PERSONALIZED RECOMMENDATION SYSTEM FOR LEISURE TIME ACTIVITY USING SOCIAL MEDIA DATA

M Ramashini[1], DP Jayathunga, A Gowthamy, N Rammiya, and U Kiruthiga

Department of Computer Science and Technology, Uva Wellassa University, Badulla, Sri Lanka
[1]ramashini@uwu.ac.lk

Abstract- In today's digital world wherever there's associate degree endless variety of content to be consumed like books, videos, game, movies, music, etc., finding the content of one's interest has become associate degree deadening task. A personalized recommendation system for leisure activities is vital in our social life due to its strength in providing enhanced entertainment. Our system has the ability to recommend leisure time activities, to a new user and others by using their social media data. It gathers all the important information, such as popularity, liking and disliking, required for recommendation. It also takes a minimum of new user information without connecting to social networks. It generates recommendations for the user based on his/her behaviour on social media. Such a system will counsel a group of films, books, music, TV shows, games and places to users supported their interest and private data using Collaborative filtering and Content-based filtering. Similarity, index is measured by using Pearson correlation and Cosine based similarity and Tanimoto Coefficient based Similarity. The planned system has the flexibility to advocate leisure activity to a brand new user furthermore because the others by mistreatment social media knowledge. It effectively reduces the complexity of the search space for users and attracts more and more users to the Internet, which increases the profits of site owners.

Keywords- Recommender system, Cold start user, Big data, Personalized, Hadoop.

I. INTRODUCTION

Recommender Systems ar new generation net tool that gives the power to know a person's style and notice new, fascinating content from the massive variety data} accessible on the online or in alternative electronic information supply supported the pattern between their likes and rating of various things. Similarly, a time off activity recommendation system provides A level of comfort and personalization that helps the user act higher with the system and acquire fascinating suggestions to pay his time off in usefully. This paper presents an outline of the Recommender Systems that ar presently operating within the domain of time off activities like business enterprise, movie, book, music, games, and television programs. we have a tendency to conceive to develop net application and mobile application. With the appearance of massive information, it's become tough to method the huge quantity of knowledge for recommendation (Ankit Kumar Das, 2014).Due to this reason, Hadoop is employed to use for quantifiability, responsibility and quicker process.

The term massive information is outlined by four dimensions pictured by four V's (Volume, Variety, Velocity, and Veracity). Volume is pictured by the number of text information that we have a tendency to ar mistreatment for account to come up with recommendation. selection represents completely different completely different} form of information extracted from different sources like blogs, Facebook, Twitter in addition as totally different review and opinion sites. Speed represents the seed of knowledge generation on the net. these days everyone is connected through the net in addition increasing the recognition of ecommerce sites has become the most reason for increasing the speed of text information generation on the online. Truthfulness represents the trait of the information. Persistently review, opinion, feedbacks ar manipulated or sponsored by totally different stakeholders

of on-line business for his or her personal interest (Holms, 2012) (Sean Owen, 2012) (Jai Prakash Verma, 2015).

Hadoop provides its own classification system known as HDFS (Hadoop Distributed File System). after we deploy our text information on Hadoop classification system, Hadoop distributes all the information in numerous clusters and performs operations parallel. Hadoop conjointly keeps multiple copies of knowledge just in case of hardware failure (Holms, 2012) (Sean Owen, 2012) (Jai Prakash Verma, 2015).

II. PROJECT DEFINITION

Recommendation Systems are primarily directed towards people World Health Organization are having lack of ample personal expertise or ability to judge the possibly overwhelming range of other things that an online web site, for instance, may offer. In most cases, folks are two-faced with decisions and really giant information volumes, and looking out all of them is out of user' capability. This downside is named data overload (A. Felfering, 2007). In the busy world everyone gets leisure time rarely. They wish to spend the valuable free time in useful manner but unfortunately they waste the time and resources by searching better choice and traditional recommendation system suggests items based on rating and reviews. But user may not satisfy with that suggestions.

The recommendation Systems are used in different fields such as tourism, movie, book, music, game, TV shows etc.

