British Journal of Applied Science & Technology
12(5): 1-9, 2016, Article no.BJAST.21466
ISSN: 2231-0843, NLM ID: 101664541

SCIENCEDOMAIN *international*
www.sciencedomain.org

A Hybrid Recommendation Architecture for Nigerian Online Stores

A. T. Olaniran[1*], I. O. Awoyelu[1*], A. O. Amoo[1] and B. O. Akinyemi[1]

[1]Department of Computer Science and Engineering, Obafemi Awolowo University, Ile-Ife, Osun State, Nigeria.

Authors' contributions

This work was carried out in collaboration between all authors. Authors ATO, AOA and BOA designed the study. Authors ATO and IOA reviewed literatures on the study and wrote the first draft of the manuscript. Authors IOA, AOA and BOA managed the literature searches. All authors read and approved the final manuscript.

Article Information

DOI: 10.9734/BJAST/2016/21466
Editor(s):
(1) Orlando Manuel da Costa Gomes, Professor of Economics, Lisbon Accounting and Business School (ISCAL), Lisbon Polytechnic Institute, Portugal.
Reviewers:
(1) Jeevarathnam P Govender, Durban University of Technology, South Africa.
(2) Thi Huong Lan, Ho, Hue University, Vietnam.
(3) Wen-Shan Lin, National Chiayi Unversity, Taiwan.
Complete Peer review History: http://sciencedomain.org/review-history/11954

Short Research Article

Received 18th August 2015
Accepted 1st October 2015
Published 24th October 2015

ABSTRACT

Online retailing, a business activity in e-commerce, is the process of selling physical goods on the internet while online shopping is the process of buying and selling goods and services over the internet. Item recommendations, however, is a business strategy in e-commerce employed in promoting goods sold on an online store. Majority of the online stores in Nigeria have their shopping systems implemented similar to online shopping systems used Developed Countries. Much focus have been placed on the provision of non-personalized recommendations in Nigerian online stores, as ratings information needed for personalized recommendations is sparse. These systems are mostly a hybrid of content-based and collaborative filtering approaches to recommendation generation. The use of content-based, collaborative and demographic hybrids have not been fully explored and implemented. This paper, however, proposes a hybrid item recommendation architecture that combines the content-based, collaborative and demographic filtering approaches in a mixed hybridization strategy for provision of recommendations on online stores. The architecture provides collaborative recommendations using the vector similarity and adjusted cosine similarity measure. The proposed system will go a long way in providing adequate item recommendation in Nigerian stores.

**Corresponding author: E-mail: abiola_olaniran03@yahoo.com, iawoyelu@oauife.edu.ng;*

Keywords: Online stores; item recommendation; ratings information; purchase data; demographic data.

1. INTRODUCTION

Electronic Commerce (EC) is the process of carrying out business transactions via telecommunications networks, especially the Internet [1]. It is the conduct of financial transactions by electronic means. Preference for traditional shopping is gradually being reduced in Nigeria due to the ease of convenience provided by online shopping. However, online shopping is still slowly being adopted. This is due to the various challenges often encountered by consumers while shopping online, such as poor item recommendations, delayed delivery, poor customer retention schemes, insecure payment infrastructure, poor return and refund policies, unaffordable internet access charges.

E-commerce took off sharply in Nigeria with Konga, Jumia and the likes, retailing local and foreign goods. The key drivers of the Nigerian retail market system are large market size, youthful population and urbanization, while the key risk factors are naira devaluation (which often raises the cost of imported goods), poor infrastructure, difficulty in sourcing products, political and security risks [2].

To effectively carry out online shopping transactions, retailers organize online catalogs of products and take orders through their websites. They also have to accept payments in secure environments, send merchandise to customers and manage customer data. For customers to consider online products, a recommender system has to be used. Recommender systems have been defined by [3] as software tools and techniques providing suggestions to a user. They have become fundamental applications in electronic commerce and information access. They provide suggestions/recommendations on online stores that effectively prune large information spaces so that users can be directed towards items that best meet their needs and preferences.

