RESEARCH ARTICLE OPEN ACCESS

A Friend Recommender System for Social Networks by Life Style Extraction Using Probabilistic Method -"Friendtome"

Namrata M.Eklaspur [1], Anand S.Pashupatimath [2]
M.Tech P.G Student [1], Professor [2]
Department of CSE
SDMCET, Dharwad
India-580 002

ABSTRACT

Every day we are overwhelmed with many choices and options, simultaneously recommendation systems have gained popularity in providing suggestions. Today every web application has its own recommendation system. Whereas, Recommendation systems for social networks are different from other kinds of system, since the item here are rational human beings rather than goods. Hence, the 'Social' factor has to be accounted for when making a recommendation. We considered one of the most popular social Networking sites that is Facebook as it offers impressive features. Here, we are mainly focusing on recommending friend with similar interest which is different among all the existing ones where Facebook uses social graph a friend of friend approach to recommend friend which may not be the most appropriate to reflect a user's preferences on friend selection in real life. And Netflix, Foursquare which all focus on recommending items. Hence we proposed framework Friendtome, a novel semantic based friend recommendation system for social networks.

In this paper, a social network is formally represented and taking text mining as a perspective, we have proposed a framework that will recommend friend using an efficient Algorithm. Here, we have analyzed the structure of Facebook and considering the activities of individuals got some values & computed the score of each individual based on which we have, analyzed and computed to show the percentage of similarity of life styles between users, and recommends friends to users if their life styles have high similarity.

Keywords:- Friend recommendation, life style, social networks.

I. INTRODUCTION

Social networks can be considered as a milestone in the web history with the advance in online social life. A social network is "a set of people (or organizations or other social entities) connected by a set of social relationships, such as friendship, co-working or information exchange."

Fig.1. Social Networks

Social Networking sites like FaceBook (FB) focus on building and reflecting the social networking and relationships among the community sharing similar interests. FaceBook showed tremendous changes in the way how people communicated and connected to one another. And mainly the friends are recommended based on the previous existing relationships and pick among them as friends for example, FaceBook makes use of social link analysis among those who already share common friends and recommends symmetrical users as common friends and connect people across the country. Which might be not that suitable to recommend as it doesn't reflect any user preferences on friend selection in real life.

Social networking sites have enormous data set of users, according to the current survey. Every individual social networking site makes record of the activities of users such as his/her likes; what user likes? , what user is doing? , what is user's hobby? Etc. and it has gained main area of focus in understanding the user behavior, One of the best example we might consider is FaceBook.

Hence here, in our approach we are making use of user life style as major concern for recommending friends and build relationship among the people with similar interest and help to share information or build communication among likely minded people. With Graph API provided by FaceBook developers we extract user interest which tends to discover life style of user. The contexts may also include information like topic of interests, hobbies, profession, etc. the information about the user interest and his/her profession can be used to recommend friends.

Contextual information has been recognized by researchers and practitioners in many disciplines including Ecommerce, personalized IR, ubiquitous and mobile computing, data mining, marketing and management. There are many existing e-commerce websites which have implemented recommendation systems successfully. Few of the recommending system have been described in the next part.

Inspired by text mining, we consider user profile as life document, where we treat user activities (eg:watching) as words in life document , and topics (ex: movie) as life style and document is treated as life document containing daily lives of user. example :if user A has interest in movies and user B has interest in sports ,and user C interest is movies then the user A is interested in making friend with user C then compared to user B .as user A and user C share similar life style so the user C is recommended as friend to user A. The current context has a major dimension on personal context (example: topics of interest, hobbies, profession, etc.) With information about the activity in which the user is currently involved (e.g. doing sport, working, etc.) Can be extracted with help of Graph API on FaceBook.

Thus, we model a user's daily life as life document, from which his/her life styles are extracted using the Latent Dirichlet Allocation model. We further propose a similarity metric to measure the similarity of life styles between users, and user's impact in terms of life style with a friend-matching graph. Upon receiving request, Friendtome returns a list of people with highest recommendation scores to the query user. Finally, Friendtome integrates a feedback mechanism to further improve the recommendation accuracy.

II. EXISTING SYSTEM

Online social networks have become important hubs of social activity and conduits of information. Popular social net- working sites such as FaceBook, the social news aggregator Digg, and the microblogging service Twitter have undergone explosive growth. Though FaceBook has approach of recommending friend which is based on mutual friend that makes use of friend of friend approach found to be not that appropriate.