Movie Lens is an online movie recommendation system developed by Group lens Research Group, which asks user to rate a few movies on the first login. Through these ratings, the movies are recommended to the user. Collaborative recommendations are provided supported the ratings of comparable users. Grouplens conjointly provides an intensive dataset for movies and ratings in numerous sizes (F.Maxwell harpist, 2016).

MOVREC (Manoj Kumar, 2015) could be a motion picture recommendation system bestowed by D.K. Yadav et al. supported cooperative filtering approach. Cooperative filtering makes use of data provided by user. That data is analysed and a motion picture is usually recommended to the users that square measure organized with the motion picture with highest rating initial. The system additionally

encompasses a provision for user to pick attributes on that he needs the motion picture to be suggested.

A hybrid system has been presented by Harpreet Kaur et al. (Harpreet Kaur Virk, 2015). The system uses a mixture of content yet as cooperative filtering formula. The context of the films is additionally thought-about whereas recommending. The user -user relationship yet as user - item relationship plays a significant role within the recommendation.

The Internet is a vast source of information. Whenever any person dreams of planning a tour for himself first thing he is looking for the suggestion from the web and he might faces problems like extensive information and has no idea which information is reliable and which is not. The introduction of digitization in the field of tourism has led to the collection of massive tourism related data. This big data is complicated to process using conventional tools. The data and information are an integral part of any field as it helps in understanding recent trend and patterns. Hence, there is need of a system which will overcome this Problem (Yetis, 2016).

To provide an efficient recommendation to the user, it is of prime importance to understand the user. What are his likings or preferences and much more Social media plays a significant role in understanding user. Social media is a great platform to understand user preferences using which recommendations can be provided. This will ultimately lead to user satisfaction (Miah, 2016) introduces a system which extracts user data from web sources like Flickr and provides recommendation.

There exist variety of book recommendation systems (Greg Linden, 2003) (G. Ramya, n.d.), among that the one utilized by Amazon could be a in style one. Amazon's recommender, as conferred in (Greg Linden, 2003), suggests things to a user that area unit kind of like different users' past purchased and/or rated things, i.e., things that seem within the purchase patterns of assorted users.

(G. Ramya, n.d.)Rely on a ranking-oriented cooperative filtering approach, that considers users' preferences on digital library resources extracted from users' access logs to perform the advice task. This filtering approach overcomes the matter that arises because of the poorness of express users' ratings, i.e., lack of initial info to perform the advice task, in predicting digital library materials of interest to a user.

Park and Yangtze Kiang produce a user-profile P supported individual and cluster behaviour info, like clicks, searching habits, purchases, and interest fields, for generating book recommendations. Using P, the authors calculate the geometrician distance between P and every product profile, that describes product options, and choose merchandise that their geometrician distances ar the nighest to P. extra references on book recommendation systems will be found in Digital Library (Hui Li, 2009). The authors of Ido Guy et al (Ido Guy, 2010) (Maria city Pera, n.d.). Use social-media knowledge to boost the performance of advice systems.

To overcome the issues found in above mentioned existing systems we found out that social network data is extremely helpful for cold start users, users who have not yet interacted with the items. We propose a customized recommendation try and apprehend the characteristics and preferences of the user by collection and analyzing historical behaviour to grasp what reasonably person the user is, what reasonably behaviour preference the user has, what reasonably things the user prefer to share and then on (Mohammad Yahya H. AI-Shamri, 2008) (Xu, 2014) (Jinming, 2010) and eventually perceive that user characteristics and preferences supported the foundations of the platform and suggest the knowledge and product that the user interested (Yan, 2011) (Davidsson C, 2011) it's Associate in Nursing integrated system that may be a combination of a spread of leisure activity recommendations like business, movie, book, music, game and television shows.

In our system a flick recommendation provides A level of comfort and personalization that helps the user move higher with the system and watch movies that cater to his desires. The chief purpose of our system is to advocate movies to its users supported their liked on Facebook and ratings that they supply. In movie recommendation, we wish to use a Hybrid recommender system as it avoids the cold start and sparsity problems inherent with the other models of recommender systems. A hybrid recommender system is one that mixes multiple techniques along to attain some synergism between them.

