GEVR: An Event Venue Recommendation System for Groups of Mobile Users

JASON SHUO ZHANG, University of Colorado Boulder, USA
MIKE GARTRELL, University of Colorado Boulder, USA
RICHARD HAN, University of Colorado Boulder, USA
QIN LV, University of Colorado Boulder, USA
SHIVAKANT MISHRA, University of Colorado Boulder, USA

In this paper, we present GEVR, the first Group Event Venue Recommendation system that incorporates mobility via individual location traces and context information into a *"social-based"* group decision model to provide venue recommendations for groups of mobile users. Our study leverages a real-world dataset collected using the OutWithFriendz mobile app for group event planning, which contains 625 users and over 500 group events. We first develop a novel *"social-based"* group location prediction model, which adaptively applies different group decision strategies to groups with different social relationship strength to aggregate each group member's location preference, to predict where groups will meet. Evaluation results show that our prediction model not only outperforms commonly used and state-of-the-art group decision strategies with over 80% accuracy for predicting groups' final meeting location clusters, but also provides promising qualities in cold-start scenarios. We then integrate our prediction model with the Foursquare Venue Recommendation API to construct an event venue recommendation framework for groups of mobile users. Evaluation results show that GEVR outperforms the comparative models by a significant margin.

34

CCS Concepts: • **Human-centered computing → Ubiquitous and mobile computing**.

Additional Key Words and Phrases: Group Recommendation, Event Venue, Group Location Cluster, User Mobility

ACM Reference Format:
Jason Shuo Zhang, Mike Gartrell, Richard Han, Qin Lv, and Shivakant Mishra. 2019. GEVR: An Event Venue Recommendation System for Groups of Mobile Users. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 3, 1, Article 34 (March 2019), 25 pages. https://doi.org/10.1145/3314421

1 INTRODUCTION

Online venue recommendation services, such as Foursquare, Google Maps, and Yelp, demonstrate the importance placed on recommending venues to an individual mobile user, e.g., what coffee shop or restaurant should that person visit next. However, such venue recommendation does not address the larger challenge of recommending venues to a group of mobile users who are trying to choose a place to meet, e.g., friends and colleagues trying

Authors' addresses: Jason Shuo Zhang, jasonzhang@colorado.edu, University of Colorado Boulder, 1111 Engineering Dr, Boulder, CO, 80309, USA; Mike Gartrell, Mike.Gartrell@colorado.edu, University of Colorado Boulder, 1111 Engineering Dr, Boulder, CO, 80309, USA; Richard Han, richard.han@colorado.edu, University of Colorado Boulder, 1111 Engineering Dr, Boulder, CO, 80309, USA; Qin Lv, qin.lv@colorado.edu, University of Colorado Boulder, 1111 Engineering Dr, Boulder, CO, 80309, USA; Shivakant Mishra, shivakaht.mishra@colorado.edu, University of Colorado Boulder, 1111 Engineering Dr, Boulder, CO, 80309, USA.

2474-9567/2019/3-ART34 $15.00
https://doi.org/10.1145/3314421

to decide on a restaurant for lunch or dinner. In this case, the members of the group may have a wide range of behaviors and preferences and be scattered across many locations.

This paper addresses this topic of recommending event venues to *groups of mobile users*, a service that we expect will become increasingly important as people get more and more digitally connected and organize their social gatherings via mobile technologies. While effective recommendation to individual users is already a challenging task, developing event venue recommendations for groups of mobile users is significantly more challenging due to a number of reasons. Groups of users often have a wide variety of behaviors and preferences which must be resolved to achieve the best recommendations to the group. In addition, different members of a group may be at varied locations, so mobility considerations should be taken into account when recommending a place to meet. Groups are also dynamic rather than static, and so group membership may change fluidly from event to event, unlike individual recommendation for the same target user.

Prior studies in group-based recommendation has addressed the issue of resolving differing preferences of individual members of the group in order to provide the best recommendation [8, 15, 32]. However, the role that location and mobility of different group members may play in influencing the group's final choice has not been addressed in these articles. Some scant prior research in group event venue recommendation [6, 31] has been confined only to in-lab surveys and has not studied the real-world dynamics of groups of mobile users, nor its impacts on group decision-making. We believe this is an exciting area ripe for exploration by the ubiquitous computing research community. Due to the explosion of people scheduling group events online using smart devices, we expect mobile and context-aware computing to be used extensively to reduce the friction inherent in coordination and assist groups in making decisions.

As the first step towards this direction, we present GEVR, a novel Group Event Venue Recommendation system that tackles the recommendation challenge by incorporating mobility patterns and social relations among group members. To tackle the complexity of group recommendation, GEVR splits the recommendation process into three separate steps:

(1) **Group location cluster detection:** Using individual group members' historical location traces, GEVR first identifies location clusters where the group members have colocated in the past.
(2) **Group event location cluster prediction:** After identifying these group location clusters, GEVR uses a newly designed social-based group location cluster prediction model to estimate which location cluster the group is more likely to meet at.
(3) **Event venue recommendation:** Finally, GEVR aggregates restaurants near group members' frequented locations using the Foursquare Venue Recommendation API [14], and re-ranks them based on the prediction results calculated by Step (2) to build a final top-K recommendation list.

