

# A Temporal Probabilistic Model for Dynamic Circle Recommendation in Mobile Applications

Fan-Kai Chou

Department of Computer Science  
National Chiao Tung University  
Hsinchu, Taiwan  
plapla.cs00g@nctu.edu.tw

Meng-Fen Chiang

Department of Computer Science  
National Chiao Tung University  
Hsinchu, Taiwan  
mfchiang.cs95g@nctu.edu.tw

Wen-Chih Peng

Department of Computer Science  
National Chiao Tung University  
Hsinchu, Taiwan  
wcpeng@cs.nctu.edu.tw

**Abstract**—This paper presents a novel framework for dynamic circle recommendation for a query user at a given time point from historical communication logs. We identify the fundamental factors that govern interactions and aim to automatically form friend circles for scenarios, such as, *who should I share the photo with in the early morning? Whose post should be listed on top of my Facebook Wall feed at night?* We develop a temporal probabilistic model that not only captures temporal tendencies between the query user and each friend candidate but also blends frequency and recency into circle formation. Experimental results on Enron dataset and Call Detail Records prove the effectiveness of dynamic circle formation with proposed temporal probabilistic model.

**Keywords**—Dynamic Circle, Time-Dependency, Mobile Social Network

## I. INTRODUCTION

As the emergence of on-line social networking services, users can easily share information to their friend circles via Pinterest, Tumblr, Facebook, etc with their mobile devices. Social networking services gather and syndicate these information feed to target users. Users can browse through the information feed shared by their friends. Most existing social network services generally render information feed based on recency, that is, latest feeds always appear on top of personal feed walls. Some may provide manual tools for users to explicitly adjust friend circles so that users can control how information feed are rendered on their walls or which friend circles to share information with. Such great efforts motivate us to ask one question: Is it possible to design a dynamic circle recommendation system which can automatically suggest a ranked list of friend candidates driven by both historical interaction statistics and contextual information such as time point.

The general problem of recommendation system has been widely studied [6]. Recently, several prior studies attempt to consider temporal factor in designing recommendation systems [8][3][2][12]. For example, [8] leveraged user's long-term and short-term preference for temporal recommendation. Nonetheless, non of them addresses the fact that user interactions are not always correlate with time as users present diverse variation of temporal dependency. For example, it is quite clear that some users have higher temporal dependency in sharing information. Moreover, a user may only be sensitive to certain time points during a day. In this paper, we argue that temporal tendency should be analyzed individually for each

pair of query user and friend candidate at each time point. As an evidence, Figure 2 illustrates a distribution of temporal centrality for all pairs of users. If a pair of users' interactions only fall into a few time slots during a day, they has lower entropy and thus indicating higher temporal centrality and vice versa. We observe that over 60% pairs of users' have higher time centrality in interactions (entropy  $\leq 0.5$ ), meaning the rest 40% user interactions are driven or dominated by other factors.

In this paper, we propose a framework to discover personalized dynamic circles at different time points. Specifically, given a query user, a time point, and historical communication logs, our recommendation system returns a dynamic circle, represented by a ranked list of contact users, for the query user at given time point. To achieve this, we propose a temporal probabilistic model to capture user behaviors based on three factors: frequency, recency and time-dependency. After this, we utilize the temporal probabilistic model to support two types of circle recommendation: **Seedset Generation**: single-interaction suggestion and **Friends Suggestion**: multiple interactions suggestion. The temporal probabilistic model considers the dynamic importance of each candidate user for a query user to incorporate the factor, *different users show different temporal dependency with related to a target user*. Seedset Generation enables to generate a single candidate user automatically by proposed temporal probabilistic model for two purposes: shifting the burden of query users to provide a list of users who intent to interact with at the very beginning, and the query user merely intends to interact with a single user at given time. On the other hand, Friend Suggestion aims to provide a group of friends whenever the query user intends to interact with multiple users at the same time.

