A SURVEY ON SIMILARITY BASED FRIEND RECOMMENDATION SYSTEM USING TEXT ANALYZES

#R. Nivedhitha[1]
Research Scholar,
Department of Computer Science,
Avinashilingam University,
Coimbatore, India
malarchan03@gmail.com

#C. Malarvizhi[2]
Research Scholar,
Department of Computer Science,
Avinashilingam University,
Coimbatore, India
nivedhirengaraj@gmail.com

Abstract: Social Network is an emerging E-service for content sharing sites (CSS). Sharing the content across the world they were many kind of opinion, to group the same opinion friend recommendation is most helpful one. Social network help to recommend the things to the others with the help of the link and object suggestion. They were many things were posted in the social media, the people of the same opinion where separated in various place, this paper deals with grouping the similarity opinion people, to know the highest opinion among the peoples with the help of the friend recommendation system. Today every web application has its own recommendation system. This survey attempts to provide similarity of thinks with the help of different text mining techniques in the social networking websites.

Keyword: *Social Network, content sharing, opinion analyses, similarity measure and recommendation system.*

I. INTRODUCTION

Social Networking (SN) is one of the improving technological with hundreds of millions of people participating to swapping their content through Text, media like image, audio, video, etc. Social media (SM) become one of the most important parts of our daily life as it allows us to communicate with a group of people. . It assists an exterior of self expression for users, and assists them to entertain and exchange content with other users through social media's providing E-Service. Some of the Social media are Friendster.com, Tagged.com, Xanga.com, LiveJournal, MySpace, Facebook, Twitter and LinkedIn have developed on the Internet over the past several years. In social media connecting with friends from various countries is possible, people of same country where goes outer the country because of their personal things like job, education, etc, while going from one to another place they miss their relatives, friends, neighbor, etc. for communicating those people social media helps a lot. In olden days people use social media for exchanging the communications, images, talks etc. Social networking helps people stay in touch that might not do it otherwise. It can be used to help advertise goods and services and can provide an extremely accessible

662

medium for self expression to those with access to computer. It can help families torn apart by war, divorce, etc. stay in touch easier and quicker than by some other means and a powerful engine for job searches.

Figure 1: Social Network

With the help of the recommendation system they may easily find the friend of friends. Recommendation systems that try to suggest items (e.g., music, movie, and books) to users have become more and more popular in recent years. For instance, Flipkart [1] recommends items to a user based on items the user previously visited, and items that other users are looking at. Recently, with the advance of social networking systems, friend recommendation has received a lot of attention. Most of the previous technique where focus on recommendation of friends, but this paper deals with the similarity of the opinion of the users, for eg; if the user post any post means it deals with the opinion of the other users and their similarity of the user, if same opinion was posted by the number of friends means we suggest the friends to be one circle, likewise it recommend the friends. Recommendation systems used for providing quality of customized user experiences. This help to the users to identify their interests and set the choices by predicting the usefulness degree of an item or group of items to these users. However, this is used to find the similarly

of two users, the focuses on community detection. An analysis on brain networks using multi objective functions. Recommendation systems can be separated into two areas of center: link recommendation and object suggestion. Facebook and Twitter focus on link recommendation where friend recommendations are presented to users. Flipkart and Amazon focus on the object suggestion for further items. This paper deals with the similarity recommendation to the users, same opinion people were suggested for the friends. And it helps for the same opinion were grouped in one cluster and opposite opinion where grouped into another cluster with the help of the text mining. In Social networking websites, such as Facebook are rich in texts that enable user to create various text contents in the form of comments, wall posts, social media, and blogs. Due to ubiquitous use of social networks in recent years, an enormous amount of opinion is posted via the Web. Application of text mining techniques on social networking websites can reveal significant results related to person-to-person interaction behaviors. In social networks, textual data may be large, noisy, and dynamic. Moreover, interpreting emoticons (Smile, Sad) for expressing any specific concept or emotion is still a challenging issue for researchers. Privacy and trust in online communication is also a major issue. Application of ethical values, such as integrity, veracity, in online communication is the only effective way to build trust online. Moreover, text mining techniques in conjunction with social networks can be used for finding general opinion about any specific subject, human thinking patterns, and group identification in large-scale systems. For this confusion this paper presents the similarity recommendation of friends to verify which opinion is the highest effect.

II. LITERATURE REVIEW

From the author K. Naga Pavitra & B. Shereesha present Friendbook [5], a novel semantic-based friend recommendation system for social networks, which

recommends friends to users based on their life styles instead of social graphs. By taking advantage of sensor rich smart phone, Friendbook discovers life styles of users from user-centric sensor data, measures the similarity of life styles between users, and recommends friends to users if their life styles have high similarity. Inspired by text mining, we model a user's daily life as life documents, from which his/her life styles are extracted by using the Latent Dirichlet Allocation algorithm. We further propose a similarity metric to measure the similarity of life styles between users, and calculate users' impact in terms of life styles with a friend matching graph. This paper focuses on providing the overview about the various categories of recommendation techniques developed till now. Finally, Friendbook integrates a feedback mechanism to further improve the recommendation accuracy.