Evaluation results show that our prediction model achieves over 80% accuracy for predicting a final meeting location cluster. Furthermore, through the combination of the group location clusters identified, cluster prediction results and Foursquare Venue Recommendation API, GEVR achieves promising recommendation performance and outperforms the baseline models by a significant margin. A key contribution of this paper is that the efficacy of GEVR has been evaluated using a large-scale, real-world dataset that was collected from a mobile application called OutWithFriendz [41], which supports group event scheduling by enabling users to propose event venues and dates, invite people to join the group, and vote on venues and dates. The finalized event location and date are confirmed through location tracking in the app.

In summary, we make the following fundamental contributions in this work:

- We introduce GEVR, the first system that provides event venue recommendation to groups of mobile users. We present a novel algorithm for recommending event venues to groups of mobile users that re-prioritizes the results of the Foursquare Venue Recommendation API according to group properties, namely the likelihood of the venue residing in a location *cluster* that the group is likely to visit.

- To generate a ranking of each group's preferences for its location clusters, we separate the problem into first generating individual group members' preferences for those clusters, and then combining the individual preferences into a group ranking via a group decision strategy.
- Our analysis reveals that location familiarity, individual activeness, day of the week, and population density are strong predictors of individual cluster preferences.
- We construct an innovative *"Edge-RWR+CART"* model to incorporate these identified features for predicting individual location preference. It is not only demonstrated to be more superior to the comparative models but also flexible for adding new contextual features.
- We design a novel *"social-based"* group location prediction model, which adaptively applies different group decision strategies to groups with different social relationship strength to aggregate group members' individual location preferences.

- We collect data that is difficult to attain concerning which venues groups choose to meet at, obtaining over 500 real-world group events by over 600 users from a mobile application over one year, thereby establishing ground truth on which venues these groups chose for their gatherings.
- We evaluate the effectiveness of GEVR using our real-world dataset. Detailed evaluation results show that using our *"social-based"* group location prediction model outperforms all the commonly used and state-of-the-art baselines with a significant margin, even in the cold-start scenarios.

The remainder of this paper is organized as follows. In the next section, we discuss related work. In Section 3, we review the functions of the group event mobile application relevant to our dataset collection, and the problem formulation. Then in Section 5, we present detailed analysis of group location cluster detection and parameter selection using collected user location traces. The design of our social-based group location prediction model is introduced in Section 6, and the framework of our group event venue recommendation is described in Section 7. Performance evaluations are presented in Section 8. Finally, we highlight the key findings, discuss important implications and future improvements, and conclude this paper.

2 RELATED WORK

The problem of making recommendations to a group of users has been investigated in a number of works [8, 15, 32]. Gartrell *et al.* propose a group recommendation model that utilizes social and content interests of group members to recommend movies for groups of users [15]. The work by Cheney *et al.* presents a large-scale study of television viewing habits. They provide an analysis of how the viewing patterns shift across various group contexts and discussed the impact of these findings on the performance of a group TV program recommendation system [8]. Quintarelli *et al.* further models users' ability to direct a group's decision by using contextual influence and aggregate it to recommend TV programs for groups of users [32]. All of these works can be summarized as recommending items for groups of users, where user mobility is not much of a concern. Our problem, group event venue recommendation is more challenging as users' mobility patterns can play a significant role in group decision making. Before the actual event gathering, group members may be distributed in different geographical areas. Thus, recommending a venue that will optimize group satisfaction is more challenging and requires delicate data analysis and system design.

Our work is also generally related to the research topic of understanding the relationship between online and offline social behaviors in groups. Lane *et al.* consider various social phenomena and environmental factors and design a Networked Community Behavior framework which can exploit community-scale behavior patterns [22]. Zhang *et al.* studies how does team performance impact fan behavior on Reddit NBA fan groups [44]. The work by Park *et al.* introduce a social context-aware smart-phone notification system to help groups of users focus better on in-person social interactions [30]. Jayarajah *et al.* investigate how users deal with group co-presence to prevent conflictive situations on campus [21]. There are also a few works studying Doodle [34, 35, 47], an online event scheduling service, where the host can only suggest the meeting times for group members to vote, not venues. Our work adds to this literature by demonstrating the effects of different key factors on group event attendance and how to aggregate individual location preferences to model group decision-making.

Earlier work in event organization has investigated a series of contextual features that would impact a user's decision to attend a group event [11, 16, 26, 42, 43]. Most of these studies rely on datasets collected from popular group event services, such as Meetup, [1] Evite, [2] and Douban Events. [3] This line of research still falls into individual recommendation as the major task is to match potential users who are interested in the event. Very little existing research has explored the subject of recommending event venues for groups of mobile users.

