing. Hilal Karaman [11] proposed a movie recommendation system named as *ReMovender*. This system makes the use of ratings given by users to movies, demographic data of the users and content of the movies in order to provide high recommendations. The feature weights for different set of individuals and movies are almost universal when this recommendation system is based on assumption. The result of this thesis scores 2.28, which is slightly better than the pure content based, which scores

4.3 Ontology-based System

Ontology based techniques for web Personalization just considers the domain ontology for Personalization without considering the context information and acquaintance information of the user, which may vary with time. Personalization has made by combining the results obtained by imposing restrictions with the user model, which represents individual user interest. Muhammad Shoaib, Amna Basharat [17] proposed a framework that

supports the creation of Semantic users and resources profiles in order to map the user's personal and research interests in an ontological format. These relations are used to deliver knowledge including research articles, researchers, and research groups of related fields, about user's interests on the social web in order to create users personalization. The users' are required to keep an updated users profile. The ontology allows storing of the user, group, posts, and events profile in machine processable format. In order to enable automatically personalization feature this research illustrated the need for ontological data representation followed by ontology structure for Social network data.

Ontologies offer a way to cope with these heterogeneous representations of web resources. These comprise the backbone of the Semantic web and appear as a promising technology for implementing in particular e-Learning applications. The reason ontologies are becoming so popular is due to what they promise: *"a share and common understanding of a domain that can be communicated between people and application systems"* [18].

4.4 Context-based Personalization

These techniques use the user preferences along with context information of the user for personalizing the user request. In this Context-based system, users' result may fluctuate depends upon the request which the user has given. Context-based Context awareness computes a broad range of contextual attributes such as the user's activities, current positions, and surrounding environments. The Contexts are user profile, user activity, and user search history. Jongyi Hong et al., [19] propose an agent-based framework for providing personalized services using context history on context-aware computing. The proposed framework consists of data gathering layer, context management layer, preference management layer and application layer for providing the personalized services based on the users' preferences. The prototype was implemented according to each layer. The system that offers each user with the personalized services under the same context was displayed using PDA. This research suggests the basic direction for the provision of the personalized services on context-aware computing and utilization of context history. In that, it indicated the need or usability of context history that was not considered in the previous researches. Additionally, this research can be the basic direction of design and the guidelines of development for the context-aware computing system. However, the prototype was not completely implemented according to the proposed framework because the surveyed data is used instead of real sensor data due to the constraints of tools and time. Also, the protection of personal information or privacy needs to be considered.

4.5 Trust-based Personalization

The techniques exploit the user acquaintance information to personalize the recommendations or search. Trust based Personalization system is based on the concept that acquaintances may share similar preferences. Trust-based Personalization techniques explore an area which is not covered in both ontology-based Personalization techniques and context-based Personalization techniques and may provide some interesting recommendations which may not be provided by the prior techniques. FOAF (Friend of a Friend) is used to make a trust for Semantic web Personalization. FOAF is used to identify a person browsing habits in the website. Personalization is performed by sending user FOAF information every time a user sends the request to the server this is done by extending the HTTP Get method to include the parameters that point to the URL of the user FOAF file. Carminati Barbara, Ferrari Elena, and Perego Andrea [20] proposed a web based social network where members are able not only to specify labels, but also to rate existing labels. Both labels and ratings are then used to assess the trustworthiness of resources' descriptions and to enforce web access Personalization. FOAF profiles can be signed to grant their authenticity, and thus they can be effectively used as certified credentials. Trust computation is a feature which is supported by some online communities, mainly in order to refine recommendations. This study it supports just four descriptors (i.e., trustworthiness, vendor reliability, privacy, and child safety), and it does not give users the possibility of expressing their disagreement with claims made by other users. Moreover, differently from our proposal, explicit trust policies or trust relationships are not supported, in that the trustworthiness of rating authors is computed by the system itself.