One of the most successful applications in online shopping, according to [4], is 'personalized recommendation services'. Recommendations are aimed at supporting shoppers in their various decision making processes while carrying out shopping activities online. [5] and [6] distinguished these recommendations into personalized and non-personalized recommendations. Personalized recommendations are targeted towards meeting the individual needs of the users of a recommender system. They are generated using ratings of items on an online store. They are also sometimes generated using the demographics of shoppers, mostly for shoppers who lack sufficient ratings that can be used by the system to provide them recommendations. Non-personalized recommendations are targeted towards meeting the collective needs of all users of a recommender system. They are generated based on what other customers have said about certain products or items (product reviews), the top rated retailers on the store, items favorited by shoppers, analysis of shoppers' past buying behaviours. They are often presented as general advertisements on the store's homepage, discounts, offers and coupons.

Provision of quality item recommendations, especially personalized recommendations, has been quite challenging for Nigerian online retailers. This is because existing recommender models have been implemented majorly with foreign orientation. They do not capture the peculiar information needs and shopping preferences of Nigerian shoppers. An online shopping model system that addresses these problems is needed; hence, this paper aims at proposing a hybrid item recommendation architecture that demographically enhances the item recommendation processes of Nigerian online stores.

The proposed architecture combines the content-based, collaborative and demographic filtering processes in mixed hybridization. It also proposes demographic filtering for all shoppers on Nigerian online stores for provision of demographic recommendations. This is as opposed to existing models that consider demographic filtering for only shoppers with no prior ratings on the system. It also proposes the use of alternative data (purchase and content data) instead of ratings data in the content-based and collaborative filtering processes.

The rest of this paper is organized as follows. Section 2.0 discusses existing works on item recommendations while section 3.0 explains the proposed hybrid item recommender architecture. Section 4.0 concludes the paper.

2. EXISTING WORKS

Different techniques have been proposed for performing recommendations. Also, there are different types of filtering techniques used by recommender systems for recommendation generation. Some of these techniques are the content-based filtering, collaborative filtering, demographic filtering, hybrid filtering and rule-based filtering techniques. These techniques are explained in the following sections.

2.1 Collaborative Filtering

Collaborative filtering is the most familiar, most implemented and most established of all recommendation technologies [3]. In its simplest implementation, recommendations are provided to an active user based on items that other users with similar tastes have considered (liked, rated, or purchased) in the past.

2.2 Content-based Filtering

Content based filtering systems have their roots in information retrieval. This approach to recommendation is based on a user's specification of desired content to the system, the analysis of items previously rated by a user and profile generation. [7] described these systems as systems that analyze item descriptions to identify other items that can be of interest to the user.

2.3 Demographic Filtering

Demographic filtering recommender systems use demographic information about users as basis for providing recommendations [6]. Demographic recommendation or filtering systems aim to categorize users based on personal attributes and make recommendations based on demographic classes or profiles [8]. Demographic recommenders work similarly to collaborative ones in that they both find user similarities but use different types of data [9].

2.4 Rule-based Filtering

In a rule-based filtering system, the system uses a set of rules to recommend items based on the user's history. The system uses these rules to deliver contents to their users [10]. These rules are set by the system developer depending on what the users are allowed to do.

2.5 Hybrid Filtering

Hybrid filtering systems according to [8,10-12], combine two or more recommendation techniques to gain better performance with fewer of the drawbacks of any of the individual techniques. They capitalize on a combination of their strengths. Most commonly, collaborative filtering is combined with some other technique to avoid the cold-start or new user problem. The most common hybrid recommender systems are blends of the content-based and collaborative filtering techniques; and the blends of the content-based, collaborative and demographic filtering techniques. These techniques are often combined in order for the system to be able to alleviate the cold-start/new user problem.

