

# Personalized and Location aware Online Product Recommendation System

Nusna Khalid C K<sup>1</sup>, Sabeeha K P<sup>2</sup>

<sup>1</sup>MEA Engineering College, State Highway 39, Nellikunnu-Vengoor, Perinthalmanna, Malappuram, Kerala

<sup>2</sup>Assistant Professor, MEA Engineering College, State Highway 39, Nellikunnu-Vengoor, Perinthalmanna, Malappuram, Kerala

**Abstract:** *The large volume of information available on the web, the broad coverage of the web content, the phenomenal number of web users and their continued rapid growth has presented a major challenge to the web community. Identifying and accessing relevant information suited to individual needs is called personalized information retrieval and access. Based on user's requirements and preferences, the personalized recommendation systems are in need to provide proper recommendations. This paper proposes a product recommender system for intuitively supporting mobile users in decision making with respect to the location of both the user and product. Other recommendation techniques included in the project are: Personal interest based recommendation and Friend Circle based recommendation.*

**Keywords:** Personal Interest, Interpersonal influence, Location aware recommendation

## 1. Introduction

Recommender Systems analyze some user and item interactions to help users by recommending the most relevant, feasible and appropriate items from a wide range and pool of items and resources. New factors such as interpersonal relationships and interest in the circle of friends in social networking, provide opportunities and challenges for the recommendation system (RS) to solve the cold start and sparse problems of the data set.

Recommender systems, in the most general sense, are used to give the user of a site/function (whether that be a streaming video site like Netflix and YouTube, an online shop like Amazon or an online auction site like eBay) some added value in the form of recommendations. Some of the recommenders use only behavioural data, e.g. ratings, purchases, clicks and so on, to predict what items a customer would like. When doing predictions with little or no data about the users or items, these kind of recommenders usually can't give good, if any, recommendations. This is known as the cold-start problem. There are different kinds of systems and not everyone agrees on what classification should be used, but the most common classification divides all systems into one of three categories: Collaborative filters (CF), Content-based filters (CBF) or Hybrid filters. In 2012, a new classification is suggested that is based on the utility rather than the kind of information used by the system. The utilities are the following

- Help to Decide
- Help to Explore
- Help to Compare
- Help to Discover

Recommender systems can also be categorized as giving non-personalized or personalized recommendations. Recommender systems providing personalized recommendations looks at a user's history to help make recommendations, while the type giving non-personalized recommendations does not take that into account. Amazon is one e-commerce site that provides personal product

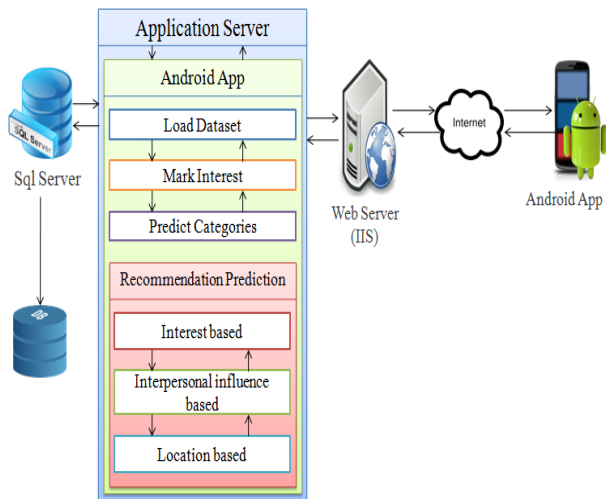
recommendations. According to [4], they use item-item collaborative filtering in their recommendation system, i.e., they use behavioural data to make recommendations in the form of for example "people who bought this also bought that". [5] mentioned that Amazon uses hybrid recommendation system, i.e. a system that uses both content data and behavioural data.

Recommender system compares a user profile to some reference characteristics and seeks to predict the 'rating' or 'preference' that a user would give to an item they had not considered. These features may be from the information item (the content-based approach) or the user's social environment (the collaborative filtering). Many of the recommender system using the technique is collaborative filtering (CF), analyze past community views to find correlations of similar users and items to suggest k personalized items to a querying user u. Many e-commerce firms have adopted the portal as the vehicle for their electronic commerce/business transactions. This is an integrative, user-centric and user-intensive approach, which offers various advantages and benefits to all stakeholders involved.