Book recommendation systems will profit industrial websites, social media sites, and digital libraries, to call a couple of, by assuaging the information acquisition method of users UN agency hunt for books that area unit appealing to them. Our recommendation system that's

supported social interactions and private interests to recommend books appealing to users. System depends on the information established on a social networking web site. User primarily {based} cooperative and item based cooperative approaches area unit accustomed suggest the books.

In our system a tourism recommendation is to suggest the destination and places of the user's visit. The Facebook Graph API is used to obtain user information. Facebook is one of the largest social media sites. People are very active on Facebook. This system uses the photos uploaded by the user, the marked photos, and the personal data of the user. Tagged photos and uploaded photos provide details such as the creation time, name, interesting place, the location of the attraction. In this system, the Google Places API is used to get a list of places to visit. After the destination has been selected, the landmark of the destination can be received.

Instead of searching different leisure time activities from many separated system we provide a facility to search different leisure time activities from one system. As a result of this our users can save their time and resources. And users easily get desirable choice among vast amount of suggestions from personalized recommendations in our system. As we said early our system user can spend their leisure time in their preferable way instead of wasting time in searching.

III. OBJECTIVES

A. Main Objectives

To introduce web application and mobile application to recommend the leisure time activities such as tourism, movie, books, music, games and TV shows.

B. Sub Objectives

I.   To process information and provide the user with potentially more relevant items by estimating their preferences.

II. To alleviate the matter of data overload, that has created a possible drawback to several web users.

III. To handle the cold start problem effectively

IV. PROJECT DEFINITION

This System facilitate the registered users to login the system through their account and update their user profile. If anyone not register to in system they can register or login through Facebook. At the time system gather and store the information of the user from Facebook and user registration. As soon as user login to the system, leisure time activity categories such as tourism movie book game music and TV show are displayed. User can choose any of them according to their wish.

In tourism recommendation, System captures the current location of the users. System provides two type recommendation.

I. According to the preference.

It recommends most similar users visited places based on the user's preference. Users also can get the detailed information about the places (Distance, Climate details, Description, nearest hospitals, restaurant, police station etc.)

System will tour plan and lists of POI. A purpose of interest or dish could be a feature on a map (or in a very geo dataset) that occupies a selected point, as opposition linear options like roads or areas of land use. The term dish is truly quite inexact, however is widely known by users of satellite navigation systems (SATNAVs), WHO are usually bestowed with choices to indicate or hide points of interest. It is also for geocaching and GIS users, however "POI" takes on totally different meanings in several GIS systems.

II. According to the shortest path.

It recommends the places where old users visited based on shortest distance.

In movie recommendation, for cold start users system will recommend the movies according to the user's preference which are given user during the registration. For registered users system will recommend high rated movie list from similar user preference. Users not only view the movie details but also give the rates to the movie. User can also view the high rate movie list.

In book recommendation system, cold start user get recommendation according to the population estimations and their preference. Ordinary user can view recommendation according to the user

based collaborative filter. If user selects a one of the book from the recommendation, he/she can view the available details of the book such as title, author, publisher, rating etc. And also gets book recommendation based on the similarity of the target book.

V. METHODOLOGY

A. Content-based Filtering Systems (CBF based systems):
In content-based filtering, items are recommended based on comparisons between item profile and user profile.

B. Collaborative filtering based systems (CF based systems):

Collaborative filtering system recommends items based on similarity measures between users and/or items.

I. User based collaborative filter Recommendations are given to user based on evaluation of items by other users form the same group, with whom he/she shares common preferences (Asanov, n.d.) (Badrul Sarwar, n.d.).

II. Item based collaborative filter

Referring to the fact that the taste of users remains constant or change very slightly similar items build neighbourhoods based on appreciations of users. Afterwards the system generates recommendations with items in the neighbourhoods that a user would prefer. (Asanov, n.d.) (Khyati Aggrawal, 2016)

C. Data Extraction

We collect user details, user likes, movie data and book details from social media data.

