Who Should I Add as a "Friend"? A Study of Friend Recommendations using Proximity and Homophily

Alvin Chin
Nokia
Building 2, No. 5 Donghuan Zhonglu
Beijing, China, 100176
alvin.chin@nokia.com

Bin Xu
Cornell University
301 College Ave.
Ithaca, New York, USA, 12850
bx55@cornell.edu

Hao Wang
Babytree.com
17-2305 Jianwai SOHO
Beijing, China, 10020
alexwhu@gmail.com

ABSTRACT
We receive many recommendations of friends in online social networks such as Facebook and LinkedIn. These friend recommendations are based usually on common friends or similar profile such as having the same interest or coming from the same company, a trait known as homophily. However, many times people do not know why they should add this friend. Should I add this friend because we met from a conference and if so, what conference? Existing friend recommendation systems cannot answer this question easily. In this paper, we create a friend recommendation system using proximity and homophily, that we conduct in the workplace and conference. Besides common friends and common interests (homophily features), we also include encounters and meetings (proximity features) and messages sent and question and answer posts (social interaction features) as reasons for adding this person as a friend. We conduct a user study to examine whether our friend recommendation is better than common friends. Results show that on average, our algorithm recommends more friends to participants that they add and more recommendations are ranked as good, compared with the common friend algorithm. In addition, people add friends due to having encountered them before in real life. The results can be used to help design context-aware recommendations in physical environments and in online social networks.

Categories and Subject Descriptors
H.5.3 **[Information Interfaces and Presentation]**: Group and Organizational Interfaces – Computer supported cooperative work; H.3.3 **[Information Search and Retrieval]** - Information filtering

General Terms
Algorithms, Measurement, Experimentation, Human Factors.

Keywords
Proximity, encounters, homophily, friend recommendations

1. INTRODUCTION
In online social networks such as Facebook, we always receive notifications of people that we may know that we should add as friends. These friend recommendations are mostly based on common friends that you and the recommended friend have in common, similar profile (such as working in the same company or went to the same school or having similar interests) and shared content. We have a tendency to add friends who are similar to us, a principle known as homophily [27]. However, very often there are some recommended friends whom we do not know why they have been recommended. The problem exists in that current recommendation systems do not record the physical context that you may have had with other recommended people.

We need to create a system and platform for recording physical context and social interactions in the real world, which can be used as additional reasons for why we should add this recommended person as our "friend". In our previous work [9, 38], we created this system for recording proximity encounters and meetings in indoor environments such as a workplace and a conference. In [36], we found that more interactions in physical context result in a higher possibility in friendship formation. In our work, we use the term "friend" to refer to a contact that the user has explicitly established a connection with, either online, offline or both.

In this paper, we describe our friend recommendation algorithm used in our system that uses proximity (specifically encounters and meetings as the physical context) and homophily in an indoor environment where people meet each other such as a workplace or conference. We hypothesize that the quality of friend recommendations based on physical context will be better than those based on common friends. To demonstrate if this is true, we perform a user study between these two friend recommendation algorithms.

Results show that our friend recommendation algorithm in the workplace recommends more friends to participants that they add (50% compared to 38% for common friends), more recommendations are ranked as good (44% compared to 32% for common friends), and more have previous acquaintance with these recommendations (69% vs. 59% for common friends), compared with the common friend algorithm. For the conference, we ask users through a survey before the conference why they add friends and also during the conference with our conference application when they add friends. We discover that the top 2 reasons for adding friends from both the survey and our conference application are the same, which are that they know each other in real life and have encountered before. This validates that friend recommendations based on encounters and meetings are useful and helps in the decision to add new friends or contacts.

Our contributions are the following. First, we build a friend recommendation system that is based on proximity using indoor physical context and homophily captured in real time. Second, we conduct a friend recommendation study where we not only examine the quality of the recommendations based on the accepted recommendations, but also ask questions related to acquaintance of the recommended individual and reasons for accepting this friend, therefore making it more clear on who users want to be friends with and why they have the preferences.

The rest of the paper is divided as follows. Section 2 describes related work on proximity and homophily in relation to friend recommendations, and Section 3 discusses about our system for friend recommendations and the features we use for proximity and homophily. Section 4 explains about our interface for friend recommendations in the workplace and the conference, followed by the details of the friend recommendation algorithm. We conduct a user study of friend recommendations in Section 5 and illustrate our results. Finally, Section 6 concludes the paper and discusses areas for future work.

2. RELATED WORK

Most of us have used friend recommendations in social networking sites such as Facebook or LinkedIn, to increase our social circle [33]. In this section, we examine the state-of-the-art in friend recommendations which can be based on homophily and proximity.

