Decision Support Systems xxx (2015) xxx–xxx

Contents lists available at ScienceDirect

Decision Support Systems

journal homepage: www.elsevier.com/locate/dss

Proposing a new friend recommendation method, FRUTAI, to enhance social media providers' performance

Zhou Zhang [a], Yuewen Liu [a,*], Wei Ding [b], Wei (Wayne) Huang [c], Qin Su [a], Ping Chen [b]

[a] School of Management, Xi'an Jiaotong University, Xi'an, 710049, China
[b] Department of Computer Science, University of Massachusetts, Boston, MA, 02125, USA
[c] Department of MIS, College of Business, Ohio University, Athens, OH, USA

A R T I C L E   I N F O

Article history:
Received 4 April 2014
Received in revised form 8 June 2015
Accepted 20 July 2015
Available online xxxx

Keywords:
Social media
Common neighbors
Friend recommendation
Users' attributes

A B S T R A C T

Social media, such as Facebook and Twitter, have grown rapidly in recent years. Friend recommendation systems, as an important emerging component of social media, may efficiently expand social media networks by proactively recommending new and potentially high-quality friends to users. Literature review has shown that prior research work on friend recommendation mainly focuses on the linking relation between users in social media but largely neglects the influence of users' attributes. In this study, we have systematically reviewed and evaluated the existing state-of-the-art friend recommendation algorithms. We introduce a new Friend Recommendation system using a User's Total Attributes Information (FRUTAI) based on the law of total probability. The proposed method can be easily extended according to the increasing number of a user's attributes with low computation cost. Furthermore, the FRUTAI is a universal friend recommendation method and can be applied in different types of social media because it does not distinguish the structure of the network. We have collected 7 million users' public information and their friend relationships from RenRen, commonly regarded as the Facebook of China. Using the real-world data from a dominant social media provider, we extensively evaluate the proposed method with other existing friend recommendation algorithms. Our experimental results have demonstrated the comparatively better performance of FRUTAI. In our empirical studies, we have observed that the performance of FRUTAI is related to the number of a user's friends. In particular, when a user has a small number of friends, the proposed FRUTAI algorithm performs better than other algorithms; when a user has a large number of friends, the overall performance of FRUTAI becomes less competitive but is still comparable to those of other providers, and its precision rate is quite outstanding. Our findings may provide some important practical implications to social media design and performance.

1. Introduction

Research on social media has become more important, attracting research from scholars of different business disciplines, such as marketing (e.g., Kumar et al., 2013), strategy (Bharadwaj et al., 2013), human resources (Urban and Boscolo 2013), finance (e.g., Røssvoll and Fritsch 2013), IS (e.g., Aral, et al., 2013), healthcare (e.g., Coustasse and Slack 2013; Lin and Vaska 2013; Yang and Yang 2013), and the public sector (Davies and Cairncross 2013; Kolb and Roberts 2013) [25–34]. Social network services, such as Facebook and Twitter in the U.S.A., have grown rapidly with innovative systems and tools in recent years. High-quality friend recommendation is crucial to the survival and growth of those social media services. At the early stage of social media, a network is small with only a limited number of users with an accountable number of friends; it is easy to browse over all or many of other users' profiles to make decisions of whether to choose some users as friends. Currently, the number of social media users has reached a very high level. In 2013, the number of users from Facebook reached 1.19 billion worldwide. It seems infeasible for a user to browse over millions of other users' homepages to make a decision of whether to choose a potential friend. To meet this new challenge, social media providers began to design friend recommendation systems, such as the “People You May Know” system on Facebook and other similar recommendation services from Twitter, which may assist users to make better decisions [18].

There is a stream of literature that focuses on the recommending models, named link prediction models [17]. These link prediction models are useful to predict the extent of the network by observed data and play a role as a basic question in social media structure. The possibility of connection also reflects the “quality of connection” between two users in the future. If there is a high possibility that a tie

* Corresponding author.
E-mail addresses: zhouzhang@stu.xjtu.edu.cn (Z. Zhang), liuyuewen@mail.xjtu.edu.cn (Y. Liu), ding@cs.umb.edu (W. Ding), huangw@ohio.edu (W.(W.) Huang), qinsu@mail.xjtu.edu.cn (Q. Su), Ping.Chen@umb.edu (P. Chen).

http://dx.doi.org/10.1016/j.dss.2015.07.008
0167-9236/© 2015 Published by Elsevier B.V.

will connect two users, this connection will be a strong tie, which means more similarity between them. The research of link prediction has both theoretical and practical values.

Existing friend recommendation methods and algorithms are, in principle, based on two different approaches—a path-based method and a friend-of-friend method [16,18]. The path-based method uses friend linkage information by implementing the concept of the well-known PageRank algorithm from Google. Due to its high computational cost, this type of algorithm is seldom used in commercial social media. The friend-of-friend (FoF) method is an efficient and widely used recommendation algorithm due to its low time complexity. The algorithm identifies potential but unlinked friends and makes recommendations. Existing FoF algorithms mainly focus on the relations between users, but overlook the users' attributes.

In this study, we systematically review and evaluate the existing state-of-the-art friend recommendation algorithms to discuss their strengths and weaknesses. We then propose a new Friend Recommendation method with a User's Total Attributes Information (FRUTAI). The proposed new FRUTAI method can help social media service providers provide a better decision-making tool for its users to choose high-quality or more preferable friends online and assist users to choose more relevant and preferable friends. This paper is such an initial research effort to integrate social media users' attributes with the law of total probability. Prior systems are largely designed for specific types of social media networks, which may not be effective to different structures of networks. FRUTAI is a generic friend recommendation method and can be applied in different types of social media. It can be extended to accommodate new set of user attributes as well.

The rest of this paper is organized as follows. Section 2 gives a brief literature review of the main existing recommendation algorithms. Section 3 presents the methodology of our proposed new algorithm. Section 4 presents a real-world case study using the proposed algorithm. Section 5 concludes the paper by discussing its potential implications to future research and practice.

2. Related literature review

2.1. Homophily and heterophily in relationships

Social media is structured by users and the ties between them. These ties reflect all of the types of relationships, such as friendship, kinship, marriage, working relationships, teacher–student relationships, and so on. The studies of network ties began in the 1920s and lasted nearly 100 years [40]. Homophily and heterophily are two principles that significantly influence the contact between users in social media. The homophily principle holds that if two people have similar attributes, they will have a greater chance of having a relationship than other dissimilar people. In contrast, heterophily refers to the preference for the different attributes, which is the opposite of homophily [41].

Researchers who focus on relationships in social media have studied the major sociodemographic dimensions such as race, gender, age, location, and education. These dimensions are also important attributes of the users in social media [40].