Scholars such as [13-15,5,16-18], who have worked on recommender systems have proposed different methods for addressing some of the problems of recommender systems, such as the cold-start problem, the new-user and new-item problems, scalability problems, information overloading problems and so on. Some of such propositions include the development of hybrid recommender systems, introduction of demographic filtering for characterizing users and find similarities between them, using standard demographic vectors for finding user similarities, using standard similarity metrics for finding user and item similarities, enhancing the content-based and collaborative filtering techniques using alternative data and using dimensionality reduction techniques to scale through large data sets.

Deshpande and Karypis [13] addressed scalability concerns in user-based collaborative recommendation systems by analyzing the user-item matrices in order to generate recommendations without the algorithm spanning through customer data on the system. They however, did not consider the reduction of customer data stored on the system so as to reduce the data the algorithm spans through. The proposed system takes care of this problem. [19], on the other hand, addressed the non-utilization of demographic information in developing recommender systems by proposing enhanced collaborative filtering algorithms using demographic correlations for finding similarities in user or item neighbourhoods. They, however, did not consider the provision of demographic recommendations as taken care of in the proposed model.

Chen and He [15] addressed the cold-start problem (sparse/no ratings) of document recommender systems by finding user similarities based on demographic vector age for recommendation of documents. This reduced user efforts through ratings as no initial ratings by new users was required by the system for recommendation generation. Even though they enhanced the collaborative filtering process using demographic vector age, they however did not consider using the demographic vector to find similarities for all users in recommendation generation.

Bobadilla et al. [17] addressed the problem of finding a standard metric to measure user similarities for users of recommender systems in finding similar users for new users of the system. They proposed a standard similarity metric by using genetic algorithms to assign values and weights to users who have similar pattern of ratings to the user for whom recommendations are to be presented. They however did not explore using vector similarity methods for finding user similarities as presented in the proposed model.

Escriche and Symeon [9] reviewed the various steps undertaken in user profiling and personalization and modeled user needs for personalization in recommender systems by generating recommendations through the calculation of item similarities with user profiles using the vector space model for generation of users *'points of interest'* i.e. items often considered. Modeling the collective needs of users using their points of interest does not give room for modeling the individual needs of users.

Chikhaoui et al. [16] addressed the cold-start/new-user problems of recommender systems by proposing using demographics age, gender, race, disabilities, educational attainment, home ownership employment and occupation for finding user similarities using a neighbouring technique to form user clusters. They were able to achieve accuracy and high coverage in the system's recommendation and their algorithm outperformed conventional filtering algorithms as well as naïve methods. They however failed to consider the extraction of demographic information which are easily obtainable and will not be considered invasive by users. [20] also addressed the cold-start problem in web page recommender systems by examining various information that can be used to determine if web pages can be recommended. They used demographic attributes to find users with similar ratings. Even though they were able to generate recommendations for users with sparse ratings, they were not specific about the demographic attributes to be used.

Jafari et al. [21] analyzed some of the problems and challenges encountered in deploying recommender systems by mining web browsing patterns for generating most visited web pages by users so as to present them as recommendations. Mining web access records alone for recommendations generating cannot be used in modeling the individual needs of users of a recommender system, hence the proposed model will explore alternative information for recommendation generation.

Mabude [22] developed an expertise recommender system to address the problem of finding experts in academic collaborations. Even though the system was able to find experts registered on the system for possible research collaborations, it however implemented just the content-based and collaborative recommendation approaches. [23] addressed the challenges of building hybrid recommender systems that use linked open ratings data from different databases. A hybrid multi-strategy approach was used to combines the recommendation results of different base and generic recommenders using stacking regression and rank (ratings) aggregation for producing a final recommendation. Even though they were able to generated ratings popularity scores for recommendation generation, they however failed to explore the use of demographic recommenders.