2.1 Recommendations Based on Homophily

Current friend recommendation systems in Facebook and LinkedIn [29] known as "People You May Know", are mainly based on common friends and similar profile characteristics such as being in the same network or same company [20]. This tendency to connect with similar people such as similar characteristics, interests and beliefs is called homophily [27], and has been validated to contribute to the usual preferential ties [27, 28]. However, the problem is that the system may still recommend people whom the user does not know so she will not add them as a friend. Being able to recommend more known people is important to improve the quality of the recommendations [7]. Shared content and interactions such as co-authored papers, patents, and comments in an enterprise system [17], is one way for finding similar like-minded people. Guy et al [19] examine nine diverse information sources from three categories (people, things and places) to define user similarity, with which people form ties guided by the homophily principle. Even strangers can be used to recommend people whom you do not know but would be beneficial for you to know [20]. Providing the reasons for why you should add a specific person as a friend [22, 35] (lacking in existing friend recommendations) and providing recommendation feedback [18] can help people in this decision making process.

Friend recommendations are usually ordered by friend recommendation score where weights are assigned to each feature in the recommendation algorithm [12]. However, recommendation systems ignore the physical interactions to associate how you may know that person [13, 24, 34]. Our work differs from the above in that we add physical context (e.g. encounters and meetings) and social interactions (e.g. messages, and question and answer posts) in addition to similar profile (e.g. common interests) and common friends.

2.2 Recommendations Based on Proximity

Previous works [16, 17, 19] show that the more social network information and sources integrated, the closer the returning people are to the ideal friend list. Thus, proximity may be utilized to recommend similar-minded people [10, 11, 31]. Cranshaw [11] shows how physical location can be used to recommend friends in online social networks that are nearby, which has been implemented in applications such as SONAR [32] and Ban.jo. Froehlich et al [14] show that there is a positive correlation between physical place preference and visiting frequency and visit time. Backstrom et al [3] uses self-reported address data from Facebook users and their network ties to measure the relationship between geography and friendship, finding that Facebook users' probability of friendship is roughly inversely proportional to their distance at medium to long-range scale, while in shorter distance scale, the probability is less sensitive to the distance.

Physical location context can be captured from positioning systems such as GPS, WiFi or RFID [25] and have been added in collaborative filtering-based recommendation algorithms [37]. However, many location-based systems use co-location when recommending friends, which only shows that two people are in the same location at the same time, but does not capture the dynamic movement of people. We encounter many people every day and the people that we encounter and meet could present opportunities to make new social connections [30]. Encounter duration and frequency can be used to infer friendship using Bluetooth [13, 31], RFID [4, 23] and WiFi [36], and has been used in commercial applications such as Aka-Aki [1]. Encounters can also be used to introduce people and infer one's social network like Serendipity [2].

In our work, priority weight is given to physical proximity when recommending the potential friends, since the people nearby that you may see and listen to, may be talking with you. Physical context in our work is not absolute location like GPS or RFID location coordinates, but rather semantic where we use encounter as the combination of two types of physical context (location and time) to represent mobility interaction. We also include social context which include the social interactions (messages and question and answer posts) between users in our friend recommendations, similar to Lo and Lin [26] where they use weighted minimum message ratio to determine friends, and Chin and Wang [8] who uses social network analysis on the conversation graph extracted from message interactions.

3. SYSTEM FOR SUPPORTING FRIEND RECOMMENDATIONS

From previous work, we build our system for supporting friend recommendations based on proximity and homophily.

3.1 Recording Proximity

Like previous work, we use encounter as the concept for defining a probable proximity interaction. We define an encounter if the distance between two people is within the encounter distance threshold ΔD and their distance lasts for at least the encounter duration threshold ΔT before they move away and are beyond the encounter distance threshold. From the individual's position, we calculate the distance between any two individuals on the same floor at the same time. We then create an *encounter graph* $G_{en}(V, E)$ where V is the set of nodes ($v_i \mid 1 \leq i < N$), N is the number of nodes and E is the set of edges ($e_{ij} \mid 1 \leq i < N, 1 \leq j < N, i \neq j$) and

• node v_i designates user i, node v_j is user j and the edge e_{ij} is a link when two users (v_i and v_j) encounter each other,
• edge e_{ij} has a timestamp attribute to define when the encounter happens called $T_{en\ start}(e_{ij})$ and when the encounter ends called $T_{en\ stop}(e_{ij})$, and
• edge e_{ij} is built only if the encounter distance $D_{en}(e_{ij})$ is less than the encounter distance threshold ΔD and the encounter duration $\Delta T_{en}(e_{ij}) = T_{en\ stop}(e_{ij}) - T_{en\ start}(e_{ij})$ is larger than ΔT (the time duration threshold that is defined to be an encounter)

The collection of all the encounters forms the *encounter network*. Note that the encounter is defined as a pairwise proximity interaction between two users where the speed of the users is less than the normal walking speed (considered to be 1.5 m/s [5]) because they likely stop to talk to each other. If the walking speed of the users is greater than the normal walking speed, we say that the users *passby* each other and record that. However, since the positioning accuracy can be off (5 meters for WiFi based on our experience) and we do not know if the encounter is a face-to-face encounter since we do not record orientation (e.g. two people may be back-to-back), the encounter may not be a true encounter. We assume that these encounters could be possible encounters, that is, there is a high probability that an encounter may have occurred. If you physically encounter a particular person more and have longer encounter duration, *common encounters* may increase the probability that you will add this person as a friend. The reason why is because there is an attractive force of similarity between the both of you, the similarities being that both of you are at the same place at the same time doing the same activity. In our system, we define the encounter distance threshold ΔD as 4 metres since Hall defines this as a proxemic interaction [21] and the encounter duration threshold ΔT is 5 minutes because intuitively if you meet with someone for an activity (like a chat), it will last at least a couple of minutes [39].