Finally, the proliferation of location-based social services in recent years has enabled a rich body of research in understanding individual movement patterns and predicting when and where people will go next, such as next place prediction [5, 28], trip planning [25, 46], ride-sharing management [1, 9, 10], and individual venue recommendation [23, 29].

3 DATASET

The dataset we use in this work is collected using the OutWithFriendz mobile application, which enables group hosts to organize events and negotiate when and where to meet with group members through an embedded voting process [41]. The app is implemented as a client-server architecture that is comprised of both Android and iOS based clients. When creating an event, the host can specify the details of the invitation, including a title, list of suggested meeting venues, times, and invited participants. After the host creates and submits a new invitation, all participants will receive a notification. Then the event scheduling process begins. Every invitee can vote for their preferred venues and times, and suggest more options. After the group consensus process ends, the host decides the final meeting venue and time. To support later negotiation, the host can still update the final venue and time after it is finalized. In the design of the OutWithFriendz mobile application, the host can make decisions based on the voting results, but he/she does not always have to obey them.

The whole dataset contains group event data collected from January 2017 to December 2017. We use two methods for data collection: (1) We advertise the mobile application in our university's student center, dining rooms, and departmental lobbies. (2) We post task assignments on two crowdsourcing platforms: Microworkers [40] and Craigslist [39]. For the users of crowdsourcing platforms, \$10–20 is paid to the group host for completing a legitimate event using OutWithFriendz, with the provisions that:

(1) The host must live in the US.
(2) The host should invite at least two other members to the invitation using the mobile app.
(3) The meeting venue and time must be finalized.
(4) Each group member should open the location service on his/her smart phone during the event scheduling process and allow location traces to be collected by the mobile app.
(5) At least half of the group members should attend the finalized event. We manually confirm this using location traces before approving workers' assignments on crowdsourcing platforms.

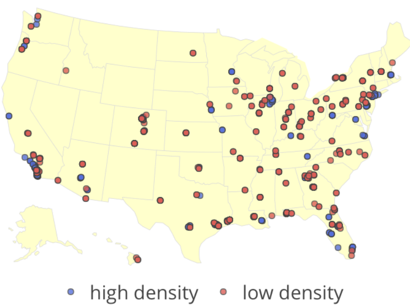
A user may create or join multiple invitations during this study. To collect user mobility data, the OutWithFriendz mobile app posts users' GPS location traces to the server every five minutes when the mobile app is running in the background or every 30 seconds when the app is running in the foreground. The experimental protocol is reviewed and conducted under IRB Protocol #12-0008.

Since our GEVR system utilizes users' location traces to provide recommendation for groups of users, to ensure sufficient data coverage, we remove group events that have members with less than three days of location

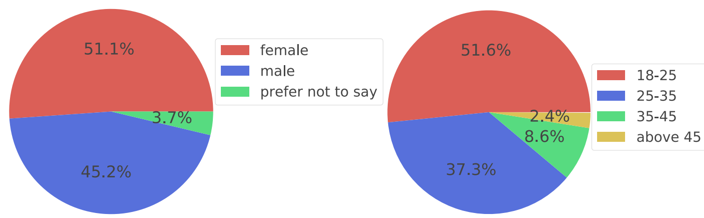
[1] https://www.meetup.com/
[2] https://www.evite.com/
[3] https://beijing.douban.com/events/

Table 1. Dataset Statistics. The users in our dataset are from two sources. Campus users are advertised from our university and volunteered to use the mobile app. Crowdsourcing users are recruited from Microworkers and Craigslist. Our users are all over the US, across 40 states and 117 cities.

	Total	Campus	Crowdsourcing
Number of users	625	75	550
Number of events	502	151	351



(a) All the suggested venues across the US.



(b) User demographics.

Fig. 1. Figure 1a shows the geographic distribution of all the suggested venues across the US. Blue marks represent places of high-density areas and red in low-density areas. Figure 1b shows the demographics (gender and age) of 341 users in the dataset who fill out the anonymized user demographic questionnaire.

trace data before the invitation was created. Also, we remove group events in which the final meeting venue is not a restaurant. [4] The basic statistics of our dataset after the filtering are shown in Table 1. The geographic distribution of all suggested venues across the US is projected in Figure 1a. All the venues are further divided into low density and high density areas based on the population density of the venue location. Here the population density information is collected from the 2016 US area development degree data from the US Census Bureau [7]. Our users are widespread over the US, covering 40 states and 117 cities. After the user study, we conduct an anonymized user demographic questionnaire and 341 study participants completed it. Figure 1b shows these users' demographic information. We also observe that all of the groups in our study are of either small or medium sized, with group sizes of 3–6. Among them, 64% of the groups have three members, 19% of the groups have four members, 12% of the groups have five members, and the remaining 5% have six members.

