ISSN No: 2348-4845
International Journal & Magazine of Engineering, Technology, Management and Research
A Peer Reviewed Open Access International Journal

A Semantic-based Friend System for Social Networks

M.Dinesh
M.Tech Student,
Department of CSE,
Sree Rama institute of Technology and Science,
Kuppenakuntla, Penuballi, Khammam,TS India.

P.Karunakar Reddy
Assistant Professor,
Department of CSE,
Sree Rama institute of Technology and Science,
Kuppenakuntla, Penuballi, Khammam,TS India.

ABSTRACT:

Existing social networking services recommend friends to users based on their social graphs, which may not be the most appropriate to reflect a user's preferences on friend selection in real life. In this paper, we present Friendbook, 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 smartphones, 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. Upon receiving a request, Friendbook returns a list of people with highest recommendation scores to the query user. Finally, Friendbook integrates a feedback mechanism to further improve the recommendation accuracy. We have implemented Friendbook on the Android-based smartphones, and evaluated its performance on both small-scale experiments and large-scale simulations. The results show that the recommendations accurately reflect the preferences of users in choosing friends.

Index Terms:

Friend recommendation, mobile sensing, social networks, life style.

INTRODUCTION:

Twenty years ago, people typically made friends with others who live or work close to themselves, such as neighbors or colleagues. We call friends made through this traditional fashion as G-friends, which stands for geographical location-based friends because they are influenced by the geographical distances between each other. With the rapid advances in social networks, services such as Facebook, Twitter and Google+ have provided us revolutionary ways of making friends. According to Facebook statistics, a user has an average of 130 friends, perhaps larger than any other time in history [2]. One challenge with existing social networking services is how to recommend a good friend to a user. Most of them rely on pre-existing user relationships to pick friend candidates. For example, Facebook relies on a social link analysis among those who already share common friends and recommends symmetrical users as potential friends. Unfortunately, this approach may not be the most appropriate based on recent sociology findings [16], [27], [29], [30]. According to these studies, the rules to group people together include: 1) habits or life style; 2) attitudes; 3) tastes; 4) moral standards; 5) economic level; and 6) people they already know.

Apparently, rule #3 and rule #6 are the mainstream factors considered by existing recommendation systems. Rule #1, although probably the most intuitive, is not widely used because users' life styles are difficult, if not impossible, to capture through web actions. Rather, life styles are usually closely correlated with daily routines and activities. Therefore, if we could gather information on users' daily routines and activities, we can exploit rule #1 and recommend friends to people based on their similar life styles.

ISSN No: 2348-4845
International Journal & Magazine of Engineering, Technology, Management and Research
A Peer Reviewed Open Access International Journal

This recommendation mechanism can be deployed as a standalone app on smartphones or as an add-on to existing social network frameworks. In both cases, Friendbook can help mobile phone users find friends either among strangers or within a certain group as long as they share similar life styles. In our everyday lives, we may have hundreds of activities, which form meaningful sequences that shape our lives. In this paper, we use the word activity to specifically refer to the actions taken in the order of seconds, such as "sitting", "walking", or "typing", while we use the phrase life style to refer to higher-level abstractions of daily lives, such as "office work" or "shopping". For instance, the "shopping" life style mostly consists of the "walking" activity, but may also contain the "standing" or the "sitting" activities.

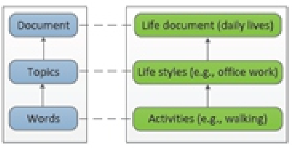

Fig. 1: An analogy between word documents and people's daily lives.

Existing System:

Most of the friend suggestions mechanism relies on pre-existing user relationships to pick friend candidates. For example, Facebook relies on a social link analysis among those who already share common friends and recommends symmetrical users as potential friends. The rules to group people together include:

1)Habits or life style
2)Attitudes
3)Tastes
4)Moral standards
5)Economic level; and
6)People they already know.

Apparently, rule #3 and rule #6 are the mainstream factors considered by existing recommendation systems.