Compared with other dimensions, the influence of gender on ties starts in childhood. Smith-Lovin, and McPherson present that the homophily exists in play patterns, and they also observed that girls play in smaller groups than boys [42]. Eder and Hallinan find that the youths prefer to delete a cross-sex friend than add a cross-sex friend, which leads to gender segregation in social media [43]. On the other hand, the networks of adults are more sex-integrated. Marsden explains that when people "discuss important matters with" the confidants, 70% of them are sex heterogeneous [44]. However, Huckfeldt and Sprague present that when the topic is limited to politics, 84% of men choose other men to discuss it [45].

Homophily in geography is obvious because it is easy for people to have more interactions with friends who live nearby than those who live far away. Kaufer and Carley study the influence of new technologies and find that they weaken the homophily of geography [46]. Likewise, Hampton and Wellman present that with virtual technology, the community does not have to be locally based as before [47].

2.2. Recommendation algorithms

In addition to the friendship studies, there is a stream of literature named "link prediction". Traditional study is based on surveys, but the dataset is limited. Currently, we have commercial social networks, with large datasets, which make it possible to investigate connections. As a result, we have the opportunity to investigate the connection problems from other perspectives.

Existing algorithms of recommendation systems can be classified into two broad categories: recommending items and recommending people.

The traditional algorithms for recommending people, such as FoF, use only the information of friend relations in social media and do not make full use of a user's attributes. On the other hand, the traditional algorithms for recommending item, such as a content-based method, care only about a user's own information and ignore the relations between users. As a result, in our study, we propose a new method to combine the two to improve friend recommendation performance.

2.2.1. Recommending items

Many prior research works focus on recommending items in social media (e.g., [1-3,21,22]), and there are two main methods. Content-based methods exploit the history information of a user's own attributes and make recommendations accordingly. Pazzani and Billsus define a content-based recommendation system [2]. For example, the basic idea is that if someone has bought a cookbook before, there is a great chance that she will buy another cookbook.

Collaborative filtering is another widely used algorithm in item recommendation. For example, it is based on the idea that if friends of a user all buy a cookbook, she may also buy the cookbook. Pazzani compares the collaborative filtering with the basic content-based method, and then proposes a model combining collaborative filtering and content-based algorithms [1].

Adomavicius and Tuzhilin present an overview of three recommendation approaches: content-based, collaborative, and hybrid methods. They analyze their advantages and limitations [3]. In a content-based method, every item is represented by a set of features, which are used to make comparisons with a user's attributes. Although features can be attached to text documents by using retrieval techniques, some other types of files still need to assign features manually, such as image, audio, and video files. Another limitation is that this system cannot distinguish the items that share a same set of features. Furthermore, this method is limited by the existing attributes of a user that are based upon the user's prior experience. For example, if a user has not purchased a cookbook before, the system will never recommend a cookbook to her. Additionally, if she is a new user who has few attributes, the system cannot recommend an accurate list of items. The collaborative method can easily address all type of files because the recommended list of items for one user is based on the information of other users' recommendations. In addition, the domain of recommended items is not limited to a user's prior preferences. The collaborative method also has some limitations. First, as with the content-based method, a new user with little information in her preferences cannot obtain a satisfactory recommendation list. Second, it will take a long time for a system to be able to recommend a new item because the recommendation will be provided only after an item is rated by a number of users. Several hybrid methods have been developed to combine the content-based and collaborative methods to address the weaknesses of the two methods to achieve a better recommendation result [21,22].

Please cite this article as: Z. Zhang, et al., Proposing a new friend recommendation method, FRUTAI, to enhance social media providers' performance, Decision Support Systems (2015), http://dx.doi.org/10.1016/j.dss.2015.07.008

2.2.2. Recommending people

Recommending friends is an important issue in social media. Research has shown that a quality friend recommendation service may enhance connections between users, as well as the user loyalty to a social media [24]. Different from recommending items, recommending people is relatively new in social media research, and there are relatively fewer literature papers being published in this field. Friend-of-friend (FoF) and path-based approaches are the two main methods.

(1) Friend-of-friend (FoF) method
The FoF algorithm draws from the assumption that if two users in a social media network share many common friends, they may have a greater chance of becoming friends in the near future. This algorithm is also called "Common-Neighbors". Newman designs an experiment and exploits the data of paper authors in two databases over a six-year period to provide evidence for the primary idea of FoF [4]. That research also shows the proportional relation between the probability of an author having new coauthors and the number of coauthors she already has. Jin et al. use an FoF algorithm as one of the three general principles to create a simple model that describes the growth of social media [5]. The friend recommendation system on Facebook, which gives a list of the "people you may know", is also based on the FoF approach [35]. Tencent, one of the most popular social media websites in China, also mentions in their official help file that its recommendation system of the product 'Quanzi' is based on the 'common neighbor' algorithm [36].
With the continuous growth of social media, the primary Common-Neighbors model has provided for several improved algorithms, such as the Jaccard coefficient and Adamic/Adar. To prove that some factors perform better in the link prediction problem, Adamic and Adar introduced a new algorithm to calculate the similarity of two actors by analyzing text, in-links, out-links, and mailing lists on the homepages of social media [6]. The number of common friends between two actors can be used to evaluate the similarity.
Preferential attachment is another well-known model to describe the expansion of social media. Barabasi and Albert explain that a social media expands when new actors join in, and these new actors link preferentially to the old actors who already have more links [7]. Barabasi et al. (2001) study an 8-year period database of co-authorship information to find evidence of preferential attachment in the evolution of social media [8].

(2) Path-based method
Differing from the neighbor-based FoF approach, calculating the shortest path is the basic idea of path-based methods. Katz predicts the probability by the sum of all paths between two nodes. The shorter paths have more contribution than the longer paths in the link prediction [9].
Brin and Page introduce the PageRank algorithm as a key component of the Google search engine. It weighs every element within a set by the link-in and link-out numbers, and then gives a rank of all of the elements [10], based on PageRank [11,12].
Jeh and Widom propose SimRank to measure the similarity of elements using the information of their relations. SimRank combines the features of FoF and the random walk algorithm, which is also used in the PageRank algorithm [13].
Yin proposes and evaluates a framework of LINKREC, which uses the information of network structure and actors' attributes, based on the random walk with the restart algorithm [14].

2.3. More relevant literature on friend-of-friend

This study is more relevant to FoF. Hence, further review of relevant literature on FoF is conducted here. The basic assumption of FoF is that if user A and user B share a large portion of common friends in their friend lists in a social media network, they may want to be friends too. We define $\Gamma(x)$ as the set of neighbors of x and $\Gamma(y)$ as the set of neighbors of y. The three basic algorithms based on FoF can be defined as follows.