4.6 Rule-based Personalization

These techniques use rules to extract Personalization information. Here web usage data in the form of logs are used to infer the rule and make recommendations based on the matched patterns. Designing the rule leads sometimes, updating the rule made difficulty. Otherwise even the only little bit of web usage data is enough for Rule-based systems comparing other Personalization systems. Initially, a set of rules are designed and used to make personalized recommendations. From the session, user activities and characteristics of other users which are stored in web usage model are observed and using these available facets , Personalization is carried out. Barla Michal, Tvarozek Michal, and Bielikov Maria [21] described a user modeling approach based on automated acquisition of user behaviour and its successive rule-based evaluation and transformation into an ontological user model. A novel method of automated user modeling for adaptive Semantic web-based systems proposed based on comprehensive logs of user actions with Semantics and a rule formalism

designed to support potentially elaborate navigational patterns. An inference agent was designed a and implemented which was driven by a rule formalism capturing interesting patterns of user-system interaction. All knowledge the agent requires in order to process the log of events and update the user model is stored in these rules, thus making the agent highly configurable and reusable.

4.7 Hybrid Recommendation System

HRS… means combination of content based filtering, collaborative filtering or any other filtering. Innar et al., [23] have presented the design and methods used in the SMART MUSEUM platform, employing adaptive and privacy preserving user profiling. They described recommendation system relies on combining a Semantics/ontology based approach with a data mining/statistics based approach. Gizem Ozturk [12] proposed hybrid video recommendation system. Collaborative filtering algorithm makes recommendations with relating similarities between users. Sparsity and cold start problem occur in pure collaborative filtering. In order to overcome these problems, content based filtering is added. Content-based filtering uses the idea of suggesting similar items that match user preferences. In order to use content-based filtering, first, the base recommendation list is updated by removing weak recommendations. Following this, item similarities of the remaining list are calculated and new items are added to make the final recommendations. Thus, collaborative recommendations are empowered considering item similarities. Therefore, the developed hybrid system combines both collaborative and content-based approaches to produce suggestions that are more effective.

5. CONCLUSIONS

Semantic web Personalization system for web search, which not only gives user a set of personalized pages, but also gives user a list of domains the user may be interested in. Thus, the user can switch to diverse interests when the user is surfing on the web for information. Besides, the system focuses on the domains that the user is interested in, and will not waste lots of time on searching the information in the irrelevant domains.

Here this study presented the classification of Semantic web approaches for web Personalization. Semantic web incorporates machine-processable information to support user preferences in their task. When Personalization is applied to the Semantic web it recommends many advantages when compared to the traditional web. Semantic web combines Semantics with the unstructured data on the web so that intelligent techniques can be applied to get results that are more proficient. Personalization based Recommendation system is becoming increasingly popular because of social networking and Semantic web has a great potential for exploring new dimensions for Trust-based Personalization.

REFERENCES

[1] Brusilovsky P., Kobsa A., Nejdl W., The Adaptive web: Methods and Strategies of web Personalization, LNCS 4321. Springer, 2007.

[2] G. Adomavicius, A. Tuzhilin,"Personalization techniques: a process-oriented perspective, Communications of the ACM 48 (10) (2005) 83–90.

[3] Chanle Wu, Jiyian Wu, Yibo Chen†, Ming Xie, Xiaojun Guo, Weichen Xiong, "Personalized Learning Service Framework Based on Semantic Web Technologies", IEEE ©2011, 978-1-4244-9763-8/11/$26.00

[4] Mahendra Thakur , Geetika S. Pandey," Performance Based Novel Techniques for Semantic web Mining", IJCSI International Journal of Computer Science Issues, Vol. 9, Issue 1, No 1, January 2012 . ISSN (Online): 1694-0814

[5] Kruti Jani and Dr.V.M Chavda ," Personalizing Movie Recommendation Using Semantic Contents in Collaborative Filtering", International Global Journal for Research analysis , Volume-5, Issue-9, September - 2016 • ISSN No 2277 – 8160

[6] Honghua Dai, and B. Mobasher, Integrating Semantic Knowledge with web Usage Mining for Personalization', *Information Systems Journal*, 2009, 1-28.

[7] Ahu Sieg, Bamshad Mobasher, Robin Burke "Learning Ontology-Based User – Profiles: A Semantic Approach to Personalized web Search" IEEE Int. Informatics Bulletin, Nov 2007.