Edson et al. [18] predicted user preferences in a group context by evaluation the model in their earlier work where a multi-faceted hybrid recommender system i.e. a recommender system that accepts and integrates different sets of data inputs. Even though they were able to analyze the impact of demographic data in predicting user needs in a group context, they did not predict the needs of users individually as will be considered in the proposed model.

As most online recommender systems use the collaborative-filtering or content-based approaches methods or a combination of both to provide recommendations using ratings data. Both approaches have their pitfalls. [24], thus, proposed a standard architectural framework, the Semantic Enhanced Personalizer (SEP) to integrate three recommendation techniques: the original, semantic and category-based techniques to enhance the content-based and collaborative recommendation processes in document recommendation. Their framework fulfils the user-based and item-based approaches of recommendations, and also overcomes the cold-start and sparsity problems.

The SEP framework, as shown in Fig. 1, has three recommendation modules, the original, semantic, and category-based recommendation modules. Original recommendations are generated based on the content-based and collaborative recommendation approaches, and also on contextual information and document ratings of provided by users of the system. Semantic recommendations are performed using various data mining techniques such as clustering, association-rule-mining and similarity measures by storing all visited document URLs in a database. If the conditions for implicit feedback are satisfied then satisfied, performance based transposition algorithm (PBTA) is applied to the database to find the most frequented URLs so that they can be recommended.

In category-based recommendations, frequent keywords are also extracted from visited document URLs using PBTA. Strong association rules are then formed based on these keywords. The URLs containing these keywords are then recommended. The SEP framework fulfils the user-based and item-based recommendation approaches by avoiding the need for ratings information in recommendation generation.

It however presents the challenge of not being able to provide personalized recommendations to users of the system as mining documents URLs visited by users collectively is not sufficient to model the individual needs of the users of the system. Each user has needs and preferences that the system must cater for, and the SEP framework cannot attend to this.