4 FRAMEWORK OVERVIEW

Given all the information collected from the mobile app, to build an effective event venue recommendation system for groups, we tackle the challenges in three steps:

Step 1: Group location cluster detection. In this step, our objective is to gather location trace points of group members into location clusters. The detected location clusters should be able to represent group members' frequented locations. The details of group location cluster detection will be introduced in Section 5. The venues are searched and selected using the integrated Google Maps API.

[4]Groups can organize any types of events using OutWithFriendz, including going to the gym, watching a movie at the theater, or participating in outdoor activities. In this paper, we focus on dining events, as it is the dominating type in the dataset: 81% of the events are dining events.

Step 2: Group location cluster prediction. Using the group location clusters detected from Step 1, our goal of this step is to predict which location clusters the group is likely to meet at. More specifically, we model the probability of each detected location cluster being selected as the final meeting location cluster of the group. The design of our prediction model will be explained in Section 6.

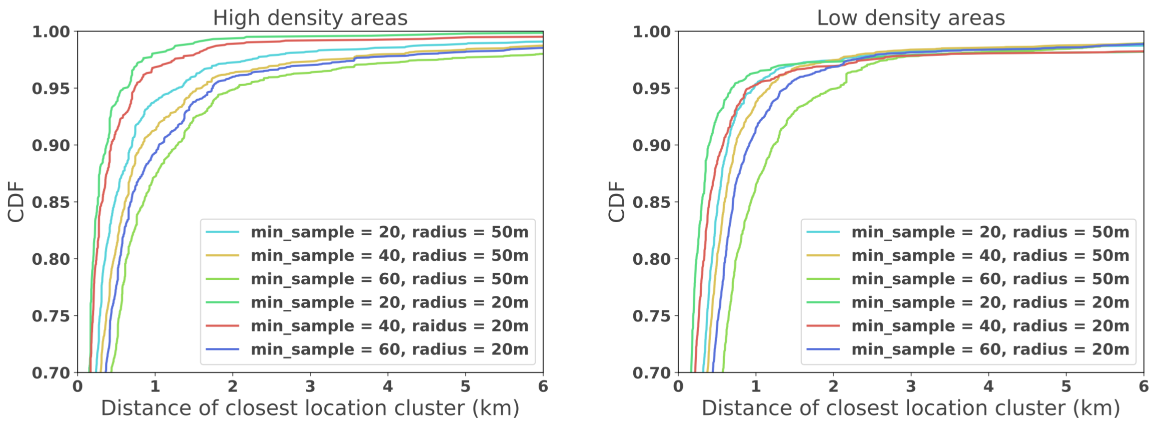
Step 3: Event venue recommendation for groups of mobile users. In this step, our goal is to recommend a list of restaurants for each group event. We use the Foursquare Venue Recommendation API to create a candidate restaurant pool near every detected location cluster generated by Step 1, and re-rank the candidate pool using the prediction results calculated by Step 2 to build a final top-N restaurant recommendation list. The construction of our group event venue recommendation framework will be described in Section 7.

To make effective event venue recommendation for groups of mobile users, the ultimate goal is to reduce and prioritize the venue recommendation list. The three steps gradually build towards this goal. We first use Step 1 to detect group members' frequented geographic areas and concentrate our recommendation on these areas instead of the whole city. After that, by modeling group members' location preferences and aggregating them using the group decision strategy in Step 2, we estimate the likelihood of each location cluster that the group will decide to meet at. Finally, in Step 3, we re-rank the candidate venues in concentrated areas by the estimated likelihood of each location cluster.

5 GROUP LOCATION CLUSTER ANALYSIS

Given a group with multiple mobile users and their location traces, our first goal is to identify *group location clusters*, i.e., concentrated regions that group members have visited. Intuitively, a location cluster may represent home, work, or a district, say downtown, near a mall, or a zoned business area where there are many restaurants. Due to the fact that user location data recorded by mobile applications usually contains errors of a few feet, it is impractical to use GPS points in location traces directly for modeling. A common approach in user mobility studies is to cluster nearby geolocation points into location clusters [3, 9, 19] and use them instead of location points when predicting human movements. Inspired by these papers, to generate group location clusters, we first combine all group members' temporally ordered location trace points and apply the widely-used DBSCAN [13] location clustering algorithm to detect group location clusters, which represent significant places that the group members visit frequently. More specifically, every location point is represented by its geographical coordinate pair (latitude and longitude), and fed directly into the DBSCAN clustering algorithm, which identifies dense neighborhoods and groups location points that are very close to each other into the same cluster to represent a location. After the clustering procedure, the ideal situation is that every location point detected from group members' location traces is assigned to a specific location cluster; nearby location points would belong to the same location cluster, and each cluster is relatively separated from each other on the map. We further represent each group location cluster using the centroid of all the location trace points that belong to that cluster. To ensure that the DBSCAN algorithm is able to find sensible and meaningful location clusters using our dataset, and the group's final meeting venue falls within one of these clusters, it is important to analyze and select appropriate values for the parameters.