DISADVANTAGES OF EXISTING SYSTEM:

Existing social networking services recommend friends to users based on their social graphs, which may not be the most appropriate to reflect a user's preferences on friend selection in real life.

Proposed System:

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 smartphones, 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.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.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.We integrate a linear feedback mechanism that exploits the user's feedback to improve recommendation accuracy.

ADVANTAGES OF PROPOSED SYSTEM:

Recommendeds potential friends to users if they share similar life styles. The feedback mechanism allows us to measure the satisfaction of users, by providing a user interface that allows the user to rate the friend list.

RELATED WORK:

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, Amazon [1] recommends items to a user based on items the user previously visited, and items that other users are looking at. Netflix [3] and Rotten Tomatoes [4] recommend movies to a user based on the user's previous ratings and watching habits. Recently, with the advance of social networking systems, friend recommendation has received a lot of attention. Generally speaking, existing friend recommendation in social networking systems, e.g., Facebook, LinkedIn and Twitter, recommend friends to users if, according to their social relations, they share common friends. Meanwhile, other recommendation mechanisms have also been proposed by researchers. For example, Bian and Holtzman [8] presented MatchMaker, a collaborative filtering friend recommendation system based on personality matching.Kwon and Kim [20] proposed a friend recommendation method using physical and social context. However, the authors did not explain what the physical and social context is and how to obtain the information. Yu et al.

ISSN No: 2348-4845
International Journal & Magazine of Engineering, Technology, Management and Research
A Peer Reviewed Open Access International Journal

[32] recommended geographically related friends in social network by combining GPS information and social network structure. Hsu et al. [18] studied the problem of link recommendation in weblogs and similar social networks, and proposed an approach based on collaborative recommendation using the link structure of a social network and content-based recommendation using mutual declared interests. Gou et al. [17] proposed a visual system, SFViz, to support users to explore and find friends interactively under the context of interest, and reported a case study using the system to explore the recommendation of friends based on people's tagging behaviors in a music community. These existing friend recommendation systems, however, are significantly different from our work, as we exploit recent sociology findings to recommend friends based on their similar life styles instead of social relations.

SYSTEM OVERVIEW:

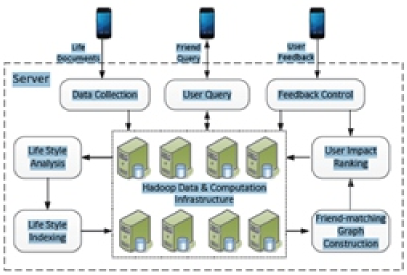

Fig. 2: System architecture of Friendbook.

In this section, we give a high-level overview of the Friendbook system. Figure 2 shows the system architecture of Friendbook which adopts a client-server mode where each client is a smartphone carried by a user and the servers are data centers or clouds. On the client side, each smartphone can record data of its user, perform real-time activity recognition and report the generated life documents to the servers. It is worth noting that an offline data collection and training phase is needed to build an appropriate activity classifier for real-time activity recognition on smartphones. We spent three months on collecting raw data of 8 volunteers for building a large training data set. As each user typically generates around 50MB of raw data each day, we choose MySQL as our low level data storage platform and Hadoop MapReduce as our computation infrastructure. After the activity classifier is built, it will be distributed to each user's smartphone and then activity recognition can be performed in real-time manner. As a user continually uses Friendbook, he/she will accumulate more and more activities in his/her life documents, based on which, we can discover his/her life styles using probabilistic topic model.

On the server side, seven modules are designed to fulfill the task of friend recommendation. The data collection module collects life documents from users' smartphones. The life styles of users are extracted by the life style analysis module with the probabilistic topic model. Then the life style indexing module puts the life styles of users into the database in the format of (life-style, user) instead of (user, life-style). A friend-matching graph can be constructed accordingly by the friend-matching graph construction module to represent the similarity relationship between users' life styles. The impacts of users are then calculated based on the friend-matching graph by the user impact ranking module. The user query module takes a user's query and sends a ranked list of potential friends to the user as response. The system also allows users to give feedback of the recommendation results which can be processed by the feedback control module. With this module, the accuracy of friend recommendation can be improved.