Sachin V Jose, Minu Lalitha Madhavu [6] analyze the Social networking sites imply friend recommendation Systems in contribution to providing better user experiences. Online friend recommendation is a rapid developing topic in web mining. Current social networking servicing recommend friends to users based on their social graphs and mutual friends , which may not be the most appropriate to reflect a user's taste on friend selection in real lifetime . In this paper propose a system that recommends friends based on the daily activities of users. Here a semantic based friend recommendation is done based on the users life styles. By using text mining, we display a user's everyday life as life archives, from which his/her ways of life are separated by using the Latent Dirichlet Allocation algorithm. At that point we discover a similarity metric to quantify the similarity of life styles between users, and ascertain users effect as far as ways of life with a similarity matching

Namrata et al [7] From their 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.

Chaitali H. Hegshetye, [8] et al discusses the recommendation in web logs and similar social networks. First, we present a collaborative recommendation using the link structure of a social network and semantic-based recommendation using mutual friend's interests. Next, we describe the application of an approach to a small representative subset of a large real-world. Friendbook discovers life styles of users from user-centric sensor data then it measures the similarity of life styles between users and recommends friends to users if their life styles have high similarity. Inspired by text mining, we present a user's daily life as life documents. This project focuses on providing the overview about the various categories of recommendation techniques developed till now.

Adomavicius and Mangina, The most important objective of recommended systems is to estimate the ratings for the items that are new for a user [9]. Ultimately, after calculating the estimated rates for the yet unrated items, an ordered list of most related items can be prepared and suggested to the target user. A number of previous studies have revealed the contribution of recommended systems in education. A collaborative filtering method was used in a research to recommend documents that will either encourage the users to expand their knowledge of a given topic [10].

III. FRIEND RECOMMENDATION

Recommendation systems are appropriate tools for provide useful and suitable recommendations in social networks. Nowadays web users are not only consumers

664

of information, but they actively participate in social networks to share their opinion in public. Social network based-recommendation uses for improve recommendation systems because of its benefits [11]. For example, as long as cold-start users are connected to the social network, it can deal with them. Main goal of this recommendation system is recommending to the similarity based one, for this similarity measures text analyses is a main tool to analyze the similarities, here we discuss the similarity measures using the algorithm.

a. Rocchio Algorithm

Rocchio Algorithm (*RA*) is used to find the relevance analyses of the people, it is based on the similarity measures, and this algorithm is used to find the similarity between the people across the world with the help of SN. It is a feedback method and is mainly used for text analyses [12]. However, the user of SN who comments their opinion for any post, this algorithm is used to find out the relevant opinion across the world for the same post. The user must have sufficient knowledge for any kind of post and to indicate relevance feedback with the help of algorithm, it can find the relevant opinion. The great drawback of this approach is, this algorithm may not work efficiently when the user spells a word in a different way. Various spelling correction techniques can be used at the cost of computation and response time, such as hashing based and context-sensitive spelling correction techniques.

b. K-Nearest Neighbour (K-NN)

K-Nearest Neighbour (*K-NN*) algorithm is used to find the nearest opinion friends in the network, similarity opinion of friends are nearby location in the network to find the friend this algorithm is helpful one to recommend to those people. This algorithm is a form of instant based learning. The algorithm label similar objects based on the neighboring feature space in the training set. The neighboring feature space may be determined by measuring the angle between the two feature vectors or by calculating the Euclidean distance between the vectors. With the help of this algorithm it finds the nearest similarity people [13].

c. Support Vector Machine (*SVM*)

SVM is one of the decision tree algorithms to find out the correlation between the objects, Support Vector Machine (*SVM*) algorithm is used to analyze data in classification analysis. It is used to classify the opinion (i.e the opinion of the people is which amongst the group) [14, 15]. Here the SVM algorithm uses both negative and positive training datasets to construct a hyper plane that separates the most common group and most dis-common group. This helps to find out the different clusters of the opinion for the post using the text analyses. It helps to find which opinion is closest to highest percentage and which one is lowest percentage.

d. Expectation Maximization (EM)

The EM algorithm is used to searches the result for a maximum likelihood. It estimates the expected values of the hidden objects. The hidden variables are the parameters of the model. In this case, we use mixture of Gaussians; hence the hidden variables are the mean and standard deviation for each Gaussian distribution [16, 17]. We start with an initial estimate of those parameters and iteratively run the algorithm to find the maximum likelihood (ML) for our estimates. The reason we are using EM is to fit the similarity opinion better, With the help of the EM algorithm it finds out the hidden opinion for the post. It iteratively process the opinion various over time to analyze the similar content in the social network for a post.