(1) Common-Neighbors
For a particular user y in a friend recommendation list for user x, its rank in the list can be calculated by the number of friends that x and y share. It is the most widely used algorithm in commercial social media. It is believed that in Facebook and RenRen, Common-Neighbors is the main idea being used in their friend recommendation systems. Eq. (1) gives how a Common-Neighbors method calculates a friend score.

$$score(x,\ y) := |\Gamma(x) \cap \Gamma(y)| \tag{1}$$

(2) Jaccard's coefficient
Salton and McGill introduce a metric to calculate the probability for information retrieval [15]. If we take friends to be recommended as features to be retrieved from, this algorithm can be used in recommendation systems [15]. The score is given by the probability that a person randomly selecting from the union of the set of neighbors of x and the set of neighbors of y, is just the overlap of them, see Eq. (2).

$$\mathrm{score}(x, y) := \frac{|\Gamma(x) \cap \Gamma(y)|}{|\Gamma(x) \cup \Gamma(y)|} \tag{2}$$

(3) Adamic/Adar
Adamic and Adar summarize a metric to calculate the similarity of two users in a social media network [15]. They sum all of the same attributes shared by two users, and the unique attributes for an entire social media network weigh more than the common ones [6]. For example, if both student A and student B take a French class (30 students in total) and a dancing class (5 students in total), we can calculate the probability that they will become friends by the information of these two classes. And because there are fewer students in the dance class, it will have more influence on the probability that they will be friends in the future. For user x, the rank of y in the friend recommendation list can be given by this algorithm, if we change item into friend, see Eq. (3).

$$\mathrm{score}(x, y) := \sum\nolimits_{z \in \Gamma(x) \cap \Gamma(y)} \frac{1}{\log|\Gamma(z)|} \tag{3}$$

3. Research methodology

3.1. Probability theory

Traditional FoF algorithms mainly utilize the information of the number of users' friends. A potential problem of using additional users' attributes is the increased computational cost due to the large number of users and user groups. We propose FRUTAI (Friend Recommendation with a User's Total Attributes Information) to efficiently and effectively utilize additional information of users' attributes with time complexity comparable to the traditional FoF algorithms.

Definition 3.1. Probability with a User's Total Attributes Information

A is a user in a social media and C is a friend candidate with the attributes $x_i (i \in \{1, \cdots, m\})$. If each probability of C's finite or countably infinite attributes x_i in a social media network where C will be the friend of A is measurable, then the total probability that the candidate C will be the friend of the user A is defined as:

$$P(A) = \sum_{i=1}^{m} P(A/x_i)P(x_i) \tag{4}$$

Please cite this article as: Z. Zhang, et al., Proposing a new friend recommendation method, FRUTAI, to enhance social media providers' performance, Decision Support Systems (2015), http://dx.doi.org/10.1016/j.dss.2015.07.008

3.2. Friend Recommendation with a User's Total Attributes Information

For a friend recommendation system, an example of a candidate friend may be $\langle x_1, x_2, \cdots, x_i, \cdots, x_m \rangle$. $x_i (i \in \{1, \cdots, m\})$ stands for the attributes of the candidate, such as gender, age, location, interest, and number of common neighbors, and these attributes may be independent or not. For example, young men may show strong interest in sports, so gender and age will actually have influence on the attribute of interest. However, Eq. (4) is defined under the condition that each user attribute is independent of the others. We argue that even if some of the attributes are not independent, we still can use Eq. (4) to calculate the total probability of friend recommendation under the strong independence assumption. The reason is that we do not use the calculated probability value to directly predict the chance that the candidate will really become a friend of a user in the future; we just use the probability values to select strong potential candidates. Our proposed friend recommendation system gives a user a list of candidate friends ranked by the probability values.

The advantage of decoupling class attributes using the strong independence assumption is that we can independently calculate each user attribute distribution quickly, and Eq. (4) can be easily extended to other social media networks that may have different sets of users' attributes. Similar to the theory behind the naïve independence assumption used in the successful naïve Bayesian classifier [23], dependence among users' attributes may likely be canceled out, and the performance of our friend recommendation system could still be strong [37–39]. Our empirical results from a case study of a real-world social media strongly support this argument.

For each attribute, we can calculate the prior probability by the data of existing friends of a user. The relation between a candidate and a user can be only one of two types: a friend or not a friend. Let y indicate a binary variable that reflects the relation between the candidate and the user. If the candidate is a friend of the user, $y = 1$; otherwise $y = 0$. Consider $x_i (i \in \{1, \cdots, m\})$ as the attributes of the user, then the probability that the user will collaborate with the candidate is:

$$P(y = 1 | \cap_1^m x_i) = 1 - \prod_1^m (1 - P(y = 1 | x_i)) \tag{5}$$

In Eq. (5), m denotes the number of users' attributes existing in a social media network. $P(y = 1 | x_i)$ denotes the prior probability for each attribute that this candidate will be a friend of a user in the future. It can be calculated by the statistical result including the information of all of the friends of the user's existing friends (friends-of-friend) and how many of them are already friends of the user. $\prod_1^m (1 - P(y = 1 | x_i))$ denotes the probability that the candidate will not be the user's friend based on all of the m attributes.

Here we give an example to explain how the total probability is calculated.

Example 1. For user A, the information of friend candidates B's and C's attributes is presented in Table 1.

Based on the information of all of user A's existing friends, we can generate Tables 2, 3, and 4 (detailed explanation on how those tables are calculated will be discussed in Section 3.3).

Based on Eq. (5), the probabilities that candidates B and C will be friends of user A are:

$$P_B = 1 - (1 - 0.12) * (1 - 0.02) * (1 - 0.25) = 0.3532 \tag{6}$$

$$P_C = 1 - (1 - 0.08) * (1 - 0.12) * (1 - 0.60) = 0.6762 \tag{7}$$

The result implies that candidate C has a greater chance of being a friend of user A in the future; hence, candidate C is ranked higher than candidate B in the recommendation list.

Table 1
Information of candidates B and C, including gender, location, and common-neighbors number between user and candidate.

	Candidate B	Candidate C
Gender	Male	Female
Location	City1	City2
Common-Neighbors	10	30

The algorithm is based on the friend-of-friend algorithm. All of the algorithms of this type limit the candidates to friends of friends, which can decrease the time complexity and have little influence on the accuracy of the recommendation result. We can see when the Common-Neighbor number decreases to 1, the probability of two people becoming friends trends to zero. In a social media network, the number of total users is uncertain, but the friends of friends are limited. This indicates that we can take the friends of friends as candidates to balance the time complexity and the accuracy of the recommendation result.

In a real-world social media network, there will be lots of users' attributes, and the number of attributes will keep increasing along with the expansion of the social media network. This algorithm can be efficiently extended with the number of users' attributes. When a new attribute is added, we just need to calculate the probability $P(x_i)$ of this attribute using the information of a database and extend the equation.