3.2 Recording Homophily

Our system is designed to allow people to find and connect with each other in indoor environments such as a workplace [38] or at a conference [39]. However, what are the underlying common factors or similar features that people take into consideration for adding others as friends? One source of homophily for adding friends, as mentioned earlier in the previous section, is *common friends* between you and another person. We tend to meet others who are friends of our friends from meetings and social events, so it is natural for us to want to be friends of our friends.

A second source of homophily is *common interests* where we like to be friends with others who have similar interests to us, again also used in friend recommendations. We also meet with people during meetings and sessions, where we exchange business cards for future contact, thus *common meetings* or *common sessions* is a third source of homophily which can be used as a reason for why people add others as friends, and this part is new which does not exist in state-of-the-art friend recommendations. Finally, besides physical interactions, social interactions may form a reason for adding others as friends (as explained in the previous section), thus we also record *messages sent* and *question and answer posts* between two people.

4. FRIEND RECOMMENDATIONS

As mentioned earlier, the problem with friend recommendations is that they do not capture the offline physical context as reasons for adding this person as a friend. In this section, we explain how we combine proximity and homophily in order to create friend recommendations. Our friend recommendation system is designed to be used where people come together not for social engagement, but through a shared interest, common goal or work environment. For example, many times we meet new friends from events where we have common interests. We explain the user interface that we use for the workplace and conference for displaying the reasons for adding an individual as a friend, followed by the friend recommendation algorithm used to rank the list of potential friends.

4.1 Interface

4.1.1 Workplace

In the workplace, people have problems trying to find and book an empty meeting room for conducting meetings, and finding people. When people meet with each other in meetings, they usually exchange business cards and begin to establish a professional relationship. However, sometimes people forget to bring their business cards, and thus they do not recall who they have met, where and when. We design an application to solve these problems, by recording the position of people in the workplace, whether they attended the meeting and with whom, and people they have encountered throughout the day [38]. By doing so, we can provide physical reasons for why a person should add this person as a "friend".

Figure 1 on the next page shows the friend recommendation interface in the workplace, which is displayed when the user selects the recommended friend from the recommendations list. It shows the number of stars for rating this person compared to others, the person's profile, the activities of the user (e.g. the meetings that the user has had), the reasons for adding this person as a friend, followed by the option to add this person as a friend, similar to Guy et al [20]. The star rating comes from ranking the scores of our friend recommendation algorithm and converting it into the number of stars. We include the recommended friend's profile and activities in order to provide additional information to the user, that can help the user to decide whether to add her as a friend.

The reasons for adding this person as a friend include the number of *common interests* (along with what those common interests are), number of *common friends* (along with who they are), number of *meetings* together (along with what those meetings are), number of *encounters* (as well as the details of the last encounter) and number of *passbys* (as well as the details of the last passby). We also include the total number of interests, friends, and meetings for reference purposes. In addition, we also add the communication media used between you and the recommended friend because online social interactions are important as well as offline social interactions (from meetings and encounters) for forming the communication basis as to why you may know this person. For the online social interactions in our system, we use the number of *messages sent* to the recommended friend and number of *answers* that the recommended friend replies to a question that the user sends, since in our system, users can send a question to a group of people which we call mobile Q&A.

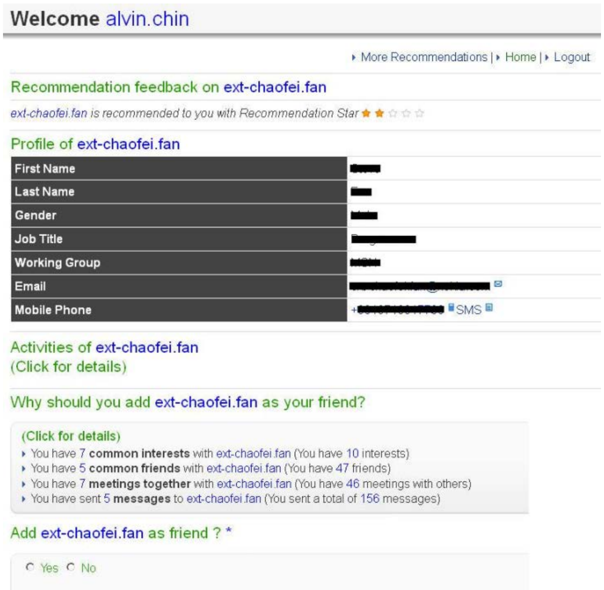

Figure1. Recommendation interface from our workplace application.