IV. COMPARISON OF VARIOUS TECHNIQUES

Algorithm	Advantage	Disadvantage
Rocchio Algorithm	Find the relevance analyses of the people	Failed in same opinion with different spelling
K-Nearest Neighbour (K-NN	Find the nearest similarities in the network	Failed in large distance and more expensive in large distance
Support Vector Machine	It is able to separate the opinion in different group.	Heavy Time taken for analyzing.
Expectation Maximization (EM)	Find hidden opinions, more accuracy, higher detection of opinion	Consumption of more energy

Table 1: Comparison analyses of different technique

V. CONCLUSION

Social network is one of the important factors in this growing world. People of today where split across the world for their personal reasons, to group up those peoples this paper provides the survey of various techniques of text mining analyses for recommending the people. It recommends the same opinion people for the post. It groups the people for recommending to their friends. This survey shows the different technique for recommending the friends. From the best of our knowledge EM method performs well to recommend the system in the social network.

VI REFERENCES

[1]. S.Thiraviya Regina Rajam1 and Dr. S.Britto.Ramesh Kumar, "SOCIAL NETWORK SERVICES: ANOVERVIEW".

[2] H. Sundaram, L. Xie, M. De Choudhury, Y. Lin, and A. Natsev,"Multimedia semantics: Interactions between content andcommunity," Proc. IEEE, vol. 100, no. 9, pp. 2737–2758, Sep. 2012.

[3] Akther, A., Alam, K. M., Kim, H.-N., & Saddik, A. E. (2012). Social Network and User Context Assisted Personalization for Recommender Systems. *Innovations in Information Technology (IIT)*, (pp. 95 - 100).

[4] Zhibo Wang, "Friendbook: A Semantic-Based Friend Recommendation System for Social Networks"

[5] K. NAGA PAVITRA & B. SHEREESHA, "Improve the Performance of Social Networks Security Based on Semantic-Based Friend Recommendation System".

[6] Sachin V Jose1, Minu Lalitha Madhavu, "A Survey Of Activity Based Friend Recommendation Services"

[7] Namrata M.Eklaspur, Anand S.Pashupatimath, "A Friend Recommender System for Social Networks by Life Style Extraction Using Probabilistic Method - "Friendtome"

[8] Chaitali H. Hegshetye, Namrata C. Bandbe, Prof. Yogesh. M. Gajmal, "Survey on Friend Recommendation System for Social Network's"

[9] Adomavicius, G. and A. Tuzhilin, toward the next generation of recommender systems: A survey of the state-of-the-art and possible extensions. Knowledge and Data Engineering, IEEE Transactions on, 2005.

[10] Mangina, E. and J. Kilbride, Evaluation of key phrase extraction algorithm and tiling process for a document/resource recommender within e-learning environments. Computers & Education, 2008.

[11] R i z w a n a i r f a n, "A Survey on Text Mining in Social Networks"

[12] Zhao, P., Han, J. & Sun, Y. 2009. P-Rank: A comprehensive structural similarity measure over

information networks. In *Proceedings of 18th ACM Conference on Information and Knowledge Management*, New York, 233-238

[13] Aci, M., Inan, C. & Avci, M. 2010. A hybrid classification method of k-nearest neighbour, bayesian method and genetic algorithm. *Expert Systems with Applications*, 37(7), 5061-5067.

[14] Chang, M. & Poon, C. K. 2009. Using phrases as features in e-mail classification. *Journal of System and Software*, 82(6), 1036-1945

[15] Chang, M. & Poon, C. K. 2009. Using phrases as features in e-mail classification. *Journal of System and Software*, 82(6), 1036-1945

[16] B. A. Frigyik, A. Kapila, and M. R. Gupta, "Introduction to the Dirichlet distribution and related processes," Dept. Elect. Eng., Univ. Washington, Seattle, WA, USA, UWEETR-2010-0006, 2010.

[17] Carmagnola, F., Vernero, F., & Grillo, P. (2009). SoNARS: A Social Networks-Based Algorithm for Social Recommender Systems. *17th International Conference, UMAP 2009, formerly UM and AH. 5535*, pp. 223-234. Trento, Italy: Springer Berlin Heidelberg.

VII AUTHOR'S BIOGRAPHY

R.NIVEDHITHA Research Scholar received her M.Sc Computer Science degree in 2015 from Kongunadu college of Arts & Science, Coimbatore. She is pursuing her M.Phil at Avinashilingam Institute for Home Science and Higher Education for Women University, Coimbatore. Her areas of interest are Data Mining, Security.