Algorithm 1. FRUTAI (Friend Recommendation with a User's Total Attributes Information)

1. **Input:** The database of friendship relations between users in a social media network; the database of the users'm attributes.
2. Construct the social media relation for a user. All of the user's existing friends are V_t; the set of persons in V_t who have already been friends of the user is V_f; the set of the other n persons in V_t will be the candidates for the friend recommendation system, and we mark it as V_c.
3. Estimate the probability $P(x_1)$ that V_t will be a friend of the user for attribute i by the statistical result of V_t and V_f. For all m attributes, we will obtain $\{P(x_1), P(x_2), \ldots, P(x_m)\}$.
4. Calculate the probability P for each of the n candidates in V_c using Eq. (5) and $\{P(x_1), P(x_2), \ldots, P(x_m)\}$.
5. Sort the n candidates by the value of probability P.
6. **Return:** Top k of the sorted n candidates as the list of friend recommendation result.

The pseudo-code of the recommendation algorithm FRUTAI is shown in Algorithm 1. In step 3, if calculating each P of the attribute costs time m and there are n attributes, the time complexity of step 3 is $O(mn)$; in step 4, if calculating each P of the candidates costs time m and if there are n candidates, the time complexity of step 4 is $O(mn)$; in step 5, we use the function "Rank( )" in SQL to sort the results and the time complexity of step 5 is $O(n \log n)$. The total time complexity of FRUTAI is $O(2mn + n \log n)$.

3.3. Prior probability

To calculate the probability for each candidate, we need to know the prior probability $P(x_i)$ of every attribute, which can be computed by the statistical result, including the information of all of the user's existing friends (friends-of-friend) and the number of them that are already

Table 2
The prior probability of gender for user A. It is generated based on the information of all of user A's existing friends.

Gender	User A
Male	0.12
Female	0.08

Table 3
The prior probability of location for user A. It is generated based on the information of all of user A's existing friends.

City	User A
City1	0.02
City2	0.12
City3	0.15
…	…

Table 5
The statistics of the gender of users' friends-of-friend number. The percentage of the real friends in the total friends-of-friend number ($P(x_{gender})$) can be generated by the number of the user's existing friends (V_t) and how many of these friends-of-friends are already friends of the user (V_f).

Gender	V_t	V_f	$P(x_{gender})$
0	a_1	b_1	b_1/a_1
1	a_2	b_2	b_2/a_2

friends of the user. Before we calculate the $P(x_i)$ for each attribute, we will first discuss the type of users' attributes. In a real-world social media network, all of the attributes can be divided into two types in this research based on the form of the $P(x_i)$, the discrete variable, and the continuous variable. The attributes such as gender, location, etc., are the discrete variables; the attributes such as number of common neighbors between user and candidate, number of candidate's friends, etc., are the continuous variables.

The discrete attributes may have several fixed variable values. By analyzing the information of all of the user's existing friends (V_t) and counting the number of these friends-of-friend persons who are already friends of the user (V_f). A table will be generated that shows the relationship between each variable value and the percentage of the real friends in the total friends-of-friend number. When the probability of a candidate friend for an attribute is calculated, we check the table and find the prior probability with respect to the particular value of a user's attribute. In addition, for different users in a real-world social media network, the $P(x_i)$ value in their own tables will be different from the other users, and it shows the diversity of users' motivation in choosing a friend. It makes the friend recommendation algorithm FRUTAI more accurate for individuals by analyzing the information.

$$P(x) = P_i, \ \ if \ \ x = x_i(i \in \{0, 1, \cdots, n\}) \tag{8}$$

Take gender, for example. A candidate can only be male or female; if the candidate is male, the gender information in the database is recorded as 1; otherwise 0, and $x_{gender} \in \{0, 1\}$. Then, the prior probability table based on the candidate's attribute of gender is shown as Table 5.

Different from discrete attributes, if we still calculate the probability table separately for each user, the data size of V_t and V_f is so small after dividing by the number of variable numbers that it will absolutely reduce the accuracy of the recommendation result. Fortunately, these types of attributes always show an obvious trend between the variable values and $P(x)$ according to the statistic result of a large amount of data. Although the users in a social media network have different personalities, this trend is always similar among those users. We can use all of the users' information for this attribute to calculate a regression function $F(x)$ of this common trend.

$$P(x) = F(x) \tag{9}$$

Table 4
The prior probability of Common-Neighbors for user A. It is generated based on the information of all of user A's existing friends.

Common-Neighbors	User A
1	0.01
2	0.01
…	…
10	0.25
…	…
30	0.60
…	…

To explain the way to use continuous attributes, we take the number of common neighbors, for example. With the database of the users' information in a social media network, we can easily know the number of common neighbors (CN) between every two users. Additionally, for each user, we can know the number of the user's friends of friends and how many of them are already friends of the user. Then, we can use regression to evaluate the $P(x_{cn})$ based on the value of number of CN and probability.

3.4. Appraisal procedure

To evaluate the performance of our proposed recommendation system (FRUTAI), we use three different measures: P@k, MRR, and MAP.

P@k (Precision@k) is a widely used method to evaluate the performance of information retrieval systems [14,16]. P@k = n/k, where k is the number of people who are recommended by the system and n is the number of true friends in a recommendation list. P@k is used to evaluate the precision of the top k persons in the recommendation list. The limitation of P@k is that this measure focuses only on the precision of friend recommendation results but is insensitive to the rank of the k persons. For example, the accuracy of the first recommended person and the accuracy of the last one have equal contribution to the value of P@k. Obviously, when we use the friend recommendation system in a real-world social media network, we always browse over the recommendation results from top to bottom. The ones on the top will have a greater chance of being noticed than the ones below. Only using P@k is not enough to reflect all of the hidden problems of the algorithms. In this paper, we choose 1, 2, 5, 10, 20, and 50 as the values of k to show the precision of the algorithms in different ranges of recommendation.

To address the limitation of P@k, MRR is proposed [14]. MRR (mean reciprocal rank) is a measure of navigational searching or question answering, which focuses on the rank of the first correct one in the recommendation list. MRR is the average of reciprocal ranks of the first correct answer for a set of queries. In the field of friend recommendation, MRR is used to evaluate the accuracy of algorithms using the rank of the first correctly recommended person. The limitation of MRR is that it focuses only on the rank of the first correct result but ignores the other correct ones. Different from information retrieval, users of a real-world social media network may intend to find more than one person as a friend when using friend recommendation system, and thus all of the correctly recommended ones are relevant and useful to them. Thus, this measure is still not good enough to evaluate algorithms.

MAP (mean average precision) takes into account the rank of all of the correct answers in the response list of a query. MAP is the mean of the average precision values for a set of queries. In a recommendation system, the first people recommended are of great importance to users, and it may impact users' satisfaction with the system. Although MAP is the most suitable measure for recommendation systems, the other two measures can also complement the measurement of the performance of algorithms. Hence we use all of the three methods to evaluate our proposed new FRUTAI system.

Please cite this article as: Z. Zhang, et al., Proposing a new friend recommendation method, FRUTAI, to enhance social media providers' performance, Decision Support Systems (2015), http://dx.doi.org/10.1016/j.dss.2015.07.008

Table 6
Examples of users' public information dataset in RenRen. This dataset stores the user's ID, name, gender, location, the number of existing friends, and the number of public blogs or micro-blogs (similar to Twitter).

