

Friends Wall: A Semantic-based Friend Recommendation System for Social Networks

Mr. Raj Agarwal¹, Mr. Swaranjeet Singh², Ms. Samiksha Bhujbal³, Ms. Pooja Jadhav⁴, Prof. U.H. Wanaskar⁵

UG Student, Comp. Engg, PVPIT, Pune, India ^{1,2,3,4}

Assistant Professor, Comp. Engg, PVPIT, Pune, India ⁵

Abstract: Present social networking services suggest friends to users based on their social activities, which may not be the most suitable to react a users taste on friend choice in real life. In this project, we present Friendswall, a semantic-based friend recommendation system for social networks, which present friends to users based on their life styles instead of social activities. By taking benefit of sensor-rich Smartphones, Friendswall discovers life styles of users, measures the resemblance of life styles between users, and suggest friends to users if their life styles have high resemblance. Inspired by text mining, in this project we design a users daily life as life documents, from which his/her life styles are taken out by using the Naive Bayes algorithm. We further aim a similarity metric to measure the resemblance of life styles between users, and estimate users impact in terms of life styles with a friend-matching graph. Upon acquiring a request, FriendsWall returns a list of people with maximum resemblance scores to the query user. At last, FriendsWall incorporate a feedback mechanism to further amend the recommendation accuracy. This project will build FriendsWall and evaluate its performance on both small-scale experiments and large-scale model.

Keywords: Friend recommendation, Mobile sensing, Social networks, Life style, Data Mining, Machine Learning.

I. INTRODUCTION

Twenty years ago, people typically made friends with others who live or work close to themselves, such as neighbours 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. 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. According to these studies, the rules to group people together include: 1) habits or life style; 2) attitudes; 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. 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.

To model daily lives properly, we draw an analogy between people's daily lives and documents, as shown in Figure 1. Previous research on probabilistic topic models in text mining has treated documents as mixtures of topics, and topics as mixtures of words. Inspired by this, similarly, we can treat our daily lives (or life documents) as a mixture of life styles (or topics), and each life style as a mixture of activities (or words). Observe here, essentially, we represent daily lives with "life documents", whose semantic meanings are reflected through their topics, which are life styles in our study.

Just like words serve as the basis of documents, people's activities naturally serve as the primitive vocabulary of these life documents. Our proposed solution is also motivated by the recent advances in smart phones, which have become more and more popular in people's lives. These smart phones (e.g., iPhone or Android-based smart phones) are equipped with a rich set of embedded sensors, such as GPS, accelerometer, microphone, gyroscope, and camera.



II. LITERATURE SURVEY

In this paper[2],in order to facilitate the introduction of transit tracking and arrival time prediction in smaller transit agencies, we investigate an automatic, smartphone-based system which we call EasyTracker. To use EasyTracker, a transit agency must obtain smartphones, install an app, and place a phone in each transit vehicle. Our goal is to require no other input. This level of automation is possible through a set of algorithms that use GPS traces collected from instrumented transit vehicles to determine routes served, locate stops, and infer schedules. In addition, online algorithms automatically determine the route served by a given vehicle at a given time and predict its arrival time at upcoming stops.

In this paper[3], we propose a method to incrementally compute PageRank for a large graph that is evolving. Our approach is quite general, and can be used to incrementally compute (on evolving graphs) any metric that satisfies the first order Markov property

General Terms: Algorithms, Measurement, Performance, Reliability, Security, Standardization, Theory, Verification.

In this paper[4],there is relatively little work on the investigation of large-scale human data in terms of multimodality for human activity discovery. In this paper, we suggest that human interaction data, or human proximity, obtained by mobile phone Bluetooth sensor data, can be integrated with human location data, obtained by mobile cell tower connections, to mine meaningful details about human activities from large and noisy datasets.