The main parameters of DBSCAN are min_sample and $radius$. Min_sample is the minimum number of points required to form a dense region, and $radius$ is the maximum distance allowed between two samples for them to be considered as in the same neighborhood. To determine a reasonable combination of min_sample and $radius$ parameter values, we tried different combinations and examined whether groups' final event venues fall within one of the group location clusters detected. Each group location cluster is represented by the centroid point of all location trace points that belong to that cluster. Figure 2 shows the cumulative distribution function (CDF) of distances between a group event's final venue and the closest group location cluster using different combinations of min_sample and $radius$ values. For comparison purpose, all groups are also divided equally into low density and



(a) Distribution of distances in high density areas.

(b) Distribution of distances in low density areas.

Fig. 2. Figure 2a and Figure 2b show the cumulative distribution function of distances between each group event's final venue and the closest group location cluster for groups in high density and low density areas, respectively. As can been seen in the figures, when setting $min\_sample = 40, radius = 20meters$, more than 90% of group events' final venues fall within the closest group location cluster (within 500 meters).

high density areas based on the population density of their centroid points. The population density information we use is collected from the 2016 US area development degree data from the US Census Bureau [7].

We find that in both high density and low density areas, when setting $min\_sample = 40, radius = 20\ meters$, more than 90% of the final event venues fall into the closest clusters (within 500 meters). When setting $min\_sample = 20, radius = 20\ meters$, locations such as highway transitions may fall within the same cluster. The other combinations fail to cover a significant proportion of group event venues in any of the group location clusters. We further confirm that by choosing $min\_sample = 40, radius = 20\ meters$, the DBSCAN algorithm is able to find sensible clusters in groups' location traces based on a visual inspection of these clusters plotted on a map. These location clusters appear to correspond to locations frequented visited by our participants, such as work, school, home, etc. Based on this analysis, we set the min_sample value as 40 and the $radius$ value as 20 meters in our group location cluster detection task.

It is also worth noting that in roughly 10% of the group events, the groups' final meeting venues did not fall into any of the group clusters detected. One reasonable explanation is that some group members may have been to this location some time ago but not recently. For example, users may have been to Costco two weeks ago, which is not covered by more recent location traces. It is also possible that groups of users decided to explore a new venue that they had never tried before. We plan to investigate this problem in more detail in the future when we collect more data.

6 SOCIAL-BASED GROUP LOCATION CLUSTER PREDICTION MODEL

Given the group location clusters determined in the previous section, our goal in this section is to design a statistical model to predict which location cluster the group is going to meet at. Motivated by prior work in group recommendation [8, 15], we divide this problem into two steps. First, we analyze strong predictors that will impact group member's meeting decisions and construct an innovative *"Edge-RWR+CART"* model to incorporate these identified features for predicting individual location preference. After that, we aggregate group members'

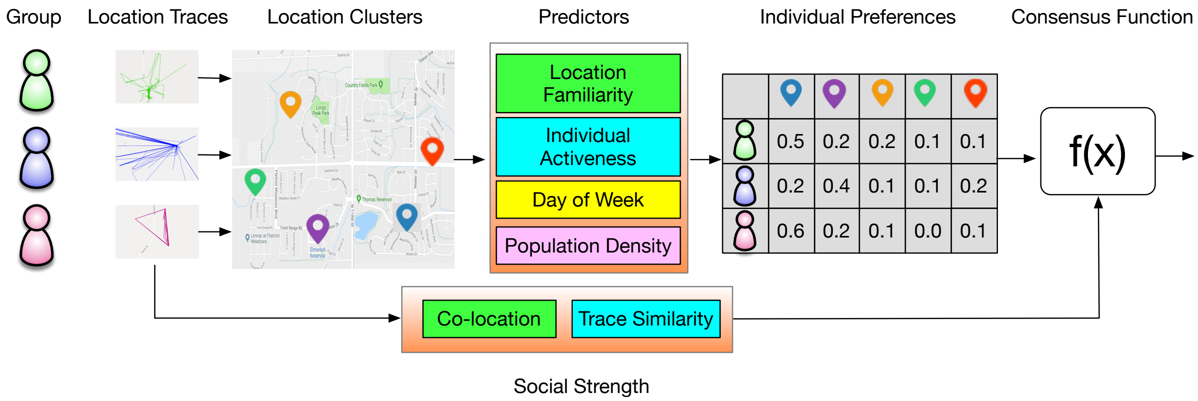


Fig. 3. The overall architecture of our group location prediction model. With the detected group location clusters, we model each group member's individual location cluster preference based on his/her location traces. Meanwhile, the location traces are also used for estimating the group's social relationship strength. After that, a newly designed consensus function that we term a *"social-based"* group decision strategy is used to aggregate individual members' location cluster preferences and predict the group's final decision.

location preferences to predict group decisions. Earlier study also suggests that the group decision strategy varies depending on the strength of social connections among group members [15]. Inspired by this observation, in the context of group event venue recommendation, we propose a novel *"social-based"* group location prediction model, which adaptively applies different group decision strategies to groups with different social relationship strength to aggregate group members' individual location preferences.