ID	ID1	ID2	ID3	…
Name	User1	User2	User3	…
Gender	Gender1	Gender2	Gender1	…
Province	Prov No.5	Prov No.3	Prov No.8	…
City	City No.23	City No.6	City No.57	…
nFriends	n_1	n_2	n_3	…
nBlogs	m_1	m_2	m_3	…
…	…	…	…	…

4. Empirical study

4.1. Data collection

To carry out experiments, we use a web crawler to collect user data from RenRen (http://www.renren.com) and store it in a database. RenRen is one of the most popular social media websites in China and has more than 200 million users. The information on RenRen can be divided into private information and public information. Public information is available to all users in RenRen. On the contrary, private information can be seen only by a user's friends in RenRen. In our study, due to legal privacy issues, we use only the public information.

First, to start, we download the information of 240 users with their attributes. We define them as D_1 nodes. Second, we extend to collect the information of 51,340 D_2 nodes that are the friends of those 240 users. Third, we keep on collecting the data of the D_2 users' friends and we call them D_3 nodes. There are 7,158,934 D_3 in total. These nodes and the edges between them form a social media structure for our case study.

Two datasets are used in the experiments. Nodes' attributes are stored in the first dataset, which contains 7 million users' public information, which includes a user's ID, name, gender, hometown, location, the number of friends, the number of public blogs or micro-blogs (similar to Twitter), whether a user sets up a barrier to prevent strangers from visiting the user's homepage, whether a user pays for more privilege on the website (a premium user), whether a user binds his/her mobile-phone, etc. The second dataset stores friend relationships between users. The data samples are shown in Tables 6 and 7. In our datasets, we have more than 7 million users' public information with their attributes. Of these 7 million users, more than 3 million users have filled in their province/state information, and Fig. 1 shows the statistics of this location information. We can see that the users are distributed among 34 provinces of China. The Jiangsu province has the largest number of users, which is over 305,000. Macao has the smallest number with 3000. The facial validity shows that the distribution is in line with the actual population of each province. Thus, our experiments have been performed on a representative dataset with quality sampling data.

Table 7
Examples of users' relations dataset in RenRen. This dataset stores friend relationships between users.

User	Friend
ID1	ID4
ID1	ID5
ID2	ID5
ID2	ID6
ID3	ID4
ID4	ID7
…	…

Fig. 1. The distribution of users' location in RenRen. The users are distributed among 34 provinces of China. The facial validity shows that the distribution is in line with the actual population of each province.

Fig. 2 shows the distribution of users' gender. 283 thousand users are female, 303 thousand users are male, and 134 thousand users do not indicate their gender.

4.2. Evaluation

Using the collected data, we have evaluated FRUTAI system's recommendation results against the other three commonly used FoF algorithms, which we mention in Section 2.3, Common-Neighbors, Jaccard's coefficient and Adamic/Adar.

We use k-fold cross validation to evaluate the result of the friend recommendation. First, we split the user's friends into 10 partitions. We take nine partitions as the training dataset and one partition as the testing dataset. The prior probability for each attribute can be calculated from the information of the training dataset, just as we mentioned in Example 1 in Section 3.2. Second, we collect the friends of the training dataset as the candidates for the friend recommendation. For each candidate, we calculate the probability that he/she will become the friend of a user by Eq. (5). After that, we can obtain a rank of the probability, and the top 100 candidates are selected as the recommendation result for the user. This list will be compared with the testing dataset and three different measures, P@k, MRR, and MAP, will be used to evaluate the performance of our proposed recommendation system

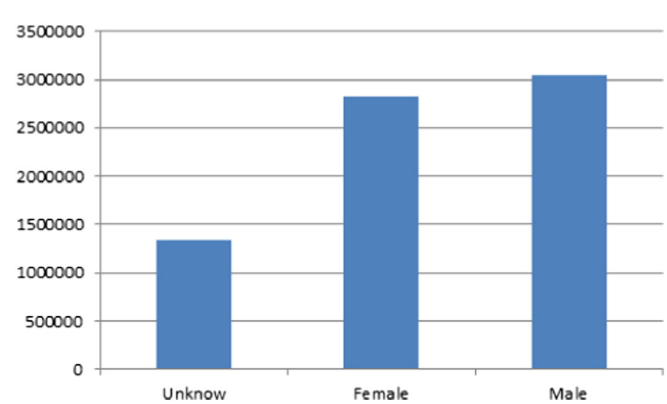

Fig. 2. The distribution of users' gender in RenRen. 283 thousand users are female, 303 thousand users are male, and 134 thousand users do not indicate their gender.

Please cite this article as: Z. Zhang, et al., Proposing a new friend recommendation method, FRUTAI, to enhance social media providers' performance, Decision Support Systems (2015), http://dx.doi.org/10.1016/j.dss.2015.07.008

Table 8
The average prior probability of gender. For a male user in RenRen, the prior probability that a male candidate will be his friend is higher than a female candidate. Similar results are obtained for a female user.

		Candidate	
		Male	Female
User	Male	0.00861	0.00589
	Female	0.00881	0.00519

(FRUTAI). Because we have 240 users in total, we will repeat the experiment 240 times and take the average numbers as the final results.

In the k-fold cross validation, we calculate the prior probability for each attribute. Tables 8 and 9 show the average value of the prior probability of gender and location. It is not easy to note whether the homophily principle or the heterophily principle plays a more important role in gender. For the location information, it seems that homophily principle plays a leading role. For the traditional study, although the researchers have tried to distinguish the homophily and the heterophily principles for decades, there is still a conflict in prior research. It is not easy to design a recommendation system based on whether the homophily or the heterophily principles apply because there is no widely accepted conclusion.

The FRUTAI is a clever algorithm that can handle this problem. The FRUTAI considers that different users may have their own preferences of attributes. Take gender, for example. Some males prefer to make friends with males, and some prefer females. The FRUTAI collects the information of each user's existing friends, calculates the prior probability of this attribute, and gives a personal recommendation result. Moreover, if the user's preference changes, it will be reflected in the information that is collected, and the prior probability and the final recommendation result will change with it.

As depicted in Fig. 3, this method of handling the data collected at a time point is commonly used in the field of friend recommendation. A limitation of this evaluation method is that the friend recommendation results that are not in the set of the partition do not mean that they are wrong because some of them may be the potential friends of a user and will be added by the user as friends in the future. Therefore, we expect that the actual precision value of the algorithms would be higher than the value in the evaluation report.

First, we pick out all of the 240 users (D_1). We randomly pick out one tenth of each D_1's friends (D_2) as D_2'' and define the other nine tenths as D_2'. We try to give a friend recommendation list D_R for each D_1 by using the information of D_2'. D_2'' is the friend recommendation target and will be compared with D_R.

4.3. Results and discussion

The link prediction results are shown in Tables 10–12.