We propose a model, called bag of multimodal behavior,that integrates the modeling of variations of location over multiple time-scales, and the modeling of interaction types from proximity. Our representation is simple yet robust to characterize real-life human behavior sensed from mobile phones, which are devices capable of capturing large-scale data known to be noisy and incomplete. We use an unsupervised approach, based on probabilistic topic models, to discover latent human activities in terms of the joint interaction and location behaviors of 97 individuals over the course of approximately a 10-month period using data from MIT's Reality

In this paper[5],the popularization of GPS-enabled mobile devices provides social network researchers a taste of cyber-physical social network in advance. Traditional link prediction methods are designed to find friends solely relying on social network information. With location and trajectory data available, we can generate more accurate and geographically related results, and help web-based social service users find more friends in the real world. Aiming to recommend geographically related friends in social network, a three-step statistical recommendation approach is proposed for GPS-enabled cyber-physical social network. By combining GPS information and social network structures, we build a pattern-based heterogeneous information network. Links inside this network reflect both people's geographical information, and their social relationships. Our approach estimates link relevance and finds promising geo-friends by employing a random walk process on the heterogeneous information network. Empirical studies from both synthetic datasets and reallife dataset demonstrate the power of merging GPS data and social graph structure, and suggest our method outperforms other methods for friends recommendation in GPS-based cyberphysical social network.

In this paper[6],over the last decade, Page Rank has gained importance in a wide range of applications and domains, ever since it first proved to be effective in determining node importance in large graphs (and was a pioneering idea behind Google's search engine). In distributed computing alone, Page Rank vector, or more generally random walk based quantities have been used for several different applications ranging from determining important nodes, load balancing, search, and identifying connectivity structures. Surprisingly, however, there has been little work towards designing provably efficient fully distributed algorithms for computing Page Rank. The difficulty is that traditional matrix-vector multiplication style iterative methods may not always adapt well to the distributed setting owing to communication bandwidth restrictions and convergence rates.

In this paper[7],we introduce a system for sensing complex social systems with data collected from 100 mobile phones over the course of 9 months. We demonstrate the ability to use standard Bluetooth-enabled mobile telephones to measure information access and use in different contexts, recognize social patterns in daily user activity, infer relationships, identify socially significant locations, and model organizational rhythms.

In this paper[8], user mobility has given rise to a variety of Web applications, in which the global positioning system (GPS) plays many important roles in bridging between these applications and end users. As a kind of human behavior, transportation modes, such as walking and driving, can provide pervasive computing systems with more contextual information and enrich a user's mobility with informative knowledge. In this article, we report on an approach based on supervised learning to automatically infer users' transportation modes, including driving,walking, taking a bus and



riding a bike, from raw GPS logs. Our approach consists of three parts: a change point-based segmentation method, an inference model and a graph-based post-processing algorithm. First, we propose a change point-based segmentation method to partition each GPS trajectory into separate segments of different transportation modes. Second, from each segment, we identify a set of sophisticated features, which are not affected by differing traffic conditions (e.g., a person's direction when in a car is constrained more by the road than any change in traffic conditions). Later, these features are fed to a generative inference model to classify the segments of different modes. Third, we conduct graph-based post processing to further improve the inference performance. This post processing algorithm considers both the commonsense constraints of the real world and typical user behaviors based on locations in a probabilistic manner. The advantages of our method over the related works include three aspects. (1) Our approach can effectively segment trajectories containing multiple.

III. PROBLEM STATEMENT

Problem Definition:-The system is to define a unique similarity metric to characterize the similarity of users in terms of life styles and then construct a friend-matching graph to recommend friends to users based on their life styles.

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

Statement of Scope:-

- A semantic-based friend recommendation system for social networks, which recommends friends to users based on their life styles instead of social graphs
- FriendsWall discovers life styles of users & measures the similarity of life styles between users, and recommends friends to users if their life styles have high similarity.
- We model a users 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 users feedback to improve recommendation accuracy.

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:

- Recommended 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

IV. DETAILED DESIGN AND IMPLEMENTATION

Admin

In this module, the Admin has to login by using valid user name and password. After login successful he can do some operations such as add details, view details, add groups, view groups, from group, list of users, view user feedback, view user query, Android mobile users, view all user ranks, view friend match graph and logout. The admin can view the registered users, and also admin can view the user feedback, view user query, and also view the android application users.