D. The cold start problem

The cold start problem appears in two variants. The first variant is related to new users. Whenever a user logs into the system for the first time, the system does not have information about the preferences of that user to provide the recommendations. The second variant is related to

new items. When a new item is added to the system, it doesn't have the user ratings. Due to lack of ratings, the item may be placed lower in the recommended items list. To calculate the similarities for neighbourhood, the historic ratings are necessary, and they are unavailable in the cold-start scenario. This makes it difficult to predict ratings for items. [26]

E. Tourism recommendation description

For any person, selecting a destination is a challenging task. The amount of information available on the Internet is enormous which leaves that person confused. In this project, the system takes care of this problem by suggesting an appropriate destination to the individual. It analyses the person's profile and provides personalized recommendations.

F. According to the preference

STEP 1(Profile Analysis): In the first step, new user profile and old user's profiles are compared. The old user list is iterated, and each old user's places visited are compared to new user's visited places. If the place is present in both old and new user's list, this place is added to the common place list. If the place in old user's list is not found in new user's list, the place is added to the different place list.

STEP 2 : After creating a list of common and different places, next step is filter the old user list. Once final old user list is ready, it is sorted according to the maximum number of common places.

STEP 3 : Now, the old user with maximum common places are taken into consideration. The list of different places is taken, and those places are suggested to the new user. STEP 4: Once destination place is finalized, the system will start looking into old user's places. Then it will pick places. To fetch point of interest, the system will use Google places API.

II. According to the shortest path

Owing to the various demerits of pure content-based and pure CF based systems, we have proposed a hybrid recommender system. This hybrid system takes advantage from both the representation of the content as well as the
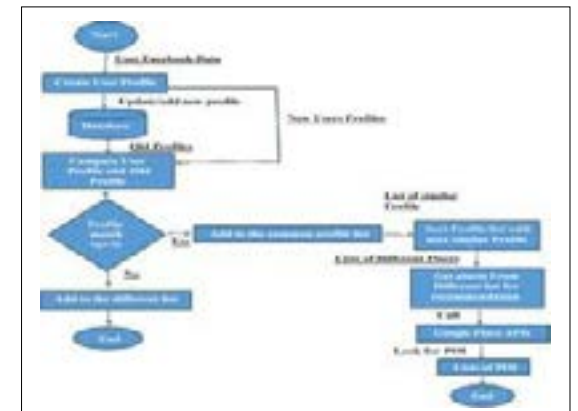

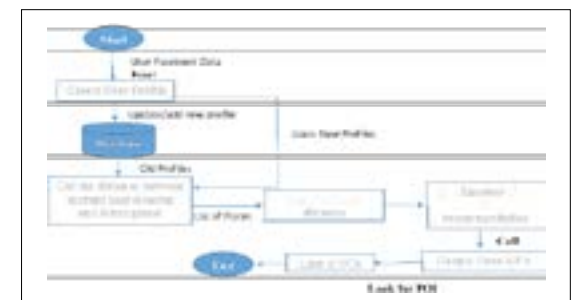*Figure 1. Tourism recommendation overview according to the preference*

Figure 2. Tourism recommendation overview according to the shortest path

similarities among users. Use a content-based filter to fill the user rating based profile matrix. Get necessary movie content for our dataset. After the preprocessing the movie content database is stored. The dataset consists of a user-rating matrix. Pseudo rating matrix which combines user rating based profile matrix with user-rating matrix. Collaborative filtering is then applied to this full pseudo user-rating matrix to make recommendation for an active user.

Step1: Extract data from Facebook and get minimum information of user's preference about movie. Store the details

Step2: If user is cold start user then create a user rating based profile matrix using content based filtering.

Step3: Weight all users with respect to similarity with the active user. Similarity between users is measured as the Pearson correlation between their ratings vectors. Where, m - total number of items.

$$P_{a,u} = \frac{\sum_{i=1}^{m}(r_{a,i} - \bar{r}_a) \times (r_{u,i} - \bar{r}_u)}{\sqrt{\sum_{i=1}^{m}(r_{a,i} - \bar{r}_a)^2} \times \sqrt{\sum_{i=1}^{m}(r_{u,i} - \bar{r}_u)^2}}$$

Step4: Select n users that have the highest similarity with the active user.

These users form the neighbourhood.

Step5: Compute a prediction rate of each movie for current user from a weighted combination of the selected neighbour's ratings.