Location-aware recommender system that uses location-based ratings to produce recommendations. The traditional recommendation system is not considered spatial properties of users nor items. Experimental analysis using real and synthetic data sets i.e., Yelp data set, show that the proposed system is efficient, scalable, and provide better quality advice than techniques used in traditional recommender systems.

## 2. The Approach

The overall architecture of the location aware online product recommendation system is shown in fig :



**Figure 1:** Architecture diagram

The SQL Server is a relational database management system (RDBMS) developed by Microsoft. This product is built for the basic function of storing retrieving data as required by other applications. It can be run either on the same computer or on another across a network. An application server is a component-based product that resides in the middle-tier of a server centric architecture. It provides middleware services for state maintenance and security, as well as data access and persistence. Java application servers are based on the Java 2 Platform, Enterprise Edition (J2EE). Here we have two application server, Web application and Android application.

A Web application is a client-server software application, in which the client (or user interface) runs in a Web browser. Common web applications include web mail, online retail sales, online auctions, wikis, instant messaging services and many other functions. An Android app is a software application running on the Android platform. Because the Android platform is built for mobile devices, a typical Android app is designed for a smart phone or a tablet PC running on the Android OS. The android application displays Personal interest based product recommendation, Interpersonal influence based product recommendation and Location based product recommendation.

The proposed system will solves the cold start problem of recommendation system. The cold start refers to the new user into the recommendation system with less historical behavior. If a new user to the system, the proposed system will recommend product based on their friend's search history.

A Web server is an HTTP program that uses HTTP (Hypertext Transfer Protocol) to provide users with a web page file based on the user's request (forwarded by the computer's HTTP client). Dedicated computers and appliances may be referred to as Web servers as well. Cascading Style Sheets (CSS) is a simple mechanism for adding styles (such as fonts, colors, spacing) to a Web document. These pages contain information about how to learn and use CSS and available software. Although it is most commonly used to set the visual style of web and user interfaces written in HTML and XHTML, the language can be applied to any XML document, including pure XML,

SVG, and XUL, and is suitable for presenting media in voice or other ways. With HTML and JavaScript, CSS is the cornerstone technology used by most sites for creating visually appealing web pages, user interfaces for Web applications, and many mobile application user interfaces.

The Internet is a global system that uses the Internet Protocol Suite (TCP / IP) in a global network of interconnected computer networks. It is a network of networks linked by a wide range of electronic, wireless and optical networking technologies, consisting of private, public, academic, business and government local to global networks. The Internet has a wide range of information resources and services such as interconnected hypertext documents and file sharing applications for the World Wide Web (WWW), e-mail, telephone and point-to-point networks.

A Web browser, commonly known as a browser, is a software application for retrieving, presenting, and traversing information resources on the World Wide Web. The information resource is identified by a Uniform Resource Identifier (URI / URL) that can be a web page, image, video, or other content. Hyperlinks present in resources enable users easily to navigate their browsers to related resources. Although the browser is primarily intended for use with the World Wide Web, they can also be used to access information provided in a Web server or file system in a private network.

The Android application is a software application running on the Android platform. Because the Android platform is built for mobile devices, the typical Android app is designed for smart phones or tablets running on Android.

### 3. Recommendation Techniques

The proposed system consists of mainly 3 segments based on the factors considering for recommendation, which are 1) Interest based recommendation, 2) Interpersonal influence based recommendation and 3) Location based recommendation.

#### 1) Interest based recommendation

In personal interest based product recommendation, recommend products based on their interest category. These interest category are mined using decision tree approach. The steps involved in mining user interest are follows:

##### a) Attend questionnaire

The system will provide some questions, which include user age, job description, salary, gender etc. After attending the questionnaire, the system will store it to the corresponding database.

##### b) Load query dataset:

The query dataset will load in order to make predictions. After that the system will predict interest category using decision tree approach.

**c) Plot decision tree:**

After loading the query dataset the system will plot corresponding decision tree. Which makes the prediction of user interest based on the query data.

**d) Predict category**

After plotting decision tree, we get predictions based on user interest. Then the system will discover the category that the user interested in. After mining the user interest category, the system will recommend product / services according to the user interest.

**e) Decision Tree**

Decision Tree Classifier is a simple and widely used classification technique. It applies to the intuitive idea of solving the classification problem. Decision tree classifiers have a series of well-designed questions about test record attributes. Each time it receives an answer, a follow-up question is asked until a conclusion about the class label of the record is reached. Constructing the optimal decision tree is the key problem in the decision tree classifier. Usually, decision trees can be constructed from a given set of attributes. A stop condition is also needed to terminate the tree-growing process. One possible strategy is to continue to expand a node until all records belong to the same class, or all records have the same attribute value. Prediction is one of the most important usages of decision tree models. Using the tree model derived from historical data, it's easy to predict the result for future records.