4.1.2 Conference

In an academic conference, people decide which sessions and papers they wish to attend, and meet and connect with people that have similar research interests as them. However, the problem is which sessions to go to and who are the people that they should meet? We design a conference application that solves this problem using proximity and homophily that revolves around the conference schedule [9]. We incorporate a friend recommendation system for suggesting contacts to add and to connect with, based on this. Figure 2 shows the list of friend recommendations from our conference application [9]. By selecting any of the people in the list, the person's profile is displayed along with all the things that the user and this person have in common. Note that this inferface is different than the workplace because we want to make it easier for users to use, so that even if other people are not recommended, the user can still see what they have in common with that person and then decide whether to add that person as a friend. For the list of common things that the user and the other person have in common, we show the number of *encounters* (along with the details of the last encounter), number of *common sessions* attended (along with the details of the sessions attended), number of *common contacts* (along with the names of the contacts), and the number of *common interests* (along with the list of the interests).

4.2 Friend Recommendation Algorithm

The list of recommended people comes from our friend recommendation algorithm where we recommend individuals who are ranked in the top 10. The features that we use for our recommendation algorithm are the ones explained earlier in Section 3.2 and we explain how we calculate them below.

Common interests (*ci*): In the user's profile of our application, users can specify their interests from a pre-defined list and if user A and user B have the same interest then user B may likely be a recommended friend for user A. This is similar to the SONAR algorithm of Guy et al [19].

(a) (b) (c)

Figure 2. Recommendation interface from the conference application where (a) shows the place where to select the recommendations, (b) shows the list of recommended people and (c) shows the things in common between the user and the selected person (only shown is encounters, sessions attended and contacts, interests are not shown).

Common friends (*cf*): We use the common friends algorithm which is used in popular social networking sites such as Facebook and LinkedIn's "People You May Know" feature, where if user A and user B both have the same friend, then user B can likely be a recommended friend for user A.

Common meetings (*cm*): These are the meetings that users attend together. Two users have common meetings together if they are meeting participants (in our workplace application), or if they are located in the same room at the same time during a session (in our conference application).

Encounters (*e*): If user A and user B encounter each other several times a day and for several days, then user B may likely be a recommended friend for user A, due to the frequent encounters.

Passby (*p*): If user A and user B pass by each other several times a day and for several days, then user B may likely be a recommended friend. Passby differs from encounters in that in a passby, users approach each other from opposite directions and meet for a very short period of time (less than the encounter duration threshold) before passing each other and going in opposite directions.

Mobile Q&A (*qa*): User A can send a question to a group of people and if any one of those people is online, they can answer that question. If user B answers several questions from user A, then user B may likely be a recommended friend for user A because of the previous social interactions. We only have this feature in the workplace application.

Messages (*m*): Users can send a message to another user. If user B sends several messages to user A, then user B may likely be a recommended friend for user A. We only have this feature in the workplace application.

From all these features, we compute our friend recommendation algorithm below which is a weighted feature algorithm where each feature f above has an associated weight w_f that is customized based on the importance of each feature to the friend recommendation.

We then define a weight vector w_i for each feature f:

$$w_i = \{w_{ci}, w_{cf}, w_{cm}, w_e, w_p, w_{qa}, w_m \mid w_{ci} + w_{cf} + w_{cm} + w_e + w_p + w_{qa} + w_m = 1, 0 < w_f < 1\} \quad (1)$$

for each user U_i that our recommendation algorithm will recommend potential friends to user U.

Next, we define the relevance vector R_i for each feature f:

$$R_i = \{R_{ci}, R_{cf}, R_{cm}, R_e, R_p, R_{qa}, R_m\} \quad (2)$$

for user U_i's relevance to user U in each feature f, where user U_i is not in the friend list of user U. Relevance R_f is measured by the Jaccard similarity of that feature f between U_i and U as

$$R_f = \frac{N_f(U_i \cap U)}{N_f(U_i \cup U)} = \frac{N_f(U_i \cap U)}{N_f(U_i) + N_f(U) - N_f(U_i \cap U)} \quad (3)$$

where N_f refers to the frequency usage or appearance count for that feature. However, other similarity measurements such as Pearson's Coefficient and Cosine Similarity, can be used to define the relevance between U_i and U in each feature space.

We define the recommended score FR_i for recommended friend U_i to user U as:

$$FR_i = w_i \bullet R_i = \{w_{ci}, w_{cf}, w_{cm}, w_e, w_p, w_{qa}, w_m\} \bullet \{R_{ci}, R_{cf}, R_{cm}, R_e, R_p, R_{qa}, R_m\} \quad (4)$$

From ranking the scores, we then recommend the top 10 recommended friends to the user.