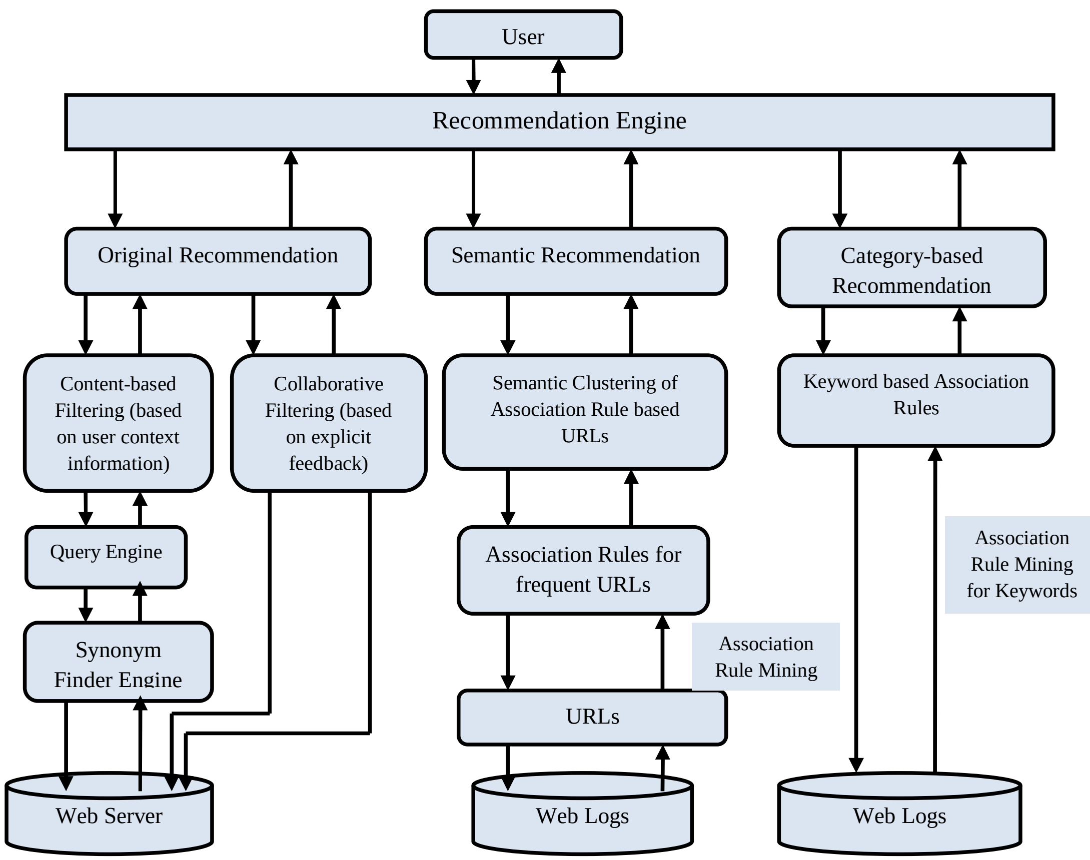

Fig. 1. Semantic Enhanced Personalizer (SEP) architecture (Sharma and Suman, [24])

In an online store scenario as considered in the proposed model, shoppers on online stores are more likely going to view a lot of pages and items to locate the items they wish to buy, compare prices, or make the decision to purchase. If these URLs are being mined to find the most visited pages so items on them can be recommended, the system is likely going to suffer from serious scalability problems when generating personalized recommendations. The visited URLs can be used in finding top most rated items, best seller items and retailers etc in generating non-personalized recommendations.

Ways of reducing scalability problems such as reducing customer data on the system and alternative information for personalized recommendation generation other than mining web pages will be considered in the proposed model so as to adequately model user needs.

There are different strategies for hybridizing filtering techniques. They have been broadly classified into seven types. These are: weighted hybridization, switching hybridization, mixed hybridization, feature combination hybridization, cascading hybridization, feature augmentation hybridization and meta-level hybridization [8,11,6]. [8,25,11,6,26] further discussed the strategies as follows.

In a weighted hybrid recommender system, the scores/ratings for a recommended item are computed from the results of all the filtering techniques used by the system. In switching hybridization, the recommender system uses some criteria to switch between filtering techniques. Mixed hybridization incorporates two or more filtering techniques at the same time, for example, content-based, collaborative filtering and demographic filtering techniques. The system is thus able to make large number of recommendations simultaneously. It is thus advisable to use "mixed" hybridization in situations where recommendations from more than one filtering techniques are to be presented together at the same time. Mixed hybridization also helps avoid the 'new-user' or 'cold-start' problem. This is because recommendations from other filters can be presented by the system to its users if data for one filtering process is not available at a time.

Feature combination hybridization is often used in mergers of content-based and collaborative filtering techniques. In cascading hybridization, one filtering technique refines the results given by another. In feature augmentation hybridization, the output of one filtering technique is used as an input feature to another. Finally, meta-level hybridization combines two or more filtering techniques by using the model generated by one as the input for another. They differ from feature augmentation hybrids in that the entire model of the first filter is used as input to the second whereas in an augmentation hybrid, a learned model is used to generate features for input to a second algorithm. The benefit of meta-level hybridization, especially for the content and collaborative hybrid recommender systems is that the learned model is always a compressed representation of user interests.

Developing online shopping models to adequately cater for the information needs of Nigerian shoppers and provide them with relevant recommendations require exploring ways of providing personalized item recommendations to Nigerian online shoppers without using their ratings information. This is because ratings information is mostly unavailable or relatively sparse on Nigerian stores as Nigerian shoppers are not enlightened about the importance of carrying out activities like rating items while shopping online. This is why Nigerian online retailers expend large costs on the promotion of their stores and the products they sell through media platforms. Despite doing this, these retailers have not been able to successfully retain their customers, attract newer customers to these stores, or even guarantee the continued online presence of these shoppers on their stores. Hence, this paper proposes the utilization of the most available data on online stores which is the shoppers' purchase data for providing them collaborative recommendations using the vector similarity and adjusted cosine similarity measure. Furthermore, the utilization of contents specified by these shoppers through searches and their records on the stores is proposed in this paper for the provision of content-based recommendations through the binary search, user profiling and association rule mining methods. For the demographic recommendations, this paper proposes obtaining demographic information from shoppers and utilizing it in the provision of recommendations to both visitors to the stores and existing shoppers on the store. However, it is imperative to note that if the demographic information required from the shoppers are those that could be considered invasive, Nigerian shoppers are not likely to provide them due to fear of their information being compromised.