Table 10 shows an overall result of the friend recommendation for the 240 D_1 users in RenRen. We can see that FRUTAI performs the best in MAP (16.97%), and some P@N (76.92% precision at P@1, 50.17% precision at P@2, and 10.83% precision at P@100). Common-Neighbors and Adamic/Adar perform well too. Their MRRs are 40.51%/41.59%, and MAPs are 16.24%/15.97%, both comparable to FRUTAI. The result of Jaccard's coefficient is acceptable, but it is worse than the other three algorithms.

Table 9
The average prior probability of location. If the candidate and the user live in a same city, the prior probability that they will become friends is higher than the situation where they live in different cities.

	Candidate	
User	Same location	Different location
	0.00511	0.00025

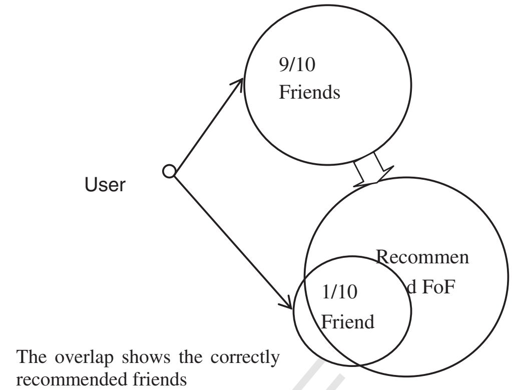

Fig. 3. Evaluation setup.

Then, we divide the D_1 users into two groups using the number of their friends, and repeat the experiments. Table 11 shows the result of the D_1 users who have fewer than 100 friends. Table 12 shows the result of the D_1 users who have more than 100 friends.

In Table 11, all of the results are worse than in Table 10, as expected. With less information of the user's friends, it is difficult to recommend friends to a user by FoF methods. The FRUTAI has the best MAP (20.69%), P@50 (2.95%). The Common-Neighbors has the best P@1 (40.68%), P@2 (16.27%), P@5 (10.51%), and P@10 (6.36%). The result of Adamic/Adar is not as good as Common-Neighbors and FRUTAI, but is still comparable. The result of Jaccard's coefficient is much worse than other two algorithms.

In Table 12, all of the results are better than Table 10. The Common-Neighbors beats the other three algorithms in most of the indices (MRR 44.17%, P@5 49.03%, P@10 36.74%, P@50 22.29%, and P@100 13.95%). The result of FRUTAI is impressively outstanding on P@1 91.43% and P@2 61.71%. Because the top recommended person is always the first one browsed by a user, P@1 is the most important in P@k. The results of Adamic/Adar are comparable to FRUTAI and Common-Neighbors. Jaccard's coefficient is still worse than the other three, but the gap is evidently narrowed from the values in Table 10.

Our extensive empirical studies have shown that

(1) Overall, FRUTAI performs much better than the other algorithms. The performances of the Common-Neighbors and Adamic/Adar algorithms are better than Jaccard's coefficient;

Table 10
Overall Precision, MRR (mean reciprocal rank) and MAP (mean average precision) results of algorithm comparison including FRUTAI, Common-Neighbors, Jaccard's coefficient, and Adamic/Adar. Higher scores (in bold) indicate better performance.

	P@1	P@2	P@5	P@10	P@50	P@100	MRR	MAP
FRUTAI	**0.7692**	**0.5017**	0.3897	0.2823	0.1719	**0.1083**	0.4121	**0.1697**
CN	0.6581	0.4957	**0.3932**	**0.2908**	**0.1737**	**0.1083**	0.4051	0.1624
JAC	0.5000	0.4171	0.3436	0.2675	0.1649	0.1069	0.3736	0.1340
ADA	0.6154	0.4744	0.3782	0.2812	0.1679	0.1076	**0.4159**	0.1597

Table 11
Overall Precision, MRR (mean reciprocal rank) and MAP (mean average precision) results of algorithms comparison including FRUTAI, Common-Neighbors, Jaccard's coefficient and Adamic/Adar (friends < 100). Higher scores (in bold) indicate better performance.

	P@1	P@2	P@5	P@10	P@50	P@100	MRR	MAP
FRUTAI	0.3390	0.1593	0.1000	0.0627	**0.0295**	0.0164	0.3287	**0.2069**
CN	**0.4068**	**0.1627**	**0.1051**	**0.0636**	0.0281	0.0158	0.2963	0.1739
Jaccard	0.1186	0.0610	0.0492	0.0305	0.0183	0.0112	0.1901	0.0997
Ada	0.2373	0.1288	0.0847	0.0576	0.0281	**0.0169**	**0.3430**	0.1530

Table 12
Overall Precision, MRR (mean reciprocal rank) and MAP (mean average precision) results of algorithms comparison including FRUTAI, Common-Neighbors, Jaccard's coefficient and Adamic/Adar (friends > 100). Higher scores (in bold) indicate better performance.

	P@1	P@2	P@5	P@10	P@50	P@100	MRR	MAP
FRUTAI	**0.9143**	**0.6171**	0.4874	0.3563	0.2199	0.1393	0.4402	0.1572
CN	0.7429	0.6080	**0.4903**	**0.3674**	**0.2229**	**0.1395**	**0.4417**	0.1586
Jaccard	0.6286	0.5371	0.4429	0.3474	0.2143	0.1391	0.4350	0.1456
Ada	0.7429	0.5909	0.4771	0.3566	0.2151	0.1382	0.4404	**0.1619**

(2) When a user has relatively fewer friends (e.g., <100), FRUTAI performs better than Adamic/Adar and Common-Neighbors, and much better than Jaccard's coefficient;

(3) When a user has relatively more friends (e.g., >100), the performance of FRUTAI, Common-Neighbors and Adamic/Adar are comparable. Jaccard's coefficient is still the worst. The precision of FRUTAI is impressively outstanding with the top recommended results.

Different from other FoF algorithms, the FRUTAI utilizes the user's attributes to improve the accuracy of the prediction. As we explain in Section 3.2, the prior probability for each attribute that this candidate will be a friend of a user in the future can be calculated by the statistical result including the information of all of the friends of the user's existing friends (friends-of-friend) and the number of them that are already friends of the user. It leads to a correlation of the number of the user's existing friends and the accuracy of the recommendation result. When the number of the user's existing friends increases (which is a trend in the social media), the precision of the recommendation result will be better.

5. Conclusions

In this paper, we propose a new friend recommendation method and algorithm, FRUTAI, to enhance social media services and performances. We compare the newly proposed FRUTAI method/algorithm with other FoF algorithms using a real-world social media network. Our results show that FRUTAI performs best overall. Our study also finds out that the performance of all of these friend recommendation methods may depend on the number of a user's existing friends. When the number of existing friends falls to less than 100, the result of Jaccard's coefficient may be unacceptable, and Adamic/Adar performs worse but is still acceptable. By contrast, Common-Neighbors and FRUTAI keep performing well. Furthermore, FRUTAI keeps its strong performance when the number of existing friends increases, while other algorithms may not be able to do so.