**2) Interpersonal influence based recommendation.**

In Interpersonal influence based recommendation, consider user's relationship to the other user's. Here recommend product which are interested by the user's friends with higher rating. The proposed system will solves the cold start problem of recommendation system. The cold start refers to the new user into the recommendation system with less historical behavior. If a new user to the system, the proposed system will recommend product based on their friend's interest.

**3) Location based recommendation.**

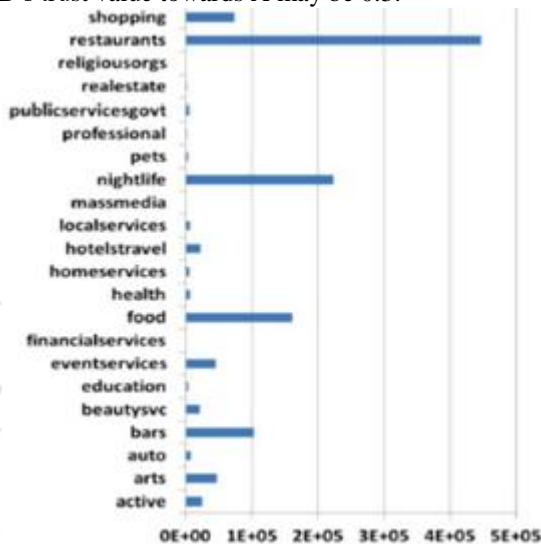
In location based product recommendation consider the geo location of both user and item. Which mainly helpful for travelers. For example, When a user is travelling to a new place, he wants to purchase the product which are popular in that specific region, then the proposed system will recommend the products by considering the recommendation of persons in that region.

**4. Dataset**

**Yelp Dataset**

Yelp is a local directory service with social networking and user reviews. It is one of the most popular consumer review websites and has more than 71 million monthly unique visitors as of January 2012. It combines local reviews and social networking functionality to create a local online community. Yelp allows real people to contribute their own reviews. This body of social reviews creates a participatory culture where anyone can share their insight and suggestions and add to a body of "collective intelligence" on local businesses, here call items, by using their personal

knowledge and skill sets to post, revise, and assign them numeric ratings in the range of 1 to 5. Essentially, this form of intelligence allows people to actively participate and share their knowledge with other users, especially their friends. Different from the Epinion datasets, the friendships between users are bidirectional: if user A is in user B's friend list, then user B is also in user A's friend list. However, the influence between users with friendship is unidirectional: if user A's trust value towards B is 0.5, then user B's trust value towards A may be 0.3.



**Figure 2:** Distribution of the ratings in the 22 categories of Yelp dataset. [10]

Here crawled nearly 60 thousand user's circles of friends and their rated items from November 2012 to January 2013. First collect some active users in New York as origin, 43 and then crawl these user's friends to build the sub-networks of Yelp. Except the user without any rated history (at least one rated item), the dataset consists of ratings from 10,555 users who rated a total of 1,783,922 items from 22 big categories (actually, Yelp has 26 big categories, the other four categories are with a few user and items). The average number of user ratings is about 169. The distribution of the ratings cross all categories is plotted in Fig. 2. The top three popular categories are Restaurants, Night Life, and Shopping. Merge the similar categories like Restaurants and Food, Night Life and Bars.

Category	User Count	Item Count	Rating Count	Sparsity	r'
Active Life	5327	7495	24395	6.110e-4	4.021
Beauty and Spas	5466	8495	21345	4.597e-4	3.937
Home Services	2500	3213	5180	6.449e-4	3.707
Hotels & Travel	4712	5883	21658	7.813e-4	3.824
Night Life	4000	21337	99878	1.170e-3	3.594
Pets	1624	1672	3093	1.139e-3	3.975
Restaurants	2000	32725	91946	1.405e-3	3.677
Shopping	3000	16154	33352	6.882e-4	3.819

**Figure 3:** Yelp Data: Statistic of the Test Categories [10]



To test the applicability of the proposed model, choose eight categories based on the popularity distribution of the 22 categories. There are three most popular categories, three common categories (Active Life, Beauty and Spas, Hotels and Travel) and two less rating but interesting categories (Home Services and Pets). Finally, the eight categories are selected as dataset. More detail of this dataset can be found from website of SMILES LAB7. Here tried to remove personal information manually before sharing this dataset. Fig. 3 is a statistic of users and items in the eight categories, where is user's average rating value in c.