Recommending dynamic circle is useful in many applications. For example, dynamic circles can be utilized to enhance the ranking results for content-based on-line social networking services (e.g., Pinterest, Instagram, Facebook), where the information feed for each user can be adjusted based on the dynamic circle. Moreover, it can be used in location sharing services (e.g., Foursquare), where the ranking of locations can be adjusted based on a user's dynamic circle at particular time point.

Our contributions are as follows.

- We propose a framework to discover personalized dynamic circles for a query user at given time point.

- We propose a temporal probabilistic model to capture user’s interaction tendency at different time point.
- We identify three fundamental factors in user interactions and propose approaches to support: single-interaction suggestion and multiple interactions suggestion by integrating the temporal probability model into the-state-of-the-art ranking model.
- We conduct experiments on real datasets to demonstrate the effectiveness of our framework and report empirical insights.

This paper is organized as follows. Section II introduces our temporal probabilistic model and then discusses the two types of dynamic circle recommendation system. Section III shows the experimental results using the two real datasets. Section IV presents the related work for this paper. Section V concludes this paper.

## II. DYNAMIC CIRCLE RECOMMENDATION SYSTEM

Existing friends suggestion approach [5] heavily relies on query users to provide partial contact users (i.e., a seed set) to derive a ranked list of contact users. This increases inconvenience for query users, especially for mobile users, to efficiently communicate with multiple users at the same time. Motivated by this, we propose a framework for automatic dynamic circle recommendation without letting query users to provide prior information such as core contact users. The system framework overview is illustrated in Fig. 1. Our system consists of two phases: Seedset Generation and Friends Suggestion. Seedset Generation automatically derives a set of core users (referred to as seedset) with the highest probability to be contacted with the query user. Seedset Generation is achieved by mining frequent and time-dependent communication patterns from historical interaction logs. In Friends Suggestion phase, our system generates a ranked list of contact users based on the derived seedset. Once the query user chooses partial members from the list, our system updates the friend suggestion list by adding selected users to current seedset and then launching Friends Suggestion again to update the ranked list of contact users. This process continues until no more contact users can be suggested or the query user drops this session. Notice that our recommendation system provides a generic framework where the Friend Suggestion component can be replaced by other state-of-the-art algorithms to serve different requirements.

### A. Temporal Probabilistic Model

When a query user attempts to share information (e.g., photos), the query user forms a list of contact users in his/her mind. After that, query users has to manually and sequentially select the list of contact users by scanning through their friends corpus. This brings lots of unnecessary efforts. To solve this problem, we first propose Seedset Generation that uses a temporal probabilistic model to predict possible contact users as seeds for Friends Suggestion.

We claim that *if a query user interacts with a contact user in a similar time interval, this infers that the query user has*

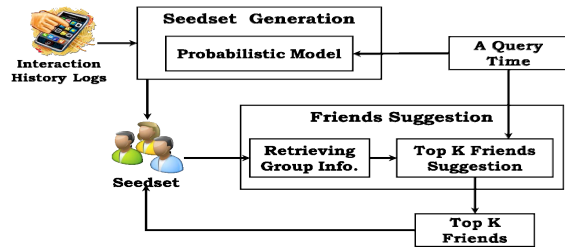


Fig. 1: Framework Overview

a higher probability to interact with the contact users in a similar time interval as well. Fig. 2 shows the distribution of entropy for every contact user, where the x-axis is entropy and the y-axis is number of contact users. We can clearly see that there are three peaks in Fig. 2, which can be explained by the fact that the lower entropy has higher time centrality and vice versa. The first peak with entropy = 0 represents that their interactions centralize in one time slot, the second peak with entropy = 0.5 means the time points of interactions distributing in two time slots and the third peak with entropy = 0.75 is distributed in three slots. We observe that the number of contact users is over 60% when entropy is no greater than 0.5, which means most of the contact users have higher time centrality for interactions. Therefore, this verifies our assumption that most interactions is highly time-dependent.