Predictions are computed as the weighted averages of deviations from the neighbour's mean:

$$P_{a,i} = \bar{r}_a + \frac{\sum_{i=1}^{m}(r_{u,i} - \bar{r}_u) \times P_{a,u}}{\sum_{i=1}^{m}P_{a,u}}$$

Where,

n - Number of users in the neighbourhood.

Step6: Sort the movie list according to the rate prediction. Step7: Display the top rate movie; there is an option user can see the top rate movie.

Step7: If user is old user then load his/her pseudo rating matrix; go to Step: 3

$$sim(i,j) = \frac{\sum_{u \in U}(R_{u,i} - \bar{R}_i)(R_{u,j} - \bar{R}_j)}{\sqrt{\sum_{u \in U}(R_{u,i} - \bar{R}_i)^2}\sqrt{\sum_{u \in U}(R_{u,j} - \bar{R}_j)^2}}.$$

Step8: In Step7 user click that movie, Display the similar movie using content based filtering based on rate. Step9: Finish.

H. Book recommendation description

Book recommendation methodology is created newly by combination of two existing methodology (Ankit Kumar Das, 2014) (Ishwari Kulkarni, 2017) and some features are added. Book recommendation relies on Item-based collaborative filter approach and user-based collaborative filter approach. Cosine based Similarity, Pearson Correlation-based Similarity and Tanimoto Coefficient based Similarity methods are used to find the average similarity.

Figure 3. Movie recommendation overview

i. Popularity Estimation

The third type of recommendation algorithm is based on popularity. In general, items that most people are interested are usually a safe choice when you don't know her particular preferences. In the case of new user, popularity becomes one possible solution to the cold start problem. Popularity of the book can be defined as the number of users that have rated that book (Ishwari Kulkarni, 2017).

ii. User type Classification

We classify the user as new user and ordinary user. If user is new user then popularity estimation algorithm is applied as collaborative filtering is not suitable. If user is ordinary user then both popularity estimation and collaborative is applied. If user has not rated any book then it is considered as new user otherwise it is ordinary user (Ishwari Kulkarni, 2017).

iii. Pearson Correlation-based Similarity

The Pearson correlation of two series is the ratio of their covariance to the product of their variances. (Ankit Kumar Das, 2014).

In this case, similarity between two items I and j is measured by computing the Pearson-correlation I, j. Let the set of users who both rated I and j are denoted by U then the correlation similarity is given by

iv. Cosine based Similarity

In this case, two items are thought of as two vectors in the m dimensional user-space. The similarity

between them is measured by computing the cosine of the angle between these two vectors. Formally, in the m×n ratings matrix, similarity between items I and j, denoted sin (i, j) is given by where − • ∥ denotes the dot-product of the two vectors (Ankit Kumar Das, 2014).

$$sim(i,j) = \cos(\vec{i}, \vec{j}) = \frac{\vec{i} \cdot \vec{j}}{\|\vec{i}\|_2 * \|\vec{j}\|_2}$$

v. Tanimoto Coefficient based Similarity

Tanimoto Coefficient Similarity is an implementation which does not take into account the preference values specified for the users. It is based on (surprise) the Tanimoto coefficient (Ankit Kumar Das, 2014).

Similarity = 2 • similarity – 1

Figure 4. Book recommendation overview

Step1: Login user classify based on user type classification method. If the target user is a new user follow Step: 02. If user is an ordinary one follow step: 03.

Step2: Books are recommended to the new user according to the popularity estimation and features that are given by user.

Step3: For ordinary user, Apply the User-Based Collaborative filter approach and find the

similarity among users. Select the users who have similarity above 75%. Collect all books which are preferred by the selected users. From These Books, Separate the Books that are not match with target preferred Books.

Step4: Separated Books are sorted by rating. Books are recommended by this sorted list. If user select one book from them then Display the available details of the books and go to step: 05 else go to step: 06

Step5: Apply the Content-based –Filter approach then calculate the Pearson Correlation-based Similarity, Cosine based Similarity, and Tanimoto Coefficient based Similarity individually. And find the average similarity. If average similarity is above 75% then that books are recommended and go to Step: 06 Else go to Step: 06

Step6: End the process.