Let us illustrate an example for calculating the friend recommendation score from Figure 1. For determining the score for ext-chaofei.fan (U_i), let us assume that number of interests $N_i(U_i) = 12$, number of friends $N_f(U_i) = 10$, number of meetings $N_{meet}(U_i) = 30$, and number of messages sent $N_m(U_i) = 50$. From Figure 1, then $N_{ci}(U_i \cap U) = 7$, $N_{cf}(U_i \cap U) = 5$, $N_{cm}(U_i \cap U) = 7$, $N_m(U_i \cap U) = 5$, $N_i(U) = 10$, $N_f(U) = 47$, $N_{meet}(U) = 46$, and $N_m(U) = 156$. Each w_i has equal weight in our algorithm, therefore $w_f = 1/7$. From the above, then $R_{ci} = 7/15$, $R_{cf} = 5/52$, $R_{cm} = 7/69$, $R_e = 0$, $R_p = 0$, $R_{qa} = 0$, $R_m = 5/201$. The friend recommendation score for ext-chaofei.fan is

$$FR_i = \frac{1}{7} * \frac{7}{15} + \frac{1}{7} * \frac{5}{52} + \frac{1}{7} * \frac{7}{69} + \frac{1}{7} * \frac{5}{201} = 0.098$$

The recommendation star rating of 2 stars shown in Figure 1 is assigned relative in score to the rest of the other 9 highest friend recommendation scores.

Our main objective here is to determine from a user perspective whether friend recommendations based on physical context (e.g. encounters and meetings) are better than friend recommendations based on common friends (which is commonly used in online social networks). To test this hypothesis, we conduct a friend recommendation user study in the office using our system and we present this in the next section below.

5. FRIEND RECOMMENDATION STUDY

To test our friend recommendation algorithm, we conduct two user studies. The first study tests our friend recommendation algorithm with the common friend recommendation algorithm in the workplace. The second study examines the reasons why users add others as friends before an event and during the event, where the event is a conference.

5.1 User Study in the Workplace

We recruit 10 employees in the office who are active users of our system and use our application frequently for booking meetings and have many position updates in the system, 8 of these are male and 2 are female. The study takes 1 hour to complete and participants are asked to perform two tasks on a phone. The first task is to evaluate up to 10 friend recommendations based on common friends, whereas the second task is to evaluate up to 10 friend recommendations based on our friend recommendation algorithm where we use only physical context (encounters and meetings) because we do not want other features other than encounters and meetings to affect the recommendation feedback from the participants, so as not to skew our results. We call our algorithm *EncounterMeet*. Therefore in Equation 4, we use equal weights where $w_m = w_e = 0.5$ and w_{ci}, w_{cf}, w_{qa}, w_p and w_m are all 0. We rank each friend recommendation with a score and present the rank to the user with the number of recommendation stars. For each friend recommendation in each task, subjects are shown the profile of the suggested friend and their activities (e.g. meetings they attended), as well as the reasons for adding the suggested friend as shown in Figure 1. Subjects are asked if the recommended person is a person that they know (in real life, in online social networks, or in their phonebook) and whether the recommendation is good, followed by the option of adding that person as a friend. Table I shows the results from the friend recommendation study and demonstrates that our friend recommendation algorithm *EncounterMeet* performs better, since the percentage of recommendations in each category is higher than the common friend recommendation algorithm.

TABLE I. RESULTS FROM FRIEND RECOMMENDATION STUDY IN THE WORKPLACE.

	Common friend	**EncounterMeet**
# of total recommendations	81	83
Average # of recommendations presented per user	8.1	8.3
% of good recommendations	32.1	44.6
% of recommended persons already known	24.7	37.3
% of recommended persons known in real life	59.4	69
% of recommended persons in phonebook	9.8	13.3
% of recommended persons in SNS	14.8	16.9
% of recommendations accepted	38.3	50.1

For the common friend algorithm, 45.5% of the users provide reasons for why the suggested friend is a good or bad recommendation. Some of the good reasons are: *"I know him from my friend", "We met in a meeting before", I may have been at a dinner evening where she was present", "She's my neighbor and colleague on the same floor", and "We are in the same group"*. The reasons specified indicate that similar profile, social relationships, co-location and physical proximity are factors in providing good recommendations. For our friend recommendation algorithm, 32.6% of the users provide reasons for why the suggested friend is a good or bad recommendation. Some of the good reasons are: "*I am more interested in knowing what type of encounters, and even common interests*", "*I know he's from the MSN team, which is a team I work with a lot*", "*my interactions with X shows the actual amount of time. This is* *important because X is already my friend and I trust his judgment.*" We can see that the reasons here are due to previous meetings, same group, and common content. Therefore, we can clearly see that physical encounters and meetings are important in addition to common friends and similar content and profile, in recommending friends.

For the common friend recommendation task, a strong majority (88%) of the people mentioned that why the recommended friend was not good was because they simply did not know that person. In particular, comments such as "*I have no idea why this person is recommended to me? What are the common interests?*" and "*I have no idea who she is. And there is very little info in her profile.*" were common responses. For our friend recommendation algorithm *EncounterMeet*, 46.7% of the people mentioned that the recommended friend was not good because they also did not know that person. Some of the reasons for not recommending the friend include: "*It should be a better recommendation. I know X already and enjoy interacting with him, but the system doesn't know that*", "*interaction distance is too long. Actual interaction time very short. Don't know why recommended*", "*I know her already, and often interact with her but she doesn't have a high recommendation.*" This means that in order to make recommendations more acceptable, the system still needs to record the social interactions, activity and social context as the evidence for recommendations.