## 5. Performance Measures

Recommendation accuracy acquired by, true positive rate, false positive rate, precision, recall and F-measures using .Net tool.

	Recommended	Not Recommended
Liked	TP	FN
Not Liked	FP	TN

**Figure 4:** Factors influencing performance of RS

### a) Precision

Precision proportion of the examples which truly have class x / Total classified as class x. Precision gives positive predictive values and it process values or product quality or exactness. So basically high precision stated the accurate results and it takes all relevant data but returns only topmost results.

$$\text{Precision} = \frac{TP}{TP + FP}$$

### b) Recall

Recall gives sensitivity of problem and it process values or product quantity or completeness. It returned most relevant and part of the documents that are relevant as result from the query. In other words, modules that are really recognize as difficult to maintain from the total number of modules.

$$\text{Recall} = \frac{TP}{TP + FN}$$

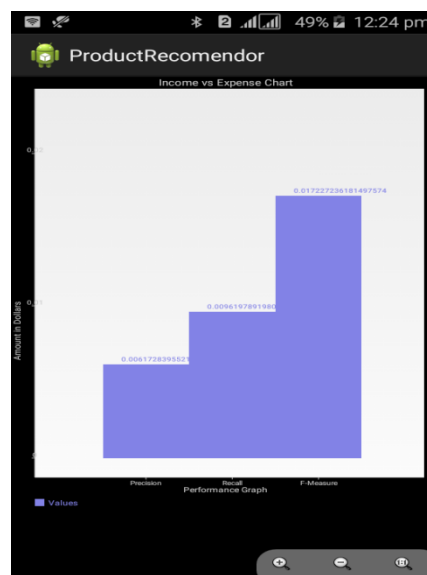
### c) F-Measure

It is a combined measure for precision and recall.

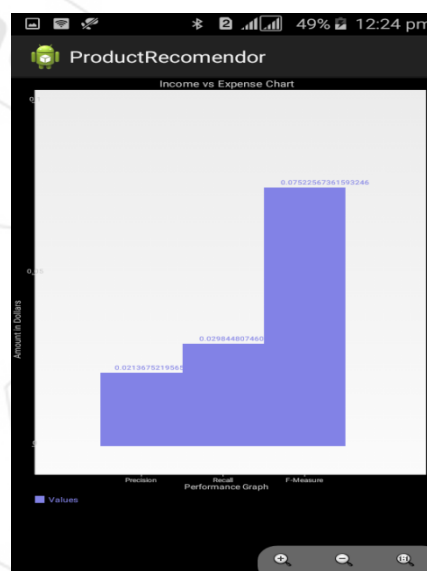
$$F\_Measure = 2 \frac{\text{Precision} \cdot \text{Recall}}{\text{Precision} + \text{Recall}}$$

### d) Comparison Graph

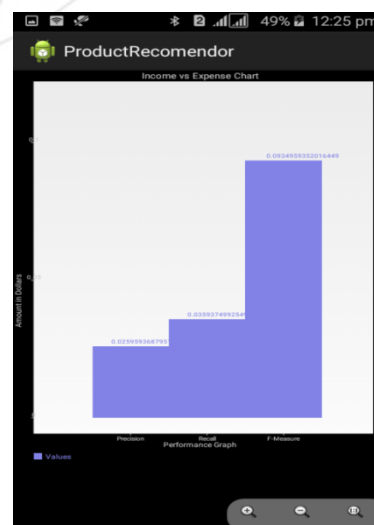
This proposed system is compared with interest based recommendation and interpersonal influence based recommendation techniques:-



**Figure 5:** Interest based recommendation: Precision, Recall, F-Measure graph



**Figure 6:** Interpersonal influence based recommendation : Precision, Recall, F-Measure graph



**Figure 7:** Location based recommendation: Precision, Recall, F-Measure graph

And the proposed system provides more accurate result.