6.1 Overall Architecture of the Model

The overall architecture of our prediction model is illustrated in Figure 3. As shown in the figure, individual group members' location traces are first merged to generate the group location clusters shown on the map. We then determine each individual's likelihood of visiting each location cluster, as shown in the table on the right side of the figure. This is accomplished by performing an analysis of the key factors influencing each individual's likelihood of visiting a location cluster, including location familiarity, individual activeness, day of the week, and population density. After determining the individual likelihoods of visiting a location cluster, we then measure the social strength of the group in order to combine the individual location preferences into a prediction of the group's likelihood of visiting each location cluster. This is depicted as the group consensus function $f(x)$ in the figure, along with the two social strength factors that we will use, co-location and trace similarity.

6.2 Social Relationship Strength

The social relationships among group members play a significant role in group decision making. Consider the following scenario: Tony and Paul are in the same group voting on a location for dinner on Friday night, and Paul does not like spicy food. If Tony knows Paul very well, he may not suggest or vote for Thai restaurants that are famous for spicy cuisine, even though he himself loves spicy food. A strong social relationship with Paul helps Tony to propose better meeting options. Here we estimate a group's social relationship strength using two factors:

- *Average number of daily meetings detected between any two group members using their location traces*: The more often two users meet in real life, the more likely that they know each other well. We use the following

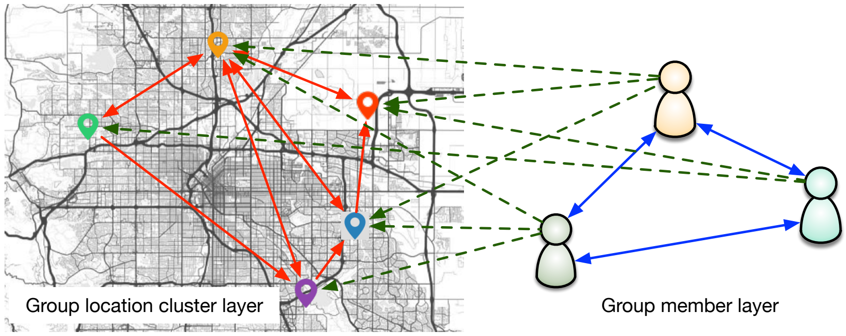


Fig. 4. An illustration of a group location cluster-group member graph. There are two layers in the graph: the group location cluster layer and the group member layer. The nodes in the location layer represent group location clusters, while nodes in the group layer represent group members. There are three types of edges in this graph: the user-cluster edge (green), cluster-cluster edge (red) and user-user edge (blue).

criteria to define a meeting in our analysis: Two members must be approximately at the same location (within 20 meters) for at least five minutes. Please note that we use GPS location points, instead of location clusters, as a more accurate estimate of group members' meetings can be obtained using this data. For example, two friends can coincidentally appear at the same farmers market for more than five minutes but not see each other, and thus this should not be counted as a meeting.

- *Average similarity of group location cluster familiarity between every pair of group members*: Two users with more similar location cluster familiarity tend to know each other better. Here we model the group location cluster familiarity of user u by vector $\mathbf{loc}(u)$, where the ith element of $\mathbf{loc}(u)$ represents the number of location trace points in the ith location cluster detected in our dataset. Furthermore, $\mathbf{loc}(u)$ is normalized per user.

Formally, we first measure the social relationship between two group members u_i and u_j as

$$\text{social}(u_i, u_j) = w_1 M(u_i, u_j) + (1 - w_1) \cdot Cosine(\mathbf{loc}(u_i), \mathbf{loc}(u_j)), \tag{1}$$

where $M(u_i, u_j)$ is the normalized average number of meetings per day detected between group members u_i and u_j, and $\mathbf{loc}(u_i)$ is the vector representing $u_i's$ location familiarity for each group location cluster. w_1 balances the relative importance of co-occurrence and location trace similarity. Normally user location co-occurrence is a much stronger signal than location trace similarity in indicating social relationship strength. Hence w_1 is set to 0.8 in our experiments based on the observations from our dataset. The similarity of two group members' location cluster familiarity vectors is measured by Cosine similarity:

$$Cosine(\mathbf{loc}(u_i), \mathbf{loc}(u_j)) = \frac{\mathbf{loc}(u_i) \cdot \mathbf{loc}(u_j)}{||\mathbf{loc}(u_i)|| \times ||\mathbf{loc}(u_j)||}. \tag{2}$$

Given the social relationship strength computed between pairs of users, a group's social relationship strength is defined as the average of all the pairwise social relationships among the group members:

$$\text{Social}(G) = \frac{2 \sum_{u_i, u_j \in G} \text{social}(u_i, u_j)}{|G|(|G| - 1)}, \tag{3}$$

where $|G|$ represents the group size.