Add Groups

In this module, the admin can add number of groups. After adding a group admin can view the all groups, and also admin can add user to particular group, after adding successful he will get a response from the server.

View friend match graph

In this module, the admin can view the matching friends. If the admin click on view friend match graph button, then all user matching graph will display with their tags such as habit graph, attitude graph, tastes graph, moral standard graph and economic level graph with all user names.

Users

In user's module, the admin can view the list of users and list of mobile users. Mobile user means android application users.

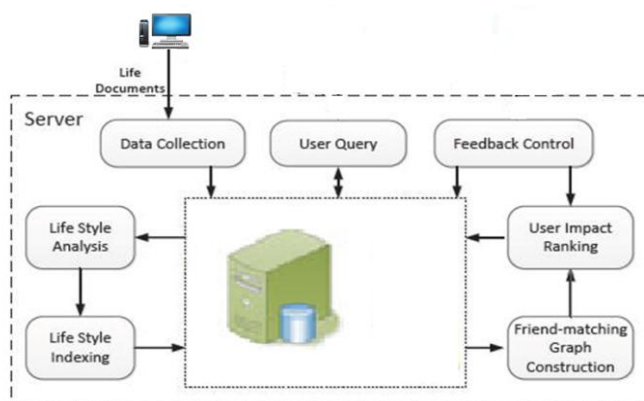


Fig: System Architecture

User

In this module, there are n numbers of users are present. User should register before doing some operations. And register user details are stored in admin module. After registration successful he has to login by using authorized user name and password. Login successful he will do some operations like search users, send friend request, view your rank, send query, send feedback, Recommend the friend and logout.

Search Users

The user can search the users based on category & sub category. And the server will give response to the user like User name, user image, E mail id, phone number, date of birth, attitude, habit, economy level and people known. If you want send friend request to particular receiver then click on request, then request will send to the user.

Send Request

In this module, the user can send friend request to another user. If the user clicks on send request button, then the friend request will send to particular user. And also user can view the all request. After accepting the request response will send to user. And also user rank will be increased based on the friends.

Recommend the friend

In this module, the user can add recommend and view recommends. If the user clicks on add recommend button, then user can select category and sub category and click on search, then server will give response to user details, then click on recommend button and select to user recommend and submit and recommend message will send to the particular user.

**Send query**

In this module, the user can send query message to another user. If user want to send query message to another user, then enter group name, about and write query and send, that query will send to particular user.

V. ALGORITHMIC STRATEGY**TRAINMULTINOMIALNB(C, D)**

```

1  V ← EXTRACTVOCABULARY(D)
2  N ← COUNTDOCS(D)
3  for each c ∈ C
4  do Nc ← COUNTDOCSINCLASS(D, c)
5     prior[c] ← Nc/N
6     textc ← CONCATENATETEXTOFALLDOCSINCLASS(D, c)
7     for each t ∈ V
8     do Tct ← COUNTTOKENSOFTERM(textc, t)
9     for each t ∈ V
10    do condprob[t][c] ←  $\frac{T_{ct}+1}{\sum_{t'}(T_{ct'}+1)}$ 
11  return V, prior, condprob

```

APPLYMULTINOMIALNB(C, V, prior, condprob, d)

```

1  W ← EXTRACTTOKENSFROMDOC(V, d)
2  for each c ∈ C
3  do score[c] ← log prior[c]
4     for each t ∈ W
5     do score[c] += log condprob[t][c]
6  return arg maxc∈C score[c]

```

VI. MATHEMATICAL MODEL

- **Set theory:**
System S = {Input, Output, Function, Success, Failure, Dataset}
- **Input:**
Input = {UP1, UP2, UPn}
where UP = User Post
- **Output:**
Output = {Recommendation List}
- **Function:**
Function = {f1, f2, f3, f4, f5, f6, f7}
where,
f1: Data collection
f2: Life style analysis module
f3: Life style indexing
f4: Friend's matching graph construction
f5: User impact ranking module
f6: User Query
f7: Feedback control module.
- **Success:** Successfully recommended friend.
- **Failure:** Unable to generate friend recommendation list. **Dataset:** {Post, Document search}