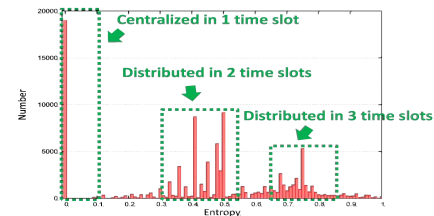


Fig. 2: Distribution of Time Centrality

Our goal is to shift away mobile user’s burden whenever they attempt to share information via their mobile devices. Specifically, we designed a temporal probabilistic model to predict a ranked list of contact users who are most likely to be interacted with the query user a given time point. There are three factors considered in developing our temporal probabilistic model:

- 1) Frequency: Receivers who have more interactions with the query user are more important than those who interact less with the query user.
- 2) Recency: More recent interactions should have more importance whereas older interactions decay over time.
- 3) Time-Dependency: If receivers always interact with the query user in a similar time interval, they should have more importance in that time interval.

*Frequency* is a straightforward yet effective measurement. Inspired from Interaction Rank [5], we unify *Frequency* with *Recency* into a single measurement as shown in Equation (1). [5] introduced a decaying parameter  $\lambda$ , to control the importance of every interaction according to its time. Namely,

every interaction decays exponentially over time with a half life  $\lambda$ . To fit the temporal probabilistic model, we form the two factors into a probability, which can be expressed as:

$$P(R_n) = \frac{\sum_{i \in I(R_n)} (\frac{1}{2})^d}{\sum_{i \in I} (\frac{1}{2})^d} \quad (1)$$

where  $P(R_n)$  is the probability of the query user interacting with  $R_n$  in the past,  $I$  is a set of all the query user's interactions, and  $I(R_n)$  is a set of all interactions between query user and  $R_n$ .  $d$  is a decay function which is expressed as  $\frac{t_{now} - t_i}{\lambda}$ , where  $t_{now}$  is the current time,  $t_i$  is the time of interaction  $i \in I$ , and a half-life parameter  $\lambda$  that assigns score 1 to an interaction at current time and decays the importance of an interaction to  $\frac{1}{2}$  with the half-life  $\lambda$ .

To incorporate the third factor, *Time-Dependency*, we formulate a conditional probability as:

$$P(R_n|t) = \frac{P(R_n \cap t)}{P(t)}. \quad (2)$$

Equation (2) shows the probability of the query user interacting with  $R_n$  in a time interval  $t$ , where  $P(R_n \cap t)$  and  $P(t)$  can be derived by the following equations:

$$P(R_n \cap t) = \frac{\sum_{i \in (I(R_n) \cap I(t))} (\frac{1}{2})^d}{\sum_{i \in I} (\frac{1}{2})^d} \quad (3)$$

and

$$P(t) = \frac{\sum_{i \in I(t)} (\frac{1}{2})^d}{\sum_{i \in I} (\frac{1}{2})^d} \quad (4)$$

where  $I(t)$  is a set of the query user's interactions in time interval  $t$ .

To take into the following three factors into consideration, *Frequency*, *Recency* and *Time-Dependency*. Intuitively, we combine  $P(R_n)$  and  $P(R_n|t)$  by a linear combination with a tunable parameter  $\alpha$ , which can be formulated as follows:

$$Score(R_n) = (1 - \alpha)P(R_n) + \alpha \cdot P(R_n|t) \quad (5)$$

where  $\alpha$  is the weight of *Time-Dependency* and the range of  $\alpha$  is between 0 and 1. In general, Equation (5) does not make sense, because when a candidate receiver  $R_n$  has higher  $P(R_n)$  and also has higher  $P(R_n|t)$ , it should be chosen with more chances. When both probabilities are not relative to each other, we should think about other methods to merge them. Calculating the mean between  $P(R_n)$  and  $P(R_n|t)$  is a good idea to balance Equation (5), since it considers the influence from not only specific time intervals but also all time intervals. We adjust Equation (5) by using two types of mean, arithmetic mean and geometric mean, and thus both equations can be expressed as follows:

$$Score_{ari}(R_n) = (1 - \alpha)P(R_n) + \alpha \cdot \frac{P(R_n) + \omega \cdot P(R_n|t)}{1 + \omega} \quad (6)$$

and

$$Score_{geo}(R_n) = (1 - \alpha)P(R_n) + \alpha \cdot \sqrt[1+\omega]{P(R_n)(P(R_n|t))^\omega} \quad (7)$$

where  $\omega$  represents the weight of a specific time interval. Notice that we only use geometric mean in our experiment since we find that geometric mean makes more sense for our

assumption: if one of  $P(R_n)$  and  $P(R_n|t)$  is much lower than the other, their mean should be closer to the lower one.

For refining Equation (7), we want to decide which value of  $\alpha$  is the best by observing data. In our observations, we find that not all receivers have high time-dependency, as some are independent of a specific time interval. In other words, receivers have different time-dependencies in different time intervals, and time-dependencies will vary from person to person. To achieve this, we change  $\alpha$  to another conditional probability which is expressed by:

$$P(t|R_n) = \frac{P(R_n \cap t)}{P(R_n)}. \quad (8)$$

Equation (8) indicates the probability of  $R_n$  interacting with the query user in time interval  $t$ . If  $P(t|R_n)$  is higher,  $R_n$  has a higher time-dependency with the query user and vice versa. We then utilize Z-score to make the importance of time-dependency be numerical, the formulation of Z-score is

$$Z(R_n) = \frac{P(t|R_n) - avg(P(t|R))}{\sigma(P(t|R))} \quad (9)$$

where  $avg(P(t|R))$  is the average probability of all  $R_n \in R$  and  $\sigma(P(t|R))$  is the standard deviation. Since Z-score may be negative, we normalize Z-score by considering the central point, namely, we normalize Z score within  $[-3, 3]$  to  $[0, 1]$ . If the Z-score is more than 3 or less than -3, it will be considered as 1.0 or 0.0. Therefore, we can reformulate Equation (7) as follows:

$$Score_{final}(R_n) = (1 - NZ(R_n)) \cdot P(R_n) + NZ(R_n) \cdot \sqrt[1+\omega]{P(R_n)(P(R_n|t))^\omega} \quad (10)$$

where  $NZ(R_n)$  is the normalized Z-score and the range is from 0 to 1.

This temporal probabilistic model has been used in Seedset Generation and we will discuss how to use it in the next subsection.

## B. Seedset Generation

Seedset Generation phase derives a set of core contact users who are highly to be the receivers with related to the query user at given time. In a sense, Seedset Generation can serve as a Friends Suggestion in a special case when query users intend to communicate with a single user instead of a group of users. In that case, Seedset Generation phase returns the potential receivers as a top- $k$  list of users. Without specific groups information, Seedset Generation adopts the introduced temporal probabilistic model mentioned before predicting which contact users in the past are most likely the receivers, merely based on specified query time. The algorithm of Seedset Generation can be summarized in Algorithm 1.

## C. Friends Suggestion

Friends Suggestion is a function which can be applied to any seed-based suggestion approach. In this subsection, we propose an enhanced suggestion approach, Friends Suggestion, with our temporal probabilistic model into state-of-the-art ranking models [5].

---

**Algorithm 1:** Seedset Generation Algorithm

---

**Input:** Query user’s history interactions  $I$  and current time interval  $t$

**Output:** The seed generations  $S$

```
1  $S = \phi$ ;  
2 foreach  $i \in I$  do  
3   Sum scores of  $i$  for temporal probabilistic model;  
4    $C = \text{GetContactUser}(i)$ ;  
5   foreach  $c \in C$  do  
6     if  $c \notin S$  then  
7       Put  $c$  into  $S$ ;  
8 foreach  $c \in S$  do  
9   Calculate all probabilities  $P(c)$ ,  $P(t)$ ,  $P(c|t)$  and  
    $P(t|c)$ ;  
10   $S[c] = \text{Score}_{final}(c)$ ;
```