6.3 Location Familiarity Modeling for Individuals

Next, we focus on what factors may influence an individual to visit a particular location cluster. Figure 3 listed four such factors that we considered, and in this section we focus first on location familiarity as a predictor, followed by the three other factors considered in the next subsection. The location familiarity of individuals within a group may vary from user to user. To model users' location familiarity more accurately, we leverage transitions between group location clusters and the social relationship between group members. Intuitively, if a user's close friend is very familiar with one location cluster, it is highly likely that this user has been there or at least heard about it before, even though a visit to this location was not detected in location trace data. Taking this into consideration, we construct a two-layer graph that effectively combines group members' social relationships and location trace information. The structure of the graph is shown in Figure 4. There are two layers in the graph: the group location cluster layer and the group member layer. The nodes in the group location cluster layer represent group location clusters, and the nodes in the group member layer represent group members. There are three types of edges in this graph: the user-cluster edge (green), the cluster-cluster edge (red), and the user-user edge (blue).

Edge-specified Random Walk with Restart (Edge-RWR). To estimate a group member's familiarity to a group location cluster, we propose a novel Edge-specified Random Walk with Restart (Edge-RWR) algorithm to calculate the relevance score between a group member and a group location cluster. Unlike the traditional RWR method that treats every edge equally, we define the transition probability of each edge follows.

The *user-cluster* edge describes the probability that a user will visit a location cluster. Instead of simply using the number of visits, we apply the widely used idea of TF-IDF [33], where higher values indicate that the user is more interested in this location cluster and this interest is significant among other group members. The same approach is utilized when estimating cluster-cluster and user-user edges' probabilities. Specifically, given m users and n location clusters of a group, the user-cluster subgraph $\mathbf{G}_{\mathrm{UC}}$ is represented by a $m \times n$ adjacency matrix $\mathbf{M}_{\mathrm{UC}}$, where $\mathbf{M}_{\mathrm{UC}} = \{p(c_j|u_i)\}, 0 \le i < m, 0 \le i < n$, and $p(c_j|u_i)$ is estimated by:

$$p(c_j|u_i) = \alpha_1 \frac{Avg(u_i, c_j)}{Avg(c_j)} + (1 - \alpha_1)\frac{1}{m}, \tag{4}$$

where $Avg(u_i, c_j)$ refers to the average number of visits per day made by user u_i to cluster c_j. $Avg(c_j)$ refers to the average number of total visits to cluster c_j per day made by all the users in the group. And $1 - \alpha_1$ is the probability that the walker will restart from a randomly selected user node.

The *cluster-cluster* edges measure the transition probabilities of users going from one location cluster to another. Given n detected location clusters, the cluster-cluster subgraph $\mathbf{G}_{\mathrm{CC}}$ is represented by a $n \times n$ adjacency matrix $\mathbf{M}_{\mathrm{CC}}$, where $\mathbf{M}_{\mathrm{CC}} = \{p(c_j|c_i)\}, 0 \le i < n, 0 \le j < n$, and $p(c_j|c_i)$ is estimated by:

$$p(c_j|c_i) = \alpha_2 \frac{Avg(c_i, c_j)}{Avg(c_j)} + (1 - \alpha_2)\frac{1}{n}, \tag{5}$$

where $Avg(u_i, c_j)$ refers to the average number of transitions per day from cluster c_i to cluster c_j. $Avg(c_j)$ refers to the average number of total transitions to cluster c_j per day from all of the group location clusters. And $1 - \alpha_2$ is the probability that the walker will restart from a randomly selected location cluster node.

The *user-user* edges represent how well two users know each other. We use $social(u_i, u_j)$, introduced in Equation 1, to model the social relationship strength between two users, u_i and u_j. Given m group members, the user-user subgraph $\mathbf{G}_{\mathrm{UU}}$ is represented by a $m \times m$ adjacency matrix $\mathbf{M}_{\mathrm{UU}}$, where $\mathbf{M}_{\mathrm{UU}} = \{p(u_j|u_i)\}, 0 \le i < m, 0 \le j < m$, and $p(u_j|u_i)$ is estimated by:

$$p(u_j|u_i) = \alpha_3 social(u_i, u_j) + (1 - \alpha_3)\frac{1}{m}, \tag{6}$$

where $1 - \alpha_3$ is the probability that the walker will restart from a randomly selected user node.