VII. RESULT ANDEVALUATION

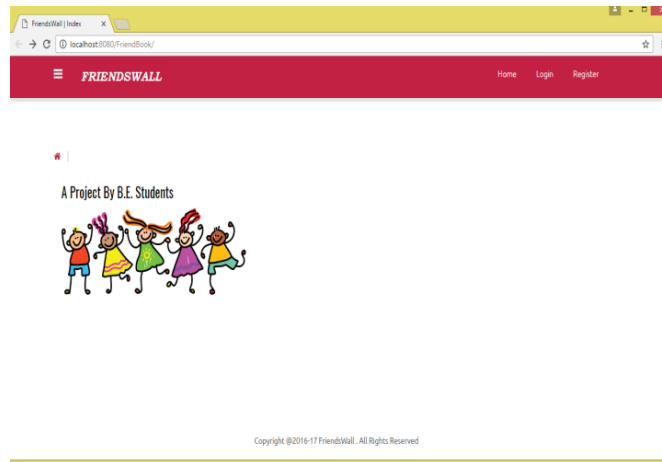


Fig1.Home Page

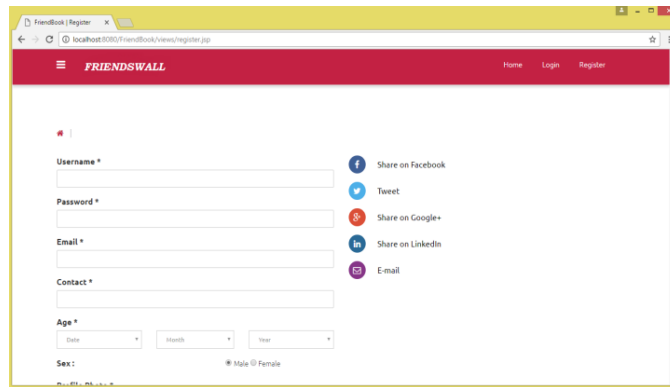


Fig2.Register Page

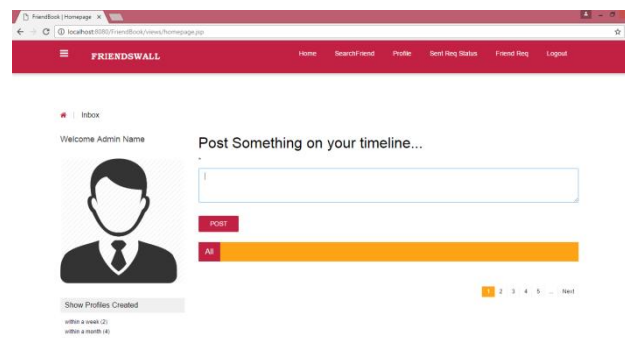


Fig3. Admin Panel

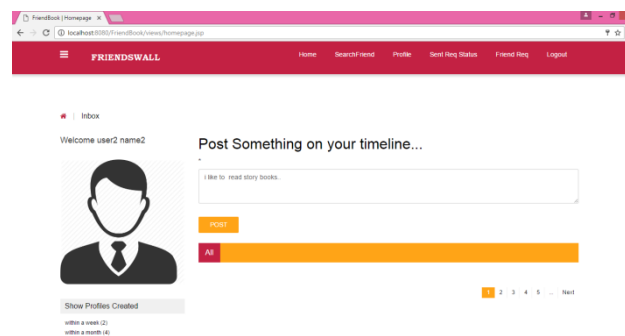


Fig4. User Profile

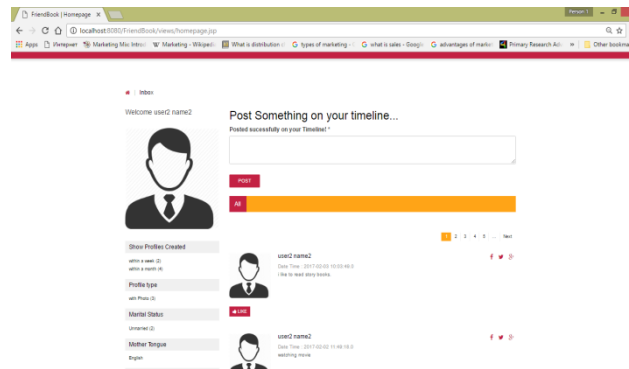


Fig5. Userpost

VIII. CONCLUSION

In this paper, we presented the design and implementation of FriendsWall, a semantic-based friend recommendation system for social networks. Different from the friend recommendation mechanisms relying on social graphs in existing social networking services. In our methodology, we exhibited the design and usage of Friend search, a semantic-based friend proposal framework for social networks. Not the same as the friend proposal instruments depending on social charts in existing social networking administrations, the outcomes demonstrated that the suggestions precisely mirror the inclinations of clients in picking companions. Past and present model, the future work can be focused on actualizing it on other social networking, and same can be utilized to manufacture a stand-alone application and access the client movement through versatile sensors.

ACKNOWLEDGMENT

We take this golden opportunity to owe our deep sense of gratitude to our project guide Prof. U. H. Wanaskar, for her instinct help and valuable guidance with a lot of encouragement throughout this paper work, right from selection of topic work up to its completion. Our sincere thanks to Head of the Department of Computer Engineering Prof.B.K..Sarkar who continuously motivated and guided us for completion of this paper. I am also thankful to all teaching and non-teaching staff members, for their valuable suggestions and valuable co-operation for partially completion of this work. We specially thank to those who helped us directly-indirectly in completion of this work successfully.

REFERENCES