We also see that from the accepted friend recommendations in Figure 3, our algorithm provided a higher percentage of good recommendations than the common friends algorithm (81% vs. 72%) and a greater percentage of accepted recommendations are known in real life (69% vs. 59%). Both algorithms provided an equal percentage of accepted recommendations that are known in SNS (social network service such as Facebook) and are in the user's phonebook. From Figure 3, a large majority of accepted friend recommendations are due to knowing the recommended person in real life (not unexpected), followed by an almost equal percentage known in SNS and in the user's phonebook.

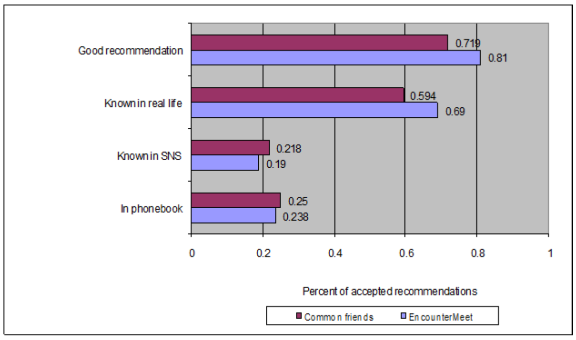

Figure 3. Percent of accepted friend recommendations based on acquaintance for common friends and EncounterMeet recommendation algorithms.

By looking at the percentage of recommended friends that the user knows, we see that for the most part, good recommendations as rated by the user, are accepted as friends. A high majority of recommended friends that are known in real life and in a user's phonebook are accepted and added as friends, and the percentage is nearly the same for each friend recommendation algorithm.

Therefore, acquaintance does affect whether a user will add this recommended friend, with real connections (contacts in phonebook and known in real life) providing the strong incentive, and is regardless of the type of friend recommendation algorithm.

5.2 User Study in the Conference

In this user study, we examine why people add others as friends in a conference environment, by using a survey before the conference and using our conference application (using encounters and meetings) during the conference, in order to determine if encounters and meetings impacts the selection process for adding friends. For the survey, we gather 29 people and ask them to select the reasons why they add others as friends using the criteria in Table II. For the conference application, using the same criteria, we also ask participants to select the reasons for adding a contact when they add contacts in our conference application (we gather a total of 571 contact relationships), to evaluate the responses. All the results are summarized in Table II.

TABLE II. REASONS FOR ADDING FRIENDS/CONTACTS FROM SURVEY BEFORE CONFERENCE AND FROM OUR APPLICATION AT THE CONFERENCE

Reason for adding friends/contacts	Survey	Our conference application	Rank (survey)	Rank (Our conference application)
Encountered before	59%	37%	2	2
Common contacts	48%	12%	3	5
Common research interests	24%	35%	5	3
Common sessions attended	7%	24%	7	4
Know each other in real life	69%	39%	1	1
Know each other online	34%	9%	4	6
Added each other as phone contact	21%	4%	6	7

From the table, we can see that the top 2 reasons for adding friends/contacts from both the survey and our conference application are the same, which are that they know each other in real life and have encountered before. This validates that friend recommendations based on encounters and meetings are useful and helps in the decision to add new friends or contacts. We also discover that having common contacts is an important reason for adding a person as a friend/contact as well as having common research interests, validating the homophily principle and social influence theory of 3 degrees of separation [6]. However, three interesting results emerge. First, common sessions that a user and another person have attended before, do not affect that user from adding that person as a friend in an online social network, however it becomes a factor during the conference when you can add that person as a friend/contact directly with our conference application. Second, to our surprise, knowing a person online does not have much of an effect on adding that person as a friend/contact in the online social network or the conference social network. Third, another surprise is that people that are in a user's phonebook are not the primary reasons for adding them as friends/contacts in online social networks or in our application. Perhaps the explanation for this is because contacts in a phonebook can be considered as strong ties and online social networks are mainly used for weak ties as evidenced by previous researchers [15], and also due to privacy concerns. Phonebook contacts are private and personal and many users do not want to share them online with others, therefore users do not want to cross the offline/online social networking boundary.

In summary, our friend recommendation algorithm used in the workplace is better than the common friend recommendation algorithm, and people primarily add others as friends due to having met them before at an event or encountering them. Thus, proximity and homophily are reasons why people usually add others as friends.

6. CONCLUSIONS

In this paper, we present a novel friend recommendation algorithm, interface and system that uses proximity and homophily that is designed for indoor environments. The recommendation uses encounters, passbys and meetings for proximity; common interests and common friends for homophily; and messages sent and question and answer posts for social interaction. We test our friend recommendation algorithm in the workplace and conference where we present users with a friend recommendation interface that shows the reasons why they should add the other person as a friend. Our hypothesis is correct in that the quality of friend recommendations based on physical context is better than those based on common friends.