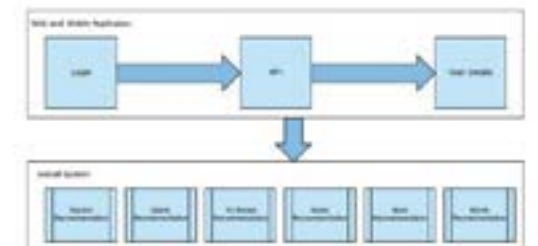## I. Methodology for overall Recommendation System

Figure 5. Methodology for overall recommendation system

VI. CONCLUSION

This paper explains how quality recommendations can be given to the user by taking minimum input from users. The system consists of Collaborative Filtering, Content Based Recommendations and Hybrid Recommendations which provide recommendations to the user. Hybrid Recommendations is one of the main modules of the system which helps overcome the drawbacks of the traditional Collaborative and Content Based Recommendations. To study the user preference we have to use social media data. Our system is a best choice for users who prefer different leisure time activities.

ACKNOWLEDGEMENT

REFERENCES

A. Felfering, G. F. a. S. T., 2007. Recommender systems. In: s.l.:IEEE Intelligent systems, pp. 18-21.

Ankit Kumar Das, P. K. T., 2014. *Book Recommender System Using Hadoop*, New Delhi: Department of Computer Science and Engineering,Amity school of engineering and technology.

Asanov, D., n.d. *Algorithms and Methods in Recommender Systems*, German: Berlin Institute of TechnologyBerlin.

Badrul Sarwar, G. K. J. K. a. J. R., n.d. *Item-Based Collaborative Filtering Recommendation Algorithms*, Minneapolis: GroupLens Research Group/Army HPC Research Center.

Davidsson C, M. S., 2011. Utilizing implicit feedback and context to recommend mobile applications from first use.

F.Maxwell Harper, J. A., 2016. The Movie Lens Datasets. 5(4).

G. Ramya, K. R., n.d. Personalized QOS based Ranking Approach for Cloud Service Selection. *International Journal of Computer Applications,* October, Volume 127, pp. 0975-8887.

Greg Linden, B. S. a. J. Y., 2003. *Amazon.com Recommendations*, s.l.: IEEE Computer Society.

Harpreet Kaur Virk, E. M. S., 2015. Analysis and Design of Hybrid Online Movie. April .5(2).

Holms, A., 2012. Hadoop in Practice, s.l.: Manning Publications co.

Hui Li, Y. G. S. K., 2009. *A Review of Digital Library Book Recommendation Models*. [Online] Available at: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=1513415 [Accessed 15 December 2009].

Ido Guy, N. Z. I. R. D. C. E. U., 2010. *Social Media Recommendation based on People and Tags*, Haifa: IBM Research Lab.

Ishwari Kulkarni, P. G. K., 2017. Book Recommendation System Using Apache Spark. *Industrial Journal of Innovative Research in Computer and Communication Engineering*, April , 5(4), pp. 2320-9801.

Jai Prakash Verma, B. P. P., 2015. *Recommendation System with Hadoop Framework*, s.l.: ResearchGate.

Jinming, H., 2010. *Application and research of collaborative filtering in e-commerce recommendation system*, s.l.: s.n.

Khyati Aggrawal, Y. S. G. J., 2016. Movie Recommandation Using Hybrid Recommendation. *International Jourrnal on Recent and Innovation Trends in Computing and Communication*, 4(12), pp. 206-209.

Manoj Kumar, D. A. S. V. K. G., 2015. A Movie Recommender System. August.Volume 124.

Maria Soledad Pera, Y.-K. N., n.d. *Generating Personalized Book Recommendations Using Data Extracted from a Social Website*, Utah: Brigham Young University.

Miah, S. J. V. H. Q. G. J. a. M. M., 2016. *A big data analytics method for tourist behaviour analysis*, s.l.: Information Management.

Sean Owen, R. A. D. F., 2012. *Mahout in Action*, s.l.: Manning Publications co.

Xu, Q. W. J. C. Q., 2014. *A novel mobile personalized recommended method based on money flow model for stock exchange*, s.l.: s.n.

Yan, B. G. C., 2011. *AppJoy :personalized mobile application discovery*. s.l., 9th international conference on Mobile systems, applications, and services.