- [1] Zhibo Wang, Jilong Liao, "Friendbook: A Semantic-Based Friend Recommendation System for Social Networks", IEEE TRANSACTIONS ON MOBILE COMPUTING, VOL. 14, NO. 3, MARCH 2015.
- [2] J. Biagioni, T. Gerlich, T. Merrifield, and J. Eriksson, "EasyTracker: Automatic transit tracking, mapping, and arrival time prediction using Smartphones," in Proc. 9th ACM Conf. Embedded Netw. Sensor Syst., 2011, pp. 68–81.
- [3] P. Desikan, N. Pathak, and J. Srivastava, V. Kumar, "Incremental page rank computation on evolving graphs," in Proc. Special Interest Tracks Posters 14th Int. Conf. World Wide Web, 2005, pp. 1094–1095.
- [4] K. Farrahi and D. Gatica-Perez, "Probabilistic mining of sociogeographic routines from mobile phone data," IEEE J. Select. Topics Signal Process., vol. 4, no. 4, pp. 746–755, Aug. 2010.
- [5] Yu, A. Pan, L.-A. Tang, Z. Li, and J. Han, "Geo-friends recommendation in GPS-based cyber-physical social network," in Proc. Int. Conf. Adv. Social Netw. Anal. Mining, 2011, pp. 361–368.
- [6] D. Sarma, A. R. Molla, G. Pandurangan, and E. Upfal, Fast Distributed Pagerank Computation. Berlin, Germany: Springer, pp. 11–26, 2013
- [7] Y. Zheng, Y. Chen, Q. Li, X. Xie, and W.-Y. Ma, "Understanding transportation modes based on GPS data for web applications," ACM Trans. Web, vol. 4, no. 1, pp. 1–36, 2010.
- [8] G. R. Arce. Nonlinear Signal Processing: A Statistical Approach. John Wiley & Sons, 2005.
- [9] L. Bian and H. Holtzman. Online friend recommendation through personality matching and collaborative filtering. Proc. of UBICOMM, pages 230-235, 2011.
- [10] M. Bishop. Pattern recognition and machine learning. Springer New York, 2006.
- [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] 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.
- [13] Amazon. <http://www.amazon.com/>.
- [14] Facebook statistics. <http://www.digitalbuzzblog.com/facebook-statistics-stats-facts-2011/>.
- [15] Netflix. <https://signup.netflix.com/>.
- [16] Rotten tomatoes. <http://www.rottentomatoes.com/>.