[8] P. Markellou, I. Mousourouli, S. Spiros, and A.Tsakalidis (Greece) 2005. Using Semantic web Mining Technologies for Personalied ELearning Experiences, Proceeding (461) web-based Education.

[9] Baoyao Zhou1, Siu Cheung Hui, and Alvis C. M. Fong 2006 web Usage Mining for Semantic web Personalization.

[10] Charbel Obeid Liris, Inaya Lahoud, Hicham El Khoury and Pierre-Antoine Champin Liris," Ontology-based Recommender System in Higher Education", *WWW '18 Companion April 23-27, 2018, Lyon, France.*© 2018 IW3C2 (International World Wide Web Conference Committee) ACM ISBN 978-1-4503-5640-4/18/04.

[11] Hilal Karaman ," A Content Based Movie Recommendation System Empowered By Collaborative Missing Data Prediction", Master Thesis, The Graduate School Of Natural And Applied Sciences Of Middle East Technical University, June 2010.

[12] Gizem Ozturk ,"A Hybrid Video Recommendation System Based On A Graph-Based Algorithm", Master thesis, Computer Engineering Department, Middle East Technical University. September 2010.

[13] Ridhika Malik, Kunjana Vasudev and Udayan Ghose,"Semantic web Mining and its application in Human Resource Management", IJCSMS International Journal of Computer Science & Management Studies, Vol. 11, Issue 02, August (2011) ISSN **(**Online): 2231 –5268

[14] Nizar R. Mabroukeh and Christie I,"Using Domain Ontology for Semantic web Usage Mining and Next Page Prediction,. Ezeife CIKM'09, pg 2–6.

[15] Y.Y. Yao, H.J. Hamilton, and Xuewei Wang, "An Intelligent Agent for web Navigation Created Using Data Mining", Volume 27 Issue 3, Pages 59 - 74. PagePrompter 2008.

[16] Punam Bedi, Harmeet Kaur, Sudeep Marwaha," Trust based Recommender System for the Semantic Web", IJCAI-07. 2677-2682

[17] Muhammad Shoaib, Amna Basharat ,"Ontology based Knowledge Representation and Semantic Profiling In Personalized Semantic Social Networking Framework", IEEE ©2010,978-1-4244-5540-9/10/$26.00

[18] J. Davies, D. Fensel, & F. van Harmelen, Introduction, *Towards the Semantic Web, Ontology- Driven Knowledge Managament*, John Wiley & Sons, 2003, 1-9.

[19] Hong Jongyi,-Ho Suh A, Eui , Kim Junyoung ,and Kim Suyeon, "Context-aware system for proactive personalized service based on context history"*, ,Expert Systems with Applications Elsevier 36 (2009) 7448–7457.*

[20] Carminati Barbara, Ferrari Elena, and Perego Andrea," Combining social networks and Semantic web technologies for personalizing web access, collaborative computing: networking, applications and worksharing" Lecture Notes of the Institute for Computer Sciences, Social Informatics and Telecommunications Engineering, Volume 10. ISBN 978-3-642-03353-7. Springer Berlin Heidelberg, (2009), p. 126

[21] Barla Michal, Tvaroˇzek Michal,and Bielikov M´Aria "A, Rule-based user characteristics acquisition from logs with Semantics for personalized web-based systems" Computing and Informatics, Vol. 28, (Sep-22 2009). 399–427

[22]Lee Jae-Won, Nam Kwang-Hyun, and Lee Sang-Goo,"Semantics based collaborative filtering", Springer-Verlag Berlin Heidelberg Software Engineering, Artificial Intelligence,(2009),201–208.

[23] Innar, Ruotsalo Tuukka, Tammet Tanel and Kuusik Alar, "Personalized context-aware recommendations in smartmuseum:combining Semantics with statistics" ,in proceedings of Third internal conference on Advances in Semantic processing,IEEE (2009) , 50 – 55.Penelope

[24] RDF, http://www.w3.org/RDF/, 01.06.2009