---

Our temporal probabilistic model can be combined with *Gmail Approach* [5] which Roth *et al.* proposed to become an enhanced suggestion approach. They considered three factors, *Frequency*, *Recency* and *Direction*, and we consider one additional factor, *Time-Dependency*. Their score function named *Interaction Rank* is:

$$\mathcal{IR}(g) = \theta_{out} \sum_{i \in I_{out}(g)} \left(\frac{1}{2}\right)^d + \sum_{i \in I_{in}(g)} \left(\frac{1}{2}\right)^d \quad (11)$$

where  $I_{out}(g)$  is the set of outgoing interactions between a query user and a group,  $I_{in}(g)$  is the set of incoming interactions and  $\theta_{out}$  is the weight of outgoing interactions to represent *Direction*. According to Equation (11), we can modify Equation (1) as the following equation:

$$P(R_n) = \frac{\theta_{out} \sum_{i \in I_{out}(R_n)} \left(\frac{1}{2}\right)^d + \sum_{i \in I_{in}(R_n)} \left(\frac{1}{2}\right)^d}{\theta_{out} \sum_{i \in I_{out}} \left(\frac{1}{2}\right)^d + \sum_{i \in I_{in}} \left(\frac{1}{2}\right)^d}. \quad (12)$$

Equations (3) and (4) have similar transformed equations, too. The scoring function in *Gmail Approach* we adopt is *Intersection Weighed Score*, which considers the intersection of group and seedset and uses them to weight the score of the group. As reported in [5], *Interaction Weighted Score* achieves the best performance among their proposals. We compare *Gmail Approach* and our approach in Section III-D.

### III. EXPERIMENT

In this section, we first verify that *Time-Dependency* is useful for recommendation, and then evaluate the quality of proposed framework: Seedset Generation and Friends Suggestion.

#### A. Data Description

In our experiment, we use two real datasets, Enron Mail<sup>1</sup> and call detail records(cdr) from Chunghwa Telecom(CHT)<sup>2</sup>. Both datasets’ basic information are shown in Table I, where we consider Enron Mail as group-based data and CHT as

---

<sup>1</sup>The Enron Mail data can be downloaded from <http://www.cs.cmu.edu/~enron/>.

<sup>2</sup>The CHT data is not in public, and Chunghwa Telecom’s website is <http://www.cht.com.tw/>.

TABLE I: Basic Information on the Enron/CHT Datasets

element	Enron	CHT
#user	65,182	76,263
#interactions	236,505	2,443,667
#group interactions	67,631	-
time	1998/01/04 - 2002/12/21	2010/08

TABLE II: Four slots of Time in 24 hours

Part I	Part II	Part III	Part IV
00:00 - 05:59	06:00 - 11:59	12:00 - 17:59	18:00 - 23:59

single-based data. In other words, we adopt Enron Mail in both Seedset Generation and Friends Suggestion evaluations, and CHT only in the Seedset Generation evaluation. We also adopt Enron Mail in *Time-Dependency* evaluation, which is discussed in the next subsection.

#### B. Experimental Setup

In time-dependency experiment, we set some constraints for choosing test query users and their contact users. There are 9,036 test query users and they have totally 102,405 contact users. The constraints are that test query users should participate in at least 10 groups and the contact users should interact with test query users at least 4 times, since we split the time in one day into four parts which is shown in Table II and 4 times could achieve the maximum entropy for four parts exactly. For Enron Mail, we chose 21,262 mails from Enron Mail to be the testing data and the others to be the training data, where the rule for choosing the testing data is that the mail should be sent to at least 2 receivers, or a *group*, and the sender of the mail had sent no less than 4 mails before. For CHT cdr, which is a single-based data, they do not have group information, since CHT cdr consists of cell phone call records and we only need to guess one person that the query user wants to call. We chose 30,295 cdr from CHT cdr to be the testing data and the others to be the training data. The testing data is all in the last day in CHT cdr.