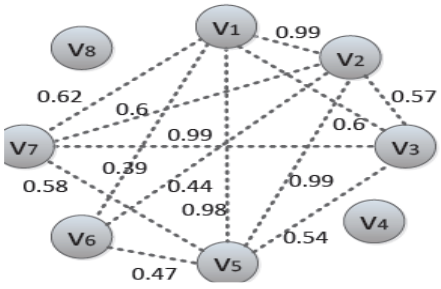


Figure 2. Friend of matching graph

Thus, this motivated to build the framework of recommending friend with similar interest. With the numbers of active users on these sites numbering in the millions or even tens of millions, identifying a people with similar interest among them becomes an important problem with application in marketing, information dissemination, search, and expertise discovery. Recommender Systems are software tools and techniques providing suggestions for items to be of use to a user [5,6].

In the existing system, many recommending systems have their own proposed framework for assigning ranks to the user activities and having various personalized recommendation. Such as Netflix for movie recommendation, Foursquare for recommending places, FaceBook for recommending friend based on mutual friends. In which recommending friend based on mutual friends is not that appropriate, these are the various disadvantages that motivated us to propose new system.

In this paper we considered FaceBook for extracting the user details such as name, interest, email id etc. and we have analyzed its structure. From our study perspective one of the important functions of this network is user interest. User interest is the process by which thoughts and actions of individual are generated and depicted in their profile and can analyze on it to identify his/her life style. This can be widely accepted in social networks.

Hence, the paper aims at fulfilling the development of the following system:

Considering, FaceBook profile data, we calculate probabilities of the topics in the user document using LDA model that is considering the probabilistic method to find dominant life style vector and then recommending to the query user with potential friend whose values are greter then certain specified threshold value.

III. METHODOLGY

A. Problem definition

In this paper we propose a method for:

"Finding out the interest of people and recommending him/her friend based on interest in social networking site: Facebook as a case i.e., the task of identifying users capable of stimulating other users to join activities/discussions in their OSN is an important issue in the present era."

Our framework consists of two main processes:

- Development of a web application
- Development & Implementation of a LDA model to find the user with similar interest and recommend friend to user.

Let's take a brief look on what we do in the above stated two processes:

1. **Development of web application**: This step involves creation of web application on the client side using HTML, Php. For a web application, we will be providing a URL for the user to search the web page. Later a login page is created through which user logs into our application i.e. Friendtome app. When a user logs into client side web app, it will be automatically directed to Facebook log in. The Facebook SDK consists of Graph API tool that provides information about the user activities if user has given permissions to access the basic information. Each time the user logs into our application, an access token is generated by Facebook for that particular user. All the user activity and access token is displayed on our application result page. Then whatever activity he has done from the base login time to the current log-in all the activities is being stored in database of the developed application and that of the further part is used for the implementation section using the data stored in database. The activities we mainly concentrating is his interest that is movies, books read, sports, etc.
2. **Development & Implementation of a Probabilistic model to find the user with similar interest and recommend friend to user**: During the web development phase, the user data is recorded into our database. The user activity from the database is accessed. An algorithm for calculating dominating life style vector of user is developed. LDA algorithm is a way of automatically discovering topics that the sentences in document has, it finds the topic by calculating the probability of words in document. Similarly in case of FaceBook we apply this method and find the dominant life style vector as below, . The life style of users is extracted by the life style analysis module with the probabilistic topic model, and then the life style indexing module puts the life styles of users into the database in the format of (life-style, user) The probabilistic topic model can be given as,

$$P(W_i|d_k)=\sum^{Z}_{j=1} p(w_i|z_j)\ p(z_j|d_k)$$

Where, w-activity

Z-life style

D –set of document and in our case as we are implementing it in facebook, d_k can be considered as 1 as we are able to fetch the topics directly by considering user activity as whole document.

The topics may be movies, books read, sports etc. and the count of these activities can be accessed based on the permission given by FaceBook developers and the people who logs in to our app and allow us to access the data to recommend them friend of similar interest among the people from our database. As we get the count values we calculate probability of each activity of user using above formula then we find dominating life style vector of user by specifying some assumed threshold value in our case we have considered it as 0.2.let us define this threshold as α (alpha) And after finding dominating life style vector of user we find similarity between the users this is done using the below formula,

$$S=S_c(i,j).S_d(i,j).$$

where i & j are number of users

S_c=is cosine similarity and

S_d= is distance similarity.