## 6. Conclusion

The proposed system focus on development and evaluation of a recommender system within the product domain. Here propose a novel location-aware recommender system built specifically to produce high-quality location-based recommendations in an efficient manner. Evaluation will be done by assessing the system functionality and comparing the recommender precision. Our conclusion from the experiment can be summarized as follows:

- 1) Location aware recommendation increases recommender precision. Which provide faster query results without loss in the quality of recommendation.
- 2) User location has proved to be an important aspect. Our experiment has shown that by integrating a mechanism work for adding location into the recommender system, it has been possible to produce recommendations that better suits the user's preferences.

In the future, going to incorporate the temporal and weather features into the recommendation system.

## References

- [1] X. Yang, H. Steck, and Y. Liu, Circle-based recommendation in online social networks, in Proceedings of the 18th ACM SIGKDD international conference on Knowledge discovery and data mining. ACM, 2012, pp. 1267–R. Caves, Multinational Enterprise and Economic Analysis, Cambridge University Press, Cambridge, 1982. (book style)
- [2] M. Jiang, P. Cui, R. Liu, Q. Yang, F. Wang, W. Zhu, and S. Yang, Social contextual recommendation, in Proceedings of the 21st ACM international conference on Information and knowledge management . ACM, 2012, pp. 45–54.
- [3] X. Qian, H. Feng, G. Zhao, and T. Mei, Personalized recommendation combining user interest and social circle, IEEE transactions on knowledge and data engineering, vol. 26, no. 7, pp. 1763–1777, 2014.
- [4] Linden, G., Smith, B., and York, J.(2003). Amazon.com recommendations: Item-to-item collaborative filtering. IEEE Internet computing, 7(1):76–80
- [5] Rao, K. N. (2008). Application domain and functional classification of recommender systems—a survey. DESIDOC Journal of Library & Information Technology, 28(3):17
- [6] R. Bell, Y. Koren, and C. Volinsky. Modeling relationships at multiple scales to improve accuracy of large recommender systems. In Proc. of KDD '07, pages 95-104, San Jose, California, USA, 2007.
- [7] G. Adomavicius and A. Tuzhilin. Toward the next generation of recommender systems: A survey of the state-of-the-art and possible extensions. IEEE Transactions on Knowledge and Data Engineering, 17:734–49, 2005.
- [8] Y. Zhang, B. Cao, and D. Y. Yeung. Multi-domain collaborative filtering. In Proceedings of the 26th Conference on Uncertainty in Artificial Intelligence (UAI), Catalina Island, California, USA, 2010.

- [9] Qian, X., Feng, H., Zhao, G., and Mei, T.(2014). Personalized recommendation combining user interest and social circle. IEEE transactions on knowledge and data engineering 26(7):1763–1777
- [10] M. Jamali and M. Ester. Trustwalker: a random walk model for combining trust-based and item-based recommendation. In Proceedings of the 15th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD), pages 397–406,
- [11] L. Liu, J. Tang, J. Han, M. Jiang, and S. Yang. Mining topic-level influence in heterogeneous networks. In Proceedings of CIKM, pages 199–208, 2010.
- [12] H. Ma, H. Yang, M. R. Lyu, and I. King, “SoRec: Social recommendation using probabilistic matrix factorization,” in Proc. 17th ACM CIKM, Napa Valley, CA, USA, 2008, pp. 931–940.
- [13] B. Sarwar, G. Karypis, J. Konstan, and J. Reidl, “Item-based collaborative filtering recommendation algorithms,” in Proc. 10th Int. Conf. WWW, Hong Kong, China, 2001, pp. 285–295.
- [14] J. Huang, X. Cheng, J. Guo, H. Shen, and K. Yang, “Social recommendation with interpersonal influence,” in Proc. 19th ECAI, Amsterdam, Netherlands, pp. 601–606, 2010.
- [15] H. Ma, H. Yang, M. R. Lyu, and I. King, “SoRec: Social recommendation using probabilistic matrix factorization,” in Proc. 17th ACM CIKM, Napa Valley, CA, USA, 2008, pp. 931–940.

## Author Profile



**Nusna Khalid C K** received the B.Tech degree in Computer Science and Engineering from Nirmala College of Engineering, Chalakudy in 2015 and doing M.Tech in Computer Science and Engineering from MEA Engineering College, Perinthalmanna.



**Sabeeha Mohammed Basheer K P** is a post graduate in Computer Science and Engineering from MES College of Engineering, Kuttippuram. She graduated from BSA Crescent Engineering college, Chennai. She is currently working as Assistant Professor in MEA Engineering College, Kerala. Her research interest lies in Data Mining and Information Systems.