Results show that our friend recommendation algorithm in the workplace recommends more friends to participants that they add, more recommendations are ranked as good, and more have previous acquaintance with these recommendations, compared with the common friend algorithm. For the conference, we discover that the top 2 reasons for adding friends/contacts are that they know each other in real life and have encountered before. This validates that friend recommendations based on encounters and meetings are useful and helps in the decision to add new friends or contacts.

One of the limitations of our work is that the sample size for the recommendation study is small, so for future work we would like to repeat the recommendation study with a larger population. In addition, the weights that we use for our friend recommendation algorithm are the same. For future work, we plan to modify the recommendation algorithm so that it learns from the user's past history of recommendations and recommendation feedback, thus automatically adjusting the feature weights (e.g., meetings is more important for adding another person as a friend vs. common friends between you and the other person), as well as adding content-based features in order to provide personalized weighted features. To improve on the recording of proximity interactions, we suggest using a peer-to-peer wireless technology such as peer-to-peer RFID tags or NFC tags that users can carry and can touch another NFC tag to actually record a real proximity interaction, rather than a centralized server-based system which has the possibility for inaccurate proximity calculations.

7. ACKNOWLEDGMENTS

We thank all the participants for using our system in the workplace and conference, and for participating in our user studies and providing us feedback. We also thank our development team for developing and maintaining the system. Without this, our research would not have been possible.

8. REFERENCES

[1] aka-aki networks gmbh: aka-aki. 2007. http://www.aka-aki.com/.

[2] Backstrom, L., Huttenlocher, D., Kleinberg, J., and Lan, X. 2006. Group formation in large social networks: membership, growth, and evolution. In *Proceedings of ACM SIGKDD* 2006, 44-54.

[3] Backstrom, L., Sun, E., and Marlow, C. 2010. Find me if you can: improving geographical prediction with social and spatial proximity. *In Proceedings of ACM WWW* 2010, 61-70.

[4] Barrat, A., Cattuto, C., Szomszor, M., Van den Broeck, W., and Alami, H. 2010. Social dynamics in conferences: analyses of data from the live social semantics application. In *Proc. The Semantic Web-ISWC 2010*, 17-33.

[5] Browning, R.C., Baker, E.A., Herron, J.A., and Kram, R. 2006. Effects of obesity and sex on the energetic cost and preferred speed of walking. *Journal of Applied Physiology* 100, 2, 390398.

[6] Cacioppo, J.T., Fowler, J.H., and Christakis, N.A. 2009. Alone in the crowd: the structure and spread of loneliness in a large social network. *Journal of Personality and Social Psychology* 97, 6, 977.

[7] Chen, J., Geyer, W., Dugan, C., Muller, M., and Guy, I. 2009. Make new friends, but keep the old: recommending people on social networking sites. In *Proc. ACM CHI*, ACM Press, 201-210.

[8] Chin, A., and Wang, H. 2010. Using Cohesive Subgroups for Analyzing the Evolution of the Friend View Mobile Social Network. In Proc. *IEEE UIC2010*, Springer.

[9] Chin, A., Xu, B., Yin, F., Wang, X., Wang., W., Fan, X., Hong, D., and Wang, Y. 2012. Using Proximity and Homophily to Connect Conference Attendees in a Mobile Social Network. In *Proc. 2nd International Workshop on PhoneCom 2012*, Distributed Computing Systems Workshops (ICDCSW), ICDCS 2012, IEEE Computer Society, 79-87.

[10] Counts, S., and Geraci, J. 2005. Incorporating physical co-presence at events into digital social networking. In *Proc. ACM CHI 2005*, ACM Press, 1308-1311.

[11] Cranshaw, J., Toch, E., Hong, J., Kittur, A., and Sadeh, N. 2010. Bridging the gap between physical location and online social networks. In *Proc. ACM UbiComp 2010,* ACM Press, 119-128.

[12] Debnath, S., Ganguly, N., and Mitra, P. 2008. Feature weighting in content based recommendation system using social network analysis. In *Proc. ACM WWW 2008*, ACM Press, 1041-1042.

[13] Eagle, N., and Pentland, A. 2005. Social serendipity: Mobilizing social software. *IEEE Pervasive Computing 4*, 2, 28-34.

[14] Froehlich, J., Chen, M., Smith, I., and Potter, F. 2006. Voting with your feet: An investigative study of the relationship

between place visit behavior and preference. In *Proc. ACM UbiComp 2006*, ACM Press, 333-350.

[15] Garton, L., Haythornthwaite, C., and Wellman, B. 1997. Studying online social networks. *Journal of Computer-Mediated Communication* 3, 1.

[16] Gilbert, E., and Karahalios, K. 2009. Predicting tie strength with social media. In *Proc. ACM CHI 2009*, ACM Press, 211-220.

[17] Guy, I., Jacovi, M., Shahar, E., Meshulam, N., Soroka, V., and Farrell, S. 2008. Harvesting with SONAR: the value of aggregating social network information. In *Proc. ACM CHI 2008*, ACM Press, 1017-1026.