Hence, a cosine similarity can be calculated as below

Between user 1 & user 2,

$$S_c(U1,U2)=COS(U1,U2)=a.b/|a||b|$$

Similarly with all the users it is calculated. And distance

Similarity is calculated as below,

$$S_d(U1,U2)=2|D1\cap D2|/|D1|+|D2|.$$

After calculating similarity value for all the user with every

Other user we store those values in matrix form from which

We recommend a friend to the user who is greater then

Some specified threshold value, we have assumed

Threshold value as 0.5 in our case and let's consider this

Recommending threshold as β(beta).

In the proposed work, we have focused on four important phases as below (Fig.2):

a. *Creating a user interface application for login*: Web applications that require authorization to access certain information. Your login page verifies a user's name and

password, places a cookie on the user's computer so he can return later, and uses database queries to retrieve the personal information for the user.

b. *Extracting user data and storing in database*: We use Graph API tools for extracting data. The advantages of Graph API over previous work are the ability to learn highly accurate extraction rules, and then we store this user information like 'name', 'email', likes, in the database that we have created.

c. *Finding dominant life style*: Depending on the activities that user has done we get certain count of the activity, then we calculate probabilities of each life style and consider those values who are greater then some specified threshold value α (alpha). In which the user interacts with the site through our application.

d. *Recommending potential friend*: We calculate the similarity between the users and recommend friends to the query user who are above certain threshold value β (beta).

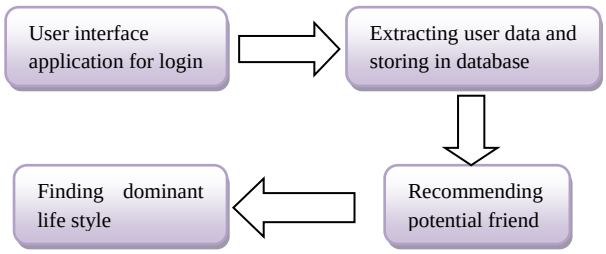


Fig.2. General architectural model

IV. PROPOSED WORK

Dependencies of our framework

- User must be logged-in into application connected to the Facebook.
- All the user activities are tracked and dumped into database along with user access permissions.
- Based on the information collected in the database potential friend is being recommended to the query user.

Main steps to find people with similar interest and recommend querying user,

The main steps are as follows:

- Develop a web application which connects to Facebook Login page through which we can host the app and users can give permissions.
- To retrieve the user data through the access tokens specified for each user.
- Based on accessing permission given by user for a web application we can get the activities performed by user.
- Develop and test a methodology to find the users with similar interest in online social networks on the basis of a simple metric of their activity level.
- We calculate the probabilistic values of each activity and find dominating life style,
- And recommend potential friend to the query user.

V. PROPOSED MODEL

In the Fig.3. we can see how the app is registerd for first time and how access token is provided for each user who uses the app. So at the time of user login the unique access tokens are provided to the app developer through browser and then the user data is continously posted from facebook to the application.

Step wise procedure of Fig.3:

STEP 1: The first step is to login into the webpage that is directed by our application.

STEP 2: Then send request to the server to access the information of user.

STEP 3: If user permits then access token is generated for each user for authentication purpose. User data is returned.

STEP 4: Here, it will post a one time-use token to token URL page.

STEP 5: The user data can be accessed using the access token given to the user.

STEP 6: User data can be served up to registration form or auto login.

STEP 7: User submits page, information is stored in database and flow is complete.