3. PROPOSED HYBRID ITEM RECOMMENDER ARCHITECTURE

The proposed hybrid architecture is shown in Fig. 2. In the architecture, the model is expected to provide each category of recommendation (content-based, collaborative, and demographic recommendations) as data is made available for its generation. For example, when a shopper registers and provides demographic information on the store, the shopper will be grouped into the cluster that best fits the shopper's demographics. These clusters will be predefined. Similarities will then be found for the shopper and other shoppers in the cluster for recommendation provision to the shopper. Furthermore, the specific items assigned to the cluster for recommendations will be recommended to the shopper.

Consequently, when the shopper purchases one or more items on the store, the shopper will get recommendations of items also purchased by other shoppers who have one or more purchased items in common with the shopper. The vector similarity method assumes that two shoppers are two vectors in an *n*-dimensional space of items on the store. One of the shoppers is the active shopper (shopper for whom recommendations are to be made) and the shopper for whom the active shopper is being compared with. If the two vectors point are in the same direction in the *n*-dimensional space, they have similar purchases. But if they are pointing in different directions in the space, they are dissimilar.

The cosine measure has the disadvantage of not finding how each of the shoppers' purchases deviates from the average purchases of the two shoppers. This disadvantage was addressed by the adjusted cosine measure, otherwise referred to as the Pearson's Correlation Coefficient (PCC).

The PCC measured how much purchases by common shoppers for a pair of items deviated from average purchases for those items. The condition for similarity is expected to be predetermined on the store e.g. 'if the shoppers have three (3) purchased items in common'. It is imperative to note that as the number of shoppers on the store increases, it is imperative