[18] Guy, I., Zwerdling, N., Carmel, D., Ronen, I., Uziel, E., Yogev, S., and Ofek-Koifman, S. 2009. Personalized recommendation of social software items based on social relations. In *Proc. ACM RecSys 2009*, ACM Press, 53-60.

[19] Guy, I., Jacovi, M., Perer, A., Ronen, I., and Uziel, E. 2010. Same places, same things, same people?: mining user similarity on social media. *In Proceedings of ACM CSCW 2010*, 41-50.

[20] Guy, I., Ur, S., Ronen, I., Perer, A., and Jacovi, M. 2011. Do You Want to Know? Recommending Strangers in the Enterprise. In *Proc. ACM CSCW 2011*, ACM Press, 285-294.

[21] Hall, T. 1963. A system for the notation of proxemic behavior. *American Anthropologist* 65, 10031026.

[22] Herlocker, J.L., Konstan, J.A., and Riedl, J. 2000. Explaining collaborative filtering recommendations. In *Proc. ACM CSCW 2000*, ACM Press, 241-250.

[23] Isella, L., Stehle, J., Barrat, A., Cattuto, C., Pinton, J.F., and Van den Broeck, W. 2010. What's in a crowd? Analysis of face-toface behavioral networks. *Journal of Theoretical Biology*.

[24] Jones, Q., Grandhi, S. A., Terveen, L., and Whittaker, S. 2004. People-to-people-to-geographical-places: the P3 framework for location-based community systems. In *Proc. ACM CSCW 2004*, ACM Press, 249-282.

[25] Liiv, I., Tammet, T., Ruotsalo, T., and Kuusik, A. 2007. Personalized Context-Aware Recommendations in SMARTMUSEUM: Combining Semantics with Statistics. In *Proc. SEMAPRO 2007*, IEEE Computer Society, 50-55.

[26] Lo, S., and Lin, C. 2006. WMR--A Graph-Based Algorithm for Friend Recommendation. In *Proc. IEEE/WIC/ACM WI 2006*, IEEE Computer Society, 121-128.

[27] McPherson, M., Smith-Lovin, L., and Cook, J.M. 2001. Birds of a feather: Homophily in social networks. *Annual review of sociology* 27, 415-444, 2001.

[28] Moody, J. 2001. Race, school integration, and friendship segregation in america. *American Journal of Sociology* 107, 679-716.

[29] Moricz, M., Dosbayev, Y., and Berlyant, M. 2010. PYMK: friend recommendation at myspace. In *Proc. SIGMOD* 2010, ACM Press, 999-1002.

[30] Paulos, E., and Goodman, E. 2004. The familiar stranger: anxiety, comfort, and play in public places. In *Proc. of the SIGCHI conference on Human factors in computing systems (CHI '04)*. ACM, New York, NY, USA, 223-230.

[31] Quercia, D., and Capra, L. FriendSensing: recommending friends using mobile phones. In *Proc. ACM RecSys 2009*, ACM Press, 273-276.

[32] SONAR. http://www.sonarapp.com/.

[33] Terveen, L., and McDonald, D. W. 2005. Social matching: A framework and research agenda. *ACM Transactions on Computer-Human Interaction (TOCHI)* 12, 401-434.

[34] Terry, M., Mynatt, E. D., Ryall, K., and Leigh, D. Social net: using patterns of physical proximity over time to infer shared interests. In *Proc. ACM CHI 2002*, ACM Press, 816-817.

[35] Tintarev, N. 2010. Explaining recommendations. PhD Thesis, University of Aberdeen, UK.

[36] Wang, H. , Chin, A., and Wang, H. 2011. Interplay between Social Selection and Social Influence on Physical Proximity in Friendship Formation, *2nd International Workshop on Social Recommender Systems*, The 2011 ACM Conference on Computer Supported Cooperative Work (CSCW 2011), Hangzhou, China, March 18, 2011.

[37] Woerndl, W., and Groh, G. Utilizing Physical and Social Context to Improve Recommender Systems. 2007. In *Proc. IEEE/WIC/ACM WI-IATW 2007*, ACM Press, 123-128.

[38] Zhu, L., Chin, A., Zhang, K., Xu, W., Wang, H., and Zhang, L. 2010. Managing Workplace Resources in Office Environments through Ephemeral Social Networks", In *Proc. of the 7th international conference on Ubiquitous intelligence and computing (UIC'10)*, Springer-Verlag, Berlin, Heidelberg, 665-679.

[39] Zuo, X., Chin, A., Fan, X., Xu, B., Hong, D., Wang, Y., and Wang, X. 2012. Connecting People at a Conference: A Study of Influence Between Offline and Online Using a Mobile Social Application. In *Proc. of the 5th IEEE International Conference on Cyber, Physical and Social Computing(CPSCom 2012)*, 8 pages, 20-23 Nov. 2012, Besancon, France, IEEE Computer Society.