We have observed that the way of utilizing information is crucial for an algorithm. Adding extra information to an algorithm does not necessarily enhance the performance of an algorithm, unless the information is integrated properly. The Common-Neighbors algorithm utilizes only the number of common neighbors. Jaccard's coefficient utilizes more information, including the number of common neighbors, the number of a user's and the candidate's friends. However, interestingly, it performs worse than the Common-Neighbors algorithm, perhaps because the three attribute numbers are integrated arbitrarily rather than properly. The Adamic/Adar algorithm also utilizes more information, including the number of friends of common neighbors. However, when the number of common friends is relatively low, introducing extra information to the algorithm may introduce too much noise; thus, the Adamic/Adar algorithm does not perform better than the Common-Neighbors algorithm. When the number of common neighbors is relatively high, the noise brought by the number of friends of common neighbors is diminished, thus Adamic/Adar algorithm performs better than Common-Neighbor algorithm. Compared with Adamic/Adar, FRUTAI efficiently utilizes users' information. It can handle all of the user attributes flexibly in a social media network. The recommendation results can be enhanced with the increase of the number of user's attributes.

The proposed FRUTAI is a generic friend recommendation method that has a flexible format that can be easily extended to adding the user's additional important attributes when needed. This friend recommendation system may enhance social media providers' performance by meeting the increasing demand of interaction between users. The friend recommendation system may also enhance the user loyalty to a social media network, which will impact the marketing position of the social media providers in the high competition of attracting more users.

There are limitations to this research. The first is that the proposed algorithm is based on an assumption of independent attributes. In future research, mechanisms of dependent attributes can be considered. The second research limitation is that the dataset used in this paper that comes from a single website. In future research, more datasets could be used to further validate the effectiveness of the proposed friend recommendation method/algorithm.

6. Uncited references Q4

[19,20,48]

Acknowledgment

Wei (Wayne) Huang would like to acknowledge financial support from the National Natural Science Foundation of China under grant number 71331005. Qin Su would like to acknowledge financial support from the National Natural Science Foundation of China under grant number 71371151. Yuewen Liu would like to acknowledge financial support from the National Natural Science Foundation of China under grant number 71301128.

References

[1] M. Pazzani, D. Billsus, Content-based recommendation systems, The adaptive web2007 325–341.
[2] M.J. Pazzani, A framework for collaborative, content-based and demographic filtering, Artificial Intelligence Review 13 (1999) 393–408.
[3] G. Adomavicius, A. Tuzhilin, Toward the next generation of recommender systems: a survey of the state-of-the-art and possible extensions, IEEE Transactions on Knowledge and Data Engineering 17 (2005) 734–749.
[4] M.E. Newman, Clustering and preferential attachment in growing networks, Physical Review E 64 (2001) 025102.
[5] E.M. Jin, M. Girvan, M.E. Newman, Structure of growing social networks, Physical Review E 64 (2001) 046132.
[6] L.A. Adamic, E. Adar, Friends and neighbors on the web, Social Networks 25 (2003) 211–230.
[7] A.-L. Barabási, R. Albert, Emergence of scaling in random networks, Science 286 (1999) 509–512.
[8] A.-L. Barabâsi, H. Jeong, Z. Néda, E. Ravasz, A. Schubert, T. Vicsek, Evolution of the social network of scientific collaborations, Physica A: Statistical Mechanics and its Applications 311 (2002) 590–614.
[9] L. Katz, A new status index derived from sociometric analysis, Psychometrika 18 (1953) 39–43.
[10] S. Brin, L. Page, The anatomy of a large-scale hypertextual Web search engine, Computer networks and ISDN systems 30 (1998) 107–117.
[11] T.H. Haveliwala, Topic-sensitive PageRank: a context-sensitive ranking algorithm for web search, IEEE Transactions on Knowledge and Data Engineering 15 (2003) 784–796.
[12] T. Haveliwala, S. Kamvar, G. Jeh, An analytical comparison of approaches to personalizing PageRank, 2003.
[13] G. Jeh, J. Widom, SimRank: A Measure of Structural-Context Similarity, Proceedings of the Eighth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining 2002, pp. 538–543.
[14] Z. Yin, M. Gupta, T. Weninger, J. Han, A Unified Framework for Link Recommendation Using Random Walks, International Conference on Advances in Social Networks Analysis and Mining (ASONAM), 2010 2010, pp. 152–159.
[15] G. Salton, M.J. McGill, Introduction to modern information retrieval, 1986.
[16] Z. Huang, X. Li, H. Chen, Link Prediction Approach to Collaborative Filtering, Proceedings of the 5th ACM/IEEE-CS Joint Conference on Digital Libraries 2005, pp. 141–142.
[17] D. Liben-Nowell, J. Kleinberg, The link-prediction problem for social networks, Journal of the American Society for Information Science and Technology 58 (2007) 1019–1031.

Please cite this article as: Z. Zhang, et al., Proposing a new friend recommendation method, FRUTAI, to enhance social media providers' performance, Decision Support Systems (2015), http://dx.doi.org/10.1016/j.dss.2015.07.008

[18] J. Chen, W. Geyer, C. Dugan, M. Muller, I. Guy, Make New Friends, But Keep the Old: Recommending People on Social Networking Sites, Proceedings of the 27th International Conference on Human Factors in Computing Systems 2009, pp. 201–210.

[19] J.S. Breese, D. Heckerman, C. Kadie, Empirical Analysis of Predictive Algorithms for Collaborative Filtering, Proceedings of the Fourteenth Conference on Uncertainty in, Artificial Intelligence 1998, pp. 43–52.

[20] H. Zhang, J. Su, Naive Bayesian Classifiers for Ranking, Machine Learning: ECML 2004, Springer 2004, pp. 501–512.

[21] M. Claypool, A. Gokhale, T. Miranda, P. Murnikov, D. Netes, M. Sartin, Combining Content-Based and Collaborative Filters in an Online Newspaper, ACM SIGIR'99. Workshop on Recommender Systems: Algorithms and Evaluation, August 1999.

[22] I. Soboroff, C. Nicholas, Combining Content and Collaboration in Text Filtering, 43 IJCAI'99 Workshop: Machine Learning for Information Filtering, August 1999.

[23] H. Zhang, The Optimality of Naive Bayes, Proceedings of the Seventeenth International Florida Artificial Intelligence Research Society Conference, Miami Beach, AAAI Press, 2004.

[24] S. Lo, C. Lin, WMR–A Graph-Based Algorithm for Friend Recommendation, Proceedings of the 2006 IEEE/WIC/ACM International Conference on Web Intelligence, 2006.

[25] A. Kaplan, M. Haenlein, Users of the world, unite! The challenges and opportunities of social media, Business Horizons 53 (2010) 59–68.