QUERY AND FRIEND RECOMMENDATION:

Before a user initiates a request, he/she should have accumulated enough activities in his/her life documents for efficient life styles analysis. The period for collecting data usually takes at least one day. Longer time would be expected if the user wants to get more satisfied friend recommendation results. After receiving a user's request (e.g., life documents), the server would extract the user's life style vector, and based on which recommend friends to the user. The recommendation results are highly dependent on users' preference. Some users may prefer the system to recommend users with high impact, while some users may want to know users with the most similar life styles.

Algorithm 1 Computing users' impact ranking

Input: The friend-matching graph G.
Output: Impact ranking vector $\mathbf{r}$ for all users.

for $i = 1$ to n **do**
  $\mathbf{r}_0(i) = \frac{1}{n}$
end for
$\delta = \infty$
$\epsilon = e^{-9}$
while $\delta > \epsilon$ **do**
  for $i = 1$ to n **do**
    $\mathbf{r}_{k+1}(i) = \sum_j \frac{1-\varphi}{n} \mathbf{r}_k(j) + \varphi \frac{\sum_j \omega(i,j) \cdot \mathbf{r}_k(j)}{\sum_j \omega(i,j)}$
  end for
  $\delta = \sum_{i=1}^{n} |\mathbf{r}_{k+1}(i) - \mathbf{r}_k(i)|$
end while
return $\mathbf{r}$

It is also possible that some users want the system to recommend users who have high impact and also similar life styles to them.

ISSN No: 2348-4845
International Journal & Magazine of Engineering, Technology, Management and Research
A Peer Reviewed Open Access International Journal

To better characterize this requirement, we propose the following metric to facilitate the recommendation, Ri(j) = _S(i; j) + (1 _)rj_ where Ri(j) is the recommendation score of user j for the query user i, S(i; j) is the similarity between user i and user j, and rj is the impact of user j. _ 2 [0; 1] is the recommendation coefficient characterizing users' preference. _ is introduced to make S(i; j) and rj in the same order of magnitude, which can be roughly set to n=10, where n is the number of users in the system. When _ = 1, the recommendation is solely based on the similarity; when _ = 0, the recommendation is solely based on the impact ranking.

Algorithm 2 Friend recommendation

Input: The query user i, the recommendation coefficient β and the required number of recommended friends from the system p.
Output: Friend list F_i.

$F_i \leftarrow \emptyset$, $Q \leftarrow \emptyset$
extracts i's life style vector $\mathbf{L}_i$ using the LDA algorithm.
for each life style z_k the probability of which in $\mathbf{L}_i$ is not zero **do**
    put users in the entry of z_k into Q
end for
for each user $j \notin Q$ **do**
    $S(i, j) \leftarrow 0$
end for
for each user j in the database **do**
    $R_i(j) = \beta S(i, j) + (1 - \beta) r_j \kappa$
end for
sort all users in decreasing order according to $R_i(j)$
put the top p users in the sorted list to F_i

Friendbook also uses GPS location information to help users find friends within some distance. In order to protect the privacy of users, a region surrounding the accurate location will be uploaded to the system. When a user uses Friendbook, he/she can specify the distance of friends before recommendation. In this way, only friends having similarity with the user within the specified distance can be recommended as friends. Privacy is very important especially for users who are sensitive to information leakage. In our design of Friendbook, we also considered the privacy issue and the existing system can provide two levels of privacy protection.

First, Friendbook protects users' privacy at the data level. Instead of uploading raw data to the servers, Friendbook processes raw data and classifies them into activities in real-time. The recognized activities are labeled by integers. In this way, even if the documents containing the integers are compromised, they cannot tell the physical meaning of the documents. Second, Friendbook protects users' privacy at the life pattern level. Instead of telling the similar life styles of users, Friendbook only shows the recommendation scores of the recommended friends with the users. With the recommendation score, it is almost impossible to infer the life styles of recommended friends.