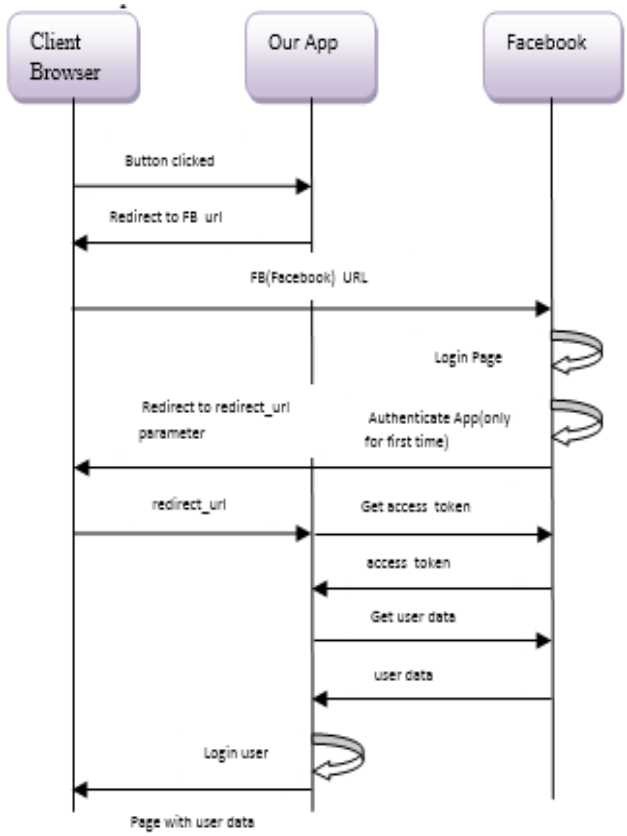


Fig.3. Creation of Facebook app & its working

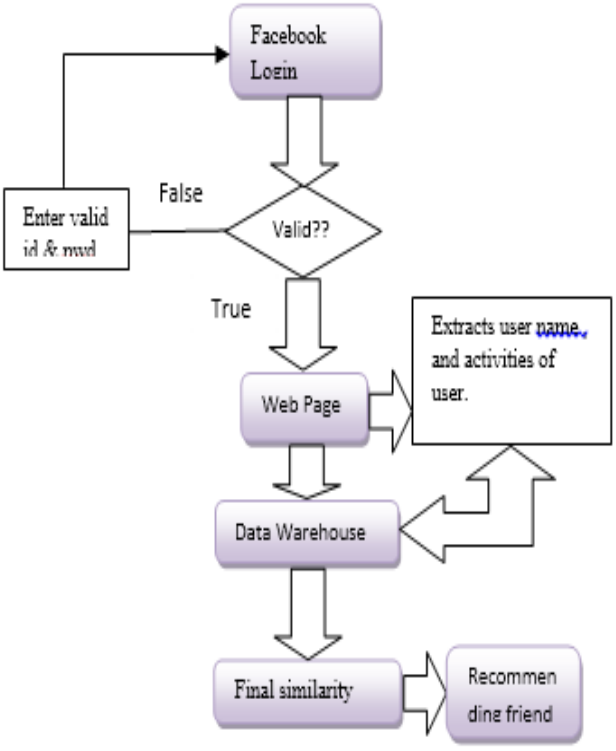


Fig.4.Architectural model of proposed system

Fig.4. shows the architectural diagram for the proposed problem; here we can see the flow of various actions that has been shown in the block diagram. The first step is log-in to the web page checks for the correct id and password and does verification. Once it is correct it is moved on to the webpage and it collects all the information such as name, email, and the activities performed by the user such as movies watched, sports liked, etc. And it is stored into the database then this is used for recommending friends.

A. Database Design

There are three main parts in database design:

1. First we have to store the user email_id , movies liked, books liked and give unique serial id to users as shown in Table.1.
2. Secondly, we collect all user permissions and group them as different attributes.
3. Next when we collect the entire user permissions those should also be stored in database as integer values because we are considering the count of user activities.

Table.1 shows the sample users taken for conducting the experiment. It shows the data which has be taken from database, it includes the user email_ID, and the activities done (Count) when logged into their FB account.

Table 1: DataBase Design

FB-user name	Sekhar Jay Krish	Sanjay Bedare	Shruthy Yeshodaran	Namrata Eklaspur	Soumya Bhat	Vani Shriengeri	Netra Pattan	Shweta Bhat
Login id	10000020013344 5	100000545136960	1000022649840 54	100000179234 927	100001493160 701	100001007660248	100001007660245	100001007660257
Movies	17	15	16	18	15	15	10	15
Books	7	20	9	4	5	2	9	9
sports	3	25	2	0	4	5	8	8

VI. IMPLEMENTATION

A. *Experimental Setup*

We are implementing using php, and HTML and running it on a Pentium – IV with 1GB of RAM and 200 GB Hard disk. The Operating system used is Windows XP. The server side script is written in php and database creator and connector used is MySQL 5.0.

Our system consists of modules as below:

1. **Creation of new user:** Steps to create a user:

a) Open your (a user who has his Fb account) Facebook account).

b) Simultaneously, Open a Graph API Explorer (From google.com), which displays the token access, user id & name of the Facebook user.

c) Click on Get Access Token

d) You will be displayed with a Select permission dialog box. Select the permissions as per user interests.

e) Next jump to the next module to work with our Login Page.

2. Login Module

In this module we create an interface to the main login page of the FaceBook site through our web application login. Here we define the form for login page of our application and also define the user cookies and sessions to keep the track of the users.