TABLE III: Parameter Settings

Parameter	Enron	CHT
$\lambda$	7 days	3 days
$\theta_{out}$	5	-
$\omega$	1	1
<i>Time Interval</i>	1 hour	1 hour

Some parameter settings we use are shown in Table III, where  $\lambda$  is a time decay parameter,  $\theta$  is an outlink weight parameter,  $\omega$  is a time dependency parameter and *Time Interval* is how many additional hours we reference next to the current hour.

#### C. Metrics

For the following experiments, we adopt normalized discounted cumulative gain(nDCG) as the measurement to evaluate. DCG measures the *gain* of a hit result based on its rank in the list, where the top rank has more gain and the lower

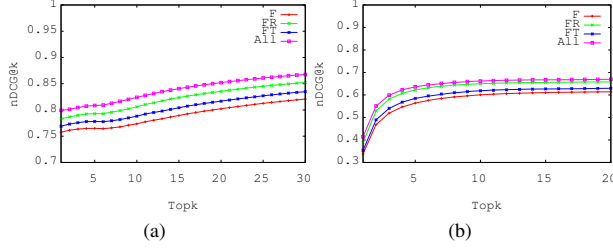


Fig. 3: Comparison for Seedset Generation in (a) Enron Mail (b) CHT cdr.

rank has less gain, although it is also hit. The formulation of DCG we use is

$$DCG_p = rel_1 + \sum_{i=2}^p \frac{rel_i}{\log_2(i)} \quad (13)$$

where  $rel_i \in \{0, 1\}$  is a relevance value and  $p$  is the length of the result list. nDCG is to normalize DCG so that we can compare result lists with different lengths. The formulation of nDCG is expressed by

$$nDCG_p = \frac{DCG_p}{IDCG_p} \quad (14)$$

where IDCG is ideal DCG which is calculated by the ideal rank of the result list.

TABLE IV: Four Methods for Comparison

Method	Symbol
Frequency	F
Frequency + Recency	FR
Frequency + Time-Dependency	FT
All factors	All

#### D. Results

**Evaluation on Seedset Generation:** Figure 3a shows the impact of each fundamental factor: Frequency, Recency and Time-dependency on Seedset Generation quality. We want to verify that our temporal probabilistic model is effective, so we separate the three factors into four methods for Seedset Generation and compare them. Each method is summarized and denoted in Table IV. In Fig. 3a, the x-axis is top- $k$  ( $1 \leq k \leq 30$ ) and the y-axis is the value of nDCG. The red line (F) that only considers frequency to determine the friend circle is considered as a baseline. The pink line (All) is our proposal which considers all factors and outperforms other models with at most 4.2% increase in accuracy compare to the baseline.

In Fig. 3b, the x-axis is the top- $k$ , ranging from 1 to 20, and the y-axis is the nDCG value. It is worth mentioning that the lines assemble the log-likelihood, because CHT cdr only has one receiver for recommendation in each record. Similar results can be found on CHT dataset. Therefore, we conclude that our temporal probabilistic model presents consistent enhancement than straightforward suggestion such as frequency or recency.

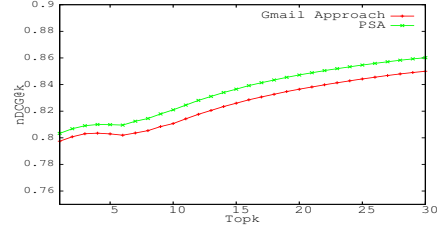


Fig. 4: Comparison of Gmail Approach and Probabilistic Suggestion Approach (PSA) for Friends Suggestion