CONCLUSIONS:

In this paper, we presented the design and implementation of Friendbook, a semantic-based friend recommendation system for social networks. Different from the friend recommendation mechanisms relying on social graphs in existing social networking services, Friendbook extracted life styles from user-centric data collected from sensors on the smartphone and recommended potential friends to users if they share similar life styles. We implemented Friendbook on the Android-based smartphones, and evaluated its performance on both smallscale experiments and large-scale simulations. The results showed that the recommendations accurately reflect the preferences of users in choosing friends.Beyond the current prototype, the future work can be four-fold. First, we would like to evaluate our system on large-scale field experiments. Second, we intend to implement the life style extraction using LDA and the iterative matrix-vector multiplication method in user impact ranking incrementally, so that Friendbook would be scalable to large-scale systems.

Third, the similaritythreshold used for the friend-matching graph is fixed in our current prototype of Friendbook. It would be interesting to explore the adaption of the threshold for each edge and see whether it can better represent the similarity relationship on the friend-matching graph. At last, we plan to incorporate more sensors on the mobile phones into the system and also utilize the information from wearable equipments (e.g., Fitbit, iwatch, Google glass, Nike+, and Galaxy Gear) to discover more interesting and meaningful life styles. For example, we can incorporate the sensor data source from Fitbit, which extracts the user's daily fitness infograph, and the user's place of interests from GPS traces to generate an infograph of the user as a "document". From the infograph, one can easily visualize a user's life style which will make more sense on the recommendation. Actually, we expect to incorporate Friendbook into existing social services (e.g., Facebook, Twitter, LinkedIn) so that Friendbook can utilize more information for life discovery, which should improve the recommendation experience in the future.

REFERENCES:

[1] Amazon. http://www.amazon.com/.

[2] Facebook statistics. http://www.digitalbuzzblog.com/facebook-statistics-stats-facts-2011/.

[3] Netfix. https://signup.netflix.com/.

[4] Rotten tomatoes. http://www.rottentomatoes.com/.

[5] G. R. Arce. Nonlinear Signal Processing: A Statistical Approach.John Wiley & Sons, 2005.

[6] B. Bahmani, A. Chowdhury, and A. Goel. Fast incremental andpersonalized pagerank. Proc. of VLDB Endowment, volume 4, pages173-184, 2010.

[7] J. Biagioni, T. Gerlich, T. Merrifield, and J. Eriksson. EasyTracker: Automatic Transit Tracking, Mapping, and Arrival Time Prediction Using Smartphones. Proc. of SenSys, pages 68-81, 2011.

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

[9] C. M. Bishop. Pattern recognition and machine learning. Springer New York, 2006.

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

[11] P. Desikan, N. Pathak, J. Srivastava, and V. Kumar. Incremental page rank computation on evolving graphs. Proc. of WWW, pages 1094-1095, 2005.

[12] N. Eagle and A. S. Pentland. Reality Mining: Sensing Complex Cocial Systems. Personal Ubiquitous Computing, 10(4):255-268, March 2006.

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

[14] 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.

[15] B. A. Frigyik, A. Kapila, and M. R. Gupta. Introduction to the dirichlet distribution and related processes. Department of Electrical Engineering, University of Washigton, UWEETR-2010-0006, 2010.

[16] A. Giddens. Modernity and Self-identity: Self and Society in the late Modern Age. Stanford Univ Pr, 1991.

[17] 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.

[18] 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.

Author's

M.Dinesh is a student of Sree Rama Institute of Technology & Science, Kuppenakuntla,Penuballi, Khammam, TS,India.Presently he is Pursuing his M.Tech (CSE) from this college His area of interests includes Information Security, Cloud Computing, Data Communication & Networks.

P. Karunakar Reddy is an efficient teacher, received M.Tech from JNTU Kakinada is working as an Assistant Professor in Department of C.S.E, Sree Rama Institute of Technology &Science, Kuppenakuntla, Penuballi, Khammam, TS,India. He has published many papers in both National & International Journals. His area of Interest includes Data Communications & Networks, Information Security, Database Management Systems, Computer Organization, C Programming and other advances in Computer Applications