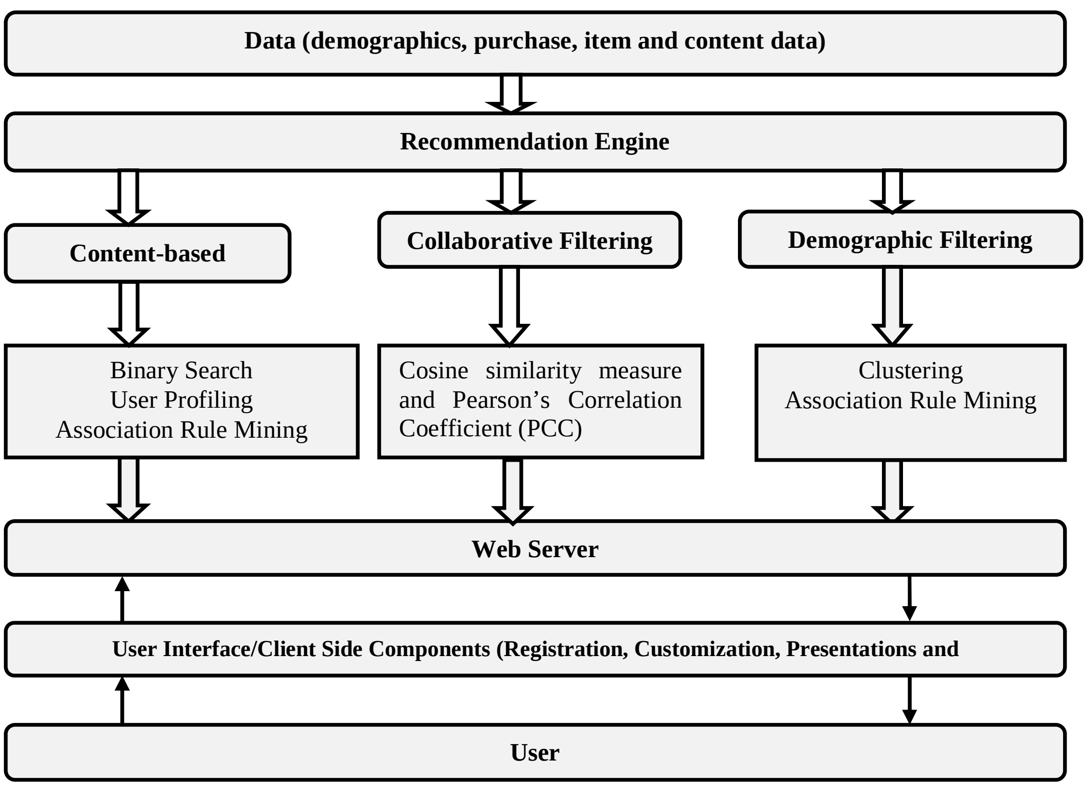

Fig. 2. Proposed hybrid item recommender architecture

for the conditions for similarity to be redefined so as to reduce the number of shoppers the active shopper will be compared with.

When a shopper provides purchase data or specifies contents (through search queries issued) on the store, the store creates a profile for the shoppers and stores the content in it. It then proceeds to find items often purchased with or associated with the items purchased by the shopper that the shopper has not yet considered, and recommends them to the shopper. Also, it recommends other items in the categories of purchase of the shopper to the shopper randomly. For the binary search, keywords from the search content specified will be extracted, which will become the input element. The entire sorted list of items is then divided into two (2), the input element (item/items being sought for) is then compared with the mid element. The system then tries to find if the sought item(s) is from the left or the right side of the list based on a conquer and divide strategy. If the item(s) sought for is found, it is returned as recommendations to the shopper. All these recommendations are to be presented together in mixed hybridization on the on the homepage of the online store.

4. CONCLUSION

The proposed approach is expected to enhance the quality of personalized item recommendations provided by Nigerian online stores and overcome the cold-start (new-user), new-item and sparsity problems. It is also expected to reduce scalability problems as only the purchase and demographic data of shoppers will be stored on the online store. It is expected to explore the use of alternative data sources such as purchase and demographic data for recommendations. It is also expected to introduce demographic recommendations on Nigerian online stores and provide the conceptual information needed to develop online stores adaptable to the peculiarities of the Nigerian shoppers.

The proposed architecture will provide a platform for providing quality personalized item recommendations on Nigerian online stores, minimize the various challenges encountered by shoppers while carrying out shopping activities online, enhance the content-based and collaborative filtering processes of Nigerian online stores, and reduce the errors of the processes.

COMPETING INTERESTS

Authors have declared that no competing interests exist.

REFERENCES

1. Zwass Z. Electronic commerce: Structures and issues. International Journal of Electronic Commerce. 1996;1(1):3-23.
2. Kgosana M, Thunstrom A, Miles K, Hunt S, Jackson D, Begg M. Fast-moving consumer goods in Africa. KPMG Africa Limited Article Publication. 2014;1(2):1-17.
3. Kabore SC. Design and implementation of a recommender system as a module for liferay portal. In Barcelona School of Computing (FIB), University Polytechnic of Catalunya (UPC). 2012;1-127.
4. Zenebe A, Ozok A, Norcio AF. Personalized recommender systems in E-Commerce and M-Commerce: A comparative study. Researchgate Publications. 2005;1(2):1-10.
5. Chu W, Park ST. Personalized recommendation on dynamic content using predictive bilinear models. In Proceedings of the 18th International Conference on World Wide Web, Madrid, Spain. 2009; 691-700.
6. Prassad R, Kumari VV. A categorical review of recommender systems. International Journal of Distributed and Parallel Systems (IJDPS). 2012;3(5):73-83.
7. Pazzani MJ, Billsus D. Content-based recommendation systems. The Adaptive Web. 2007;14(1):325-334.
8. Burke R. Hybrid recommender systems: Survey and experiments. User Modeling and User-Adapted Interaction Journal. 2002;12(1):331-370.
9. Escriche M, Symeon P. User profiling and personalization tools. WeKnowIt', Emerging, Collective Intelligence for Personal, Organizational, and Social Use Paper. 2011;1(2):1-26.
10. Kiong LT. Intelligent web techniques for web personalization. Intelligent Web Systems Paper, Massachusetts Institute of Technology, Cambridge, Massachusetts, USA. 2008;1-7.
11. Spiegel S. A hybrid approach to recommender systems based on matrix factorization. A Technical Report Technische Universistat Berlin. 2009;1-88.