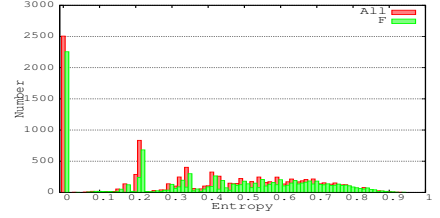


Fig. 5: Comparison of Entropy Distribution

**Evaluation on Friends Suggestion:** Fig. 4 shows the performance comparison of Gmail Approach [5] and our proposed Probabilistic Suggestion Approach (PSA). Since Gmail Approach is a seed-based suggestion approach, we use Seedset Generation to generate a seedset with  $k = 3$  and pass the input to Gmail and PSA, respectively. The final recommendation list contains the seeds which are different from the original Gmail Approach, but it will not affect the recommendation result because the seeds appears at the top of the list and they are also uncertain receivers for the query user. We use the same test data from Enron Mail as in Fig. 3a. In Fig. 4, the x-axis is top- $k$  ( $1 \leq k \leq 30$ ) and the y-axis is the nDCG value, where the red line is Gmail Approach and the green line is PSA. We can see that no matter in what situation, PSA always has higher nDCG than Gmail Approach.

**Analysis of Time-Dependency:** Finally, we provide an empirical insight of proposed framework. We choose the hit receivers, which are the correct answers for the test letters, from two different methods *All factors* and *Frequency* to calculate their time centrality. The approach for calculating time centrality is the same as Fig. 2. Fig. 5 shows that *All factors* can recommend more correct contact users because they shows higher time centrality (i.e., having more receivers whose entropy equal or less than 0.5).

## IV. RELATED WORK

Collaborative Filtering (CF) is a traditional research of recommendation system, which references the behaviors of a group of users to recommend or predict for other users. There is a survey paper [6] discussing many kinds of CF. In our opinion, modern research of recommendation systems is interested in finding new features to recommend those that have never been recommended before, especially temporal features. For example, Xiang *et al.* [8] proposed the Session-based Temporal Graph (STG) to describe user

behaviors, which are determined by long-term and short-term preferences. Long-term preferences could be considered as frequency, and short-term preferences could be considered as the relations between user behavior and time interval. Lathia *et al.* [3] proved that temporal diversity is an important feature for recommendation, since the user will rate items over time. Zheng and Li [12] provided a resource-recommendation model, which integrates both tag and tagging time information then uses CF for recommendation. The time information has a similar definition to *Recency* in this paper. Koren [2] built a time-factor model which is similar to [8], because the model considered both long-term and transient effects. Although they considered temporal features, they still not consider that user behaviors are not always related with time. There are some recommendation system studies which have tried to find other features, such as Trust Circles [11], Bayesian Network [9], Social Collaborative Filtering(SCF) [4], and the Nearest Neighbor based Top-K recommendation system [10].

Friends suggestion is sometimes similar to relationship link prediction, because if we predict when a link would be built we guess that both linked nodes will possibly communicate in that time. Sun *et al.* [7] built a relationship building time prediction model, which uses learning algorithms to fit different distributions and then gets a probability for building relationships between two nodes. But the edges are only constructed once, so we cannot use it for communication networks which change over time.

Our main idea is based on Roth *et al.* [5], who proposed a friends recommendation system for *Gmail* using group information and three criteria. *Gmail* is a well-known mail system constructed by *Google*, which may have many history records to retrieve for friends suggestion. However, the algorithm in [5] could not work effectively for sparse data, insufficient interaction history resulting in some recommendation lists to be empty. Moreover, *Time-Dependency* of user interactions is not addressed in their work. Bartel and Dewan [1] enhanced [5] with a hierarchical structure, which re-orders the recommendation list by ranking past communication group and hierarchically predicts next group. Similarly, no attention has been paid to address the problem of *Time-Dependency* of user interactions.