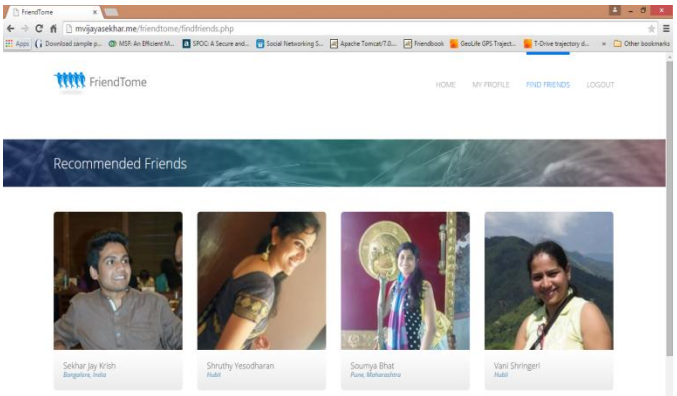
This module acts as the main means to collect both the access token of each user and also for loading various modules of Facebook through Facebook SDK for each user asynchronously.

Here when a user gives his user id & password (of Facebook). As a response to the request Facebook graph API tool will validate user details (check whether the password suits the username or not).

3. Result Module

After successful implementation, testing and deployment of the project the project's working in the user environment is recorded as the screen captures which gives the clear interpretation of results. The below screen shot gives the result of friends being recommended with similar interest the list of users is been shown.

This is demo result of recommendation. To carry out for all users we need permissions provided by the FaceBook Developers as well as the access permissions given by the application user. The objective of the system is to provide the connection among the people with likely minded .The system is tested on various end users. The system achieved an efficiency of delivering a potential friend to end user details is quiet satisfactory.



VII. CONCLUSION AND FUTURE WORK

In our approach we presented the design and implementation of Friendtome, a semantic-based friend recommendation system for social networks. Different from the friend recommendation mechanisms relying on

social graphs in existing social networking services, the results showed that the recommendations accurately reflect the preferences of users in choosing friends. Beyond the current prototype, the future work can be concentrated on implementing it on other social networking, and same can be used to build stand alone app and access the user activity through mobile sensors. Friendtome can utilize more information for life discovery, which should improve the recommendation experience in the future.

ACKNOWLEDGMENT

First and foremost, I would like to express my sincere gratitude to my project guide Prof. Anand S.Pashupathimath who has in the literal sense, guided and supervised me. I am indebted with a deep sense of gratitude for the constant inspiration and valuable guidance throughout the work. Lastly, I extend my heartfelt thanks to my friends Vijaysekhar, Swati and my parents, family for their in time kind support and helping me in doing project.

REFERENCES

[1] Prem Melville and Vikas Sindhwani, Recommendation Systems, In Encyclopedia of Machine Learning, Claude Sammut and Geoffrey Webb (Eds), Springer, 2010 Chapter No: 00338, Pg 829-838

[2] W. H. Hsu, A. King, M. Paradesi, T. Pydimarri, and T. Weninger. Collaborative and structural recommendation of friends using weblog-based social network analysis. Proc. Of AAAI Spring Symposium Series, 2006.

[3] Zhibo Wang, Jilong Liao, Qing Cao, Hairong Qi, and Zhi Wang, **"**Friendbook: A Semantic-based Friend Recommendation System for Social Networks**"**IEEE TRANSACTIONS ON MOBILE COMPUTING, VOL. 13, NO. 99, MAY2014

[4] J. Kwon and S. Kim. Friend recommendation method using physical and social context. International Journal of Computer Science and Network Security, 10(11):116-120, 2010.

[5] D. M. Blei, A. Y. Ng, and M. I. Jordan. Latent Dirichlet Allocation. Journal of Machine Learning Research, 3:993- 1022, 2003.

[6] L. Bian and H. Holtzman. Online friend recommendation through personality matching and collaborative filtering. Proc. of UBICOMM, pages 230-235, 2011.

[7] K. Farrahi and D. Gatica-Perez. Probabilistic mining of sociogeographic routines from mobile phone data. Selected Topics in Signal Processing, IEEE Journal of, 4(4):746-755, 2010.

[8] K. Farrahi and D. Gatica-Perez. Discovering Routines from Largescale Human Locations using Probabilistic Topic Models. ACM Transactions on Intelligent Systems and Technology (TIST), 2(1), 2011.

[9] L. Gou, F. You, J. Guo, L. Wu, and X. L. Zhang. Sfviz: Interestbased friends exploration and recommendation in social networks. Proc. of VINCI, page 15, 2011.