[26] V. Kumar, V. Bhaskaran, R. Mirchandani, M. Shah, Practice prize winner—creating a measurable social media marketing strategy: increasing the value and ROI of intangibles and tangibles for hokey pokey, Marketing Science 32 (2) (2013) 194–212.

[27] A. Bharadwaj, O. El Sawy, P. Pavlou, N. Venkatraman, Digital business strategy: toward a next generation of insights, MIS Quarterly 37 (2) (2013) 471–482.

[28] E. Urban Jr., R. Boscolo, Using scientific meetings to enhance the development of early career scientists, Oceanography 26 (2) (2013) 164–170.

[29] T. Røssvoll, L. Fritsch, Trustworthy and Inclusive Identity Management for Applications in Social Media, Proceedings of Human-Computer Interaction, Users and Contexts of Use 15th International Conference, HCI International 2013, Lecture Notes in Computer Science Volume 8006 2013, pp. 68–77.

[30] S. Aral, C. Dellarocas, D. Godes, Introduction to the special issue-social media and business transformation: a framework for research, Information Systems Research 24 (1) (2013) 3–13.

[31] A. Coustasse, S. Chelsea, Potential benefits of using Facebook in the healthcare industry: a literature review, Insights to a Changing World Journal 2013 (1) (2013) 41–52.

[32] Y. Lin, V. Marcus, Creating and assessing a subject-based blog for current awareness within a cancer care environment, Grey Journal (TGJ) 9 (1) (2013) 7–13.

[33] H. Yang, C. Yang, Harnessing Social Media for Drug-Drug Interactions Detection, IEEE International Conference on Healthcare Informatics (ICHI), 2013, 2013.

[34] R. Davies, G. Cairncross, Student tourism and destination choice: exploring the influence of traditional, new, and social media: an Australian case study, Tourism Culture & Communication 13 (1) (2013) 29-42(14).

[35] N. Kolb, D. Roberts, Police chief, 80 (6) (2013).

[36] http://blog.facebook.com/blog.php?post=15610312130

[37] http://quan.qq.com/help.html

[38] H. Peng, F. Long, C. Ding, Feature selection based on mutual information criteria of max-dependency, max-relevance, and min-redundancy, IEEE Transactions on Pattern Analysis and Machine Intelligence 27 (8) (2005) 1226–1238.

[39] L.I. Kuncheva, On the Optimality of Naive Bayes with Dependent Binary Features, Pattern Recognition Letters 27.72006 830–837.

[40] H. Zhang, S. Jiang, Naive Bayes for optimal ranking, Journal of Experimental & Theoretical Artificial Intelligence 20 (2) (2008) 79–93.

[41] M. McPherson, L. Smith-Lovin, J.M. Cook, Birds of a feather: homophily in social networks, Annual Review of Sociology 415–444 (2001).

[42] M.S. Granovetter, The strength of weak ties, American Journal of Sociology 1360–1380 (1973).

[43] L. Smith-Lovin, J.M. McPherson, You Are Who You Know: A Network Approach to Gender, Theory on Gender/Feminism on Theory1993 223–251.

[44] D. Eder, M.T. Hallinan, Sex differences in children's friendships, American Sociological Review 237–250 (1978).

[45] P.V. Marsden, Core discussion networks of Americans, American Sociological Review 122–131 (1987).

[46] R.R. Huckfeldt, Citizens, Politics and Social Communication: Information and Influence in an Election Campaign, Cambridge University Press, 1995.

[47] D.S. Kaufer, K.M. Carley, Communication at a Distance: The influence of Print on Sociocutural Organization and Change, Psychology Press, 1993.

[48] K.N. Hampton, B. Wellman, Examining Community in the Digital Neighborhood: Early Results from Canada's Wired Suburb, Springer, Berlin Heidelberg, 2000 194–208.

Zhou Zhang received the BS and MS degrees in the School of Electronics Engineering and Computer Science from Peking University in 2008 and 2011. Now he is a PhD candidate in the School of Management, Xi'an Jiaotong University. His research interests include social network analysis, recommender systems, and data quality.

Yuewen Liu received the PhD degrees in the College of Business from City University of Hong Kong and the School of Management from University of Science and Technology of China in 2010. From 2010 to 2011, He was a senior engineer in Tencent Technology (Shenzhen) Company Limited. Currently, he is with the Xi'an Jiaotong University. His research interests include electronic commerce and social network.

Wei Ding received her Ph.D. in Computer Science from the University of Houston in Houston, Texas in May 2008, then joined the Department of Computer Science of UMass Boston as an assistant professor in Fall 2008. Wei received her BS degree in Computer Science and Applications from Xi'an Jiaotong University and her MS degree in Software Engineering from George Mason University (find her at the Software Engineering Academic Genealogy).From 2002 to 2008, Wei had been a full-time lecturer of the Computer Science and Computer Information Systems programs at the University of Houston—Clear Lake (UHCL). Wei has an 8-year full-time working experience in banking, software development, and web technology. Wei worked as a software engineer for the Bank of China, a software testing engineer for Microsoft (China) Ltd., a systems analyst and project manager for PanSky International Holding Co. Ltd, a quality assurance team leader for , MultiCity.com, and a technical consultant and software engineer for VeriSign Inc. She is a senior member of the IEEE and a member of the ACM.

Wei (Wayne) Huang received the BS degree from Huazhong University of Science and Technology in 1984 and MS degree in the School of Management from Xi'an Jiaotong University in 1986. He received Ph.D. degrees in Information Systems from the National University of Singapore and the University of Georgia. Currently, he is with the College of Business, Ohio University as a professor of information systems. He was a fellow of Harvard University during 2009–2010. He worked as a full faculty in top-tier research universities in Australia, Singapore, Hong Kong and China. He published more than 50 research papers in international peer-review IS journals, including in top-tier journals such as MIS Quality, Journal of MIS, Communications of ACM, ACM Transactions, IEEE transactions, European Journal of Information Systems, etc.

Qin Su received the BS, MS and Ph.D. degree in the School of Management from Xi'an Jiaotong University in 1984, 1987 and 1993. She joined the School of Management of Xi'an Jiaotong University in 1993 and became a professor in 2001.

Ping Chen is an associate professor of computer engineering and the Director of Artificial Intelligence Lab at the University of Massachusetts Boston. His research interests include bioinformatics, data mining, and computational semantics. Dr. Chen has received five NSF grants and published over 50 papers in major data mining, artificial intelligence, and bioinformatics conferences and journals. Dr. Ping Chen received his BS degree on Information Science and Technology from Xi'an Jiao Tong University, MS degree on Computer Science from Chinese Academy of Sciences, and Ph.D degree on Information Technology at George Mason University.

Please cite this article as: Z. Zhang, et al., Proposing a new friend recommendation method, FRUTAI, to enhance social media providers' performance, Decision Support Systems (2015), http://dx.doi.org/10.1016/j.dss.2015.07.008