## V. CONCLUSION

In this paper, we study the problem of suggesting friends by implicit social graph and temporal importance. How to combine all the features we mentioned poses great challenges. In this paper, we propose a temporal probabilistic model to combine three factors, *Frequency*, *Recency* and *Time-Dependency*, and give them a dynamic weight which considers the problem, *different contact users have different importance of time for a query user*. Based on the temporal probabilistic model, we have designed a dynamic circle recommendation system which has two functions, Seedset Generation and Suggestion Approach. Seedset Generation can generate a set of seeds automatically by adopting the temporal probabilistic model and let the set of seeds be input to the next function. Suggestion Approach is flexible for embedding any seed-based suggestion approach, so we constructed our suggestion approach, the

Probabilistic Suggestion Approach(PSA), which applies the temporal probabilistic model and considers an additional feature, *Direction* of interactions.

We verify that *Time-Dependency* is an useful feature for suggesting friends and our experiment results also verify it. Our experiment results shows that our proposed temporal probabilistic model and dynamic circle recommendation system are effective in two real datasets, which are Enron Mail and CHT call detail records. Although we have evaluated our temporal probabilistic model using emails and phone calls, it can be applied to any interaction-based social network. Future work will include the extension of our temporal probabilistic model such as dynamically deciding the number of seeds and using user clusters. We are also interested in exploring other applications such as content-based sharing and temporal community detection.

## ACKNOWLEDGMENT

Wen-Chih Peng was supported in part by the National Science Council, Project No. 100-2218-E-009-016-MY3 and 100-2218-E-009-013-MY3, by Taiwan MoE ATU Program, by Academic Sinica, Project No. AS-102-TP-A06, and by HTC.

## REFERENCES

- [1] J. Bartel and P. Dewan. Towards hierarchical email recipient prediction.
- [2] Y. Koren. Collaborative filtering with temporal dynamics. *Communications of the ACM*, 53(4):89–97, 2010.
- [3] N. Lathia, S. Hailes, L. Capra, and X. Amatriain. Temporal diversity in recommender systems. In *Proceedings of the 33rd Annual International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR10)*, pages 210–217, 2010.
- [4] J. Noel, S. Sanner, K.N. Tran, P. Christen, L. Xie, E.V. Bonilla, E. Abbasnejad, and N. Della Penna. New objective functions for social collaborative filtering. In *Proceedings of the 21st international conference on World Wide Web*, pages 859–868. ACM, 2012.
- [5] M. Roth, A. Ben-David, D. Deutscher, G. Flysher, I. Horn, A. Leichtberg, N. Leiser, Y. Matias, and R. Merom. Suggesting friends using the implicit social graph. In *Proceedings of the 16th ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 233–242. ACM, 2010.
- [6] X. Su and T.M. Khoshgoftaar. A survey of collaborative filtering techniques. *Advances in Artificial Intelligence*, 2009:4, 2009.
- [7] Y. Sun, J. Han, C.C. Aggarwal, and N.V. Chawla. When will it happen?: relationship prediction in heterogeneous information networks. In *Proceedings of the fifth ACM international conference on Web search and data mining*, pages 663–672. ACM, 2012.
- [8] L. Xiang, Q. Yuan, S. Zhao, L. Chen, X. Zhang, Q. Yang, and J. Sun. Temporal recommendation on graphs via long-and short-term preference fusion. In *Proceedings of the 16th ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 723–732. ACM, 2010.
- [9] X. Yang, Y. Guo, and Y. Liu. Bayesian-inference based recommendation in online social networks. In *INFOCOM, 2011 Proceedings IEEE*, pages 551–555. IEEE, 2011.
- [10] X. Yang, H. Steck, Y. Guo, and Y. Liu. On top-k recommendation using social networks. In *Proceedings of the sixth ACM conference on Recommender systems*, pages 67–74. ACM, 2012.
- [11] X. Yang, H. Steck, and Y. Liu. Circle-based recommendation in online social networks. In *Proceedings of the 18th ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 1267–1275. ACM, 2012.
- [12] N. Zheng and Q. Li. A recommender system based on tag and time information for social tagging systems. *Expert Systems with Applications*, 38(4):4575–4587, 2011.