12. Junior EBS, Manzato MG, Goularte R. Evaluating the impact of demographic data on a hybrid recommender model. IADIS International Journal on WWW/Internet. 2014;12(2):149-167.
13. Deshpande M, Karypis G. Item-based Top-N recommendation algorithms. ACM Transactions on Information Systems (TOIS). 2004;22(1):143-177.
14. Vozalis M, Margaritis KG. Collaborative filtering enhanced by demographic correlations. In Proceedings of the AIAI Symposium on Profesional Practice in AI, part of the 18th World Computer Progress, Kuala Lumpur, Malaysia. 2004;1-10.
15. Chen T, He L. Collaborative filtering based on demographic attribute vector. In International Conference on Future Computer and Communication (FCC'09). 2009;225-229.
16. Chikhaoui B, Chiazzaro M, Wang S. An improved hybrid recommender system by combining predictions. In Advanced Information Networking and Applications (WAINA) 2011 IEEE Workshops. 2011; 644-649.
17. Bobadilla J, Ortega F, Hernando A. A collaborative filtering similarity measure based on singularities. Journal of Information Processing and Management. 2012;48(2):204-217.
18. Edson B, Santos J, Marecelo GM, Rudinei. Evaluating the impact of demographic data on a hybrid recommender model. IADIS International Journal on WWW/Internet. 2015;12(2):149-167.
19. Vozalis M, Margaritis KG. Enhancing collaborative filtering with demographic data: The case of item-based filtering. In Proceedings of the Fourth IEEE International Conference on Intelligent Systems Designs and Application, Brazil. 2004;1-6.
20. Safoury L, Salah A. Exploiting user demographic attributes for solving cold-start problem in recommender system. Lecture Notes on Software Engineering. 2013;1(3):303-307.
21. Jafari M, Sabzchi FS, Irani AJ. Applying web usage mining techniques to design effective web recommendation systems: A case study. Advances in Computer Science: An International Journal. 2014; 3(2):78-90.
22. Mabude CN. Development of an improved model for expertise recommendation in academic research. Unpublished M.Sc Thesis submitted to the Department of Computer Science and Engineering, Obafemi Awolowo University, Ile-Ife. 2014; 1-129.
23. Ristoski P, Mencía EL, Paulheim H. A hybrid multi-strategy recommender system using linked open data. Semantic Web Evaluation Challenge. 2014;1(2):150-156.
24. Sharma SK, Suman U. Design and implementation of architectural framework of recommender system for e-commerce. International Journal of Computer Science and Information Technology & Security (IJCSITS). 2011;1(2):53-162.
25. De-Gemmis M, Laquinta L, Lops P, Narduci F, Semeraro G. Preference learning in recommender systems. Preference Learning Journal. 2009;41(1): 1-15.
26. Jannach D, Friedrich G. Tutorial: Recommender systems. In Proceedings of the International Joint Conference on Artificial Intelligence, Beijing, China. 2013; 1-144.

Peer-review history:
The peer review history for this paper can be accessed here:
http://sciencedomain.org/review-history/11954