ALGORITHM 1: Edge-specified Random Walk with Restart (Edge-RWR)

Input: Network $G = G_{UC} + G_{CC} + G_{UU}$; starting node u_i; restart probability $\alpha_1, \alpha_2, \alpha_3$;
Output: Stationary group location cluster vector w_C for a random walk starting at u_i;

```
Let all the nodes be initialized to 0, except for a 1 for the u_i node;
Let w_U denote the weights of the column vector of users, and w_C denote the weights of the column vector of location clusters;
While (U and C have not converged):
    w_C^{k+1} = α_1 M_UC · w_U^k + α_2 M_CC · w_C^k + (1 − α_2) e_C;
    w_U^{k+1} = (1 − α_1) e_U + α_3 M_UU · w_U^k + (1 − α_3) e_U;
Output group location cluster vector w_C.
```

Time Complexity. Algorithm 1 presents the algorithm for finding the stationary vector of the random walk from a single starting user node. The time complexity of *"Edge-RWR"* is $O(k \cdot |V|^2)$, where k is the number of iterations required for convergence, and $|V|$ is the total number of nodes. According to Tong et al. [38], the ratio of the first two eigenvalues of the transition matrix specifies the rate of convergence to the stationary point.

After constructing the group location cluster-group member graph, *"Edge-RWR"* is then performed to compute the location familiarity between a user and a location cluster. The random walker will start from one node, and travel following the probability assigned to each edge to select the next transition. In addition, in order to avoid stoppage of the random walk, we assume that the random walker will in some cases jump back and restart at the same node. At each transition, the walker would either restart at one node (with probability $1 - \beta$), or move to a neighboring node (with probability β). In the two-layer graph that we constructed, the walker will start from a group member node. For each transition, the walker will either jump to an adjacent node, or jump back to the same starting user node. We set the parameter $\alpha_1 = \alpha_2 = \alpha_3 = 0.85$, which has also been widely used in prior literature. The intuition is that if a group location cluster node can be easily reached from a group member node, the user has a higher familiarity with this location cluster. Formally, the location familiarity of an individual group member u for location cluster c is modeled as

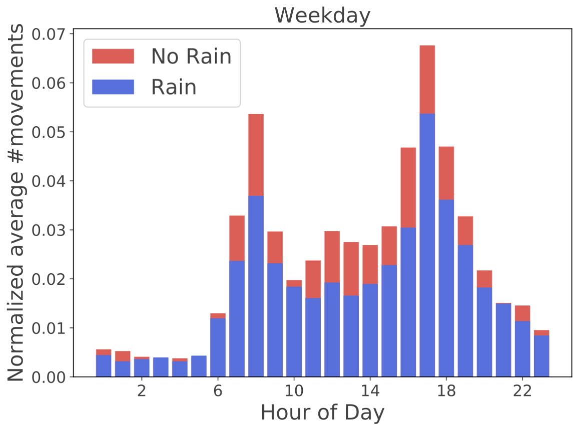
$$\text{loc\_familiarity}(u, c) = \text{PageRank}_u(c), \tag{7}$$

where $PageRank_u(c)$ is the PageRank score of group location cluster c for user u after running the Edge-RWR algorithm starting from user u.

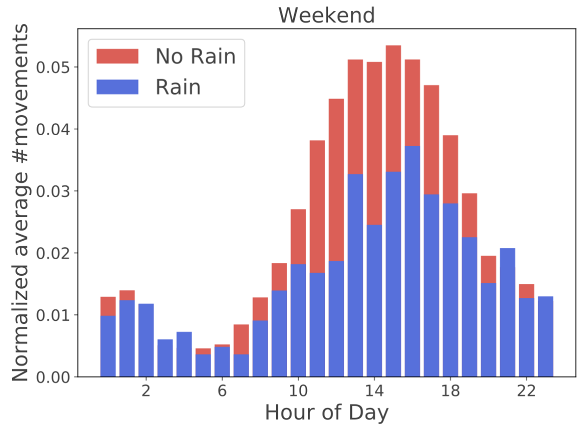
6.4 Contextual Predictors

6.4.1 User Mobility. User mobility can significantly impact daily activities. In a preference survey, Elgar *et al.* find that users with different levels of mobility differ substantially in how they value time [12]. Previous insights also suggest that users with high mobility (longer daily travel distance) tend to be more active in attending group events, and an individual's mobility has a positive correlation with her time and location availability [41]. There are two reasonable explanations for this phenomenon:

- As researchers have found, people with different modes of transportation, such as walking, taking a bus, riding a bike, and driving, exhibit significant difference in how they value time [12, 45]. Users with high mobility may travel by car or public transportation. Highly mobile users are more likely to attend group events that are far away from their current locations.
- Zhang *et al.* also found that users with high mobility are more likely to be active event attendees, as they are used to frequent meetings with friends after school or work, which generates longer travel distances [41].



(a) Number of movements by hour on weekdays.



(b) Number of movements by hour on weekends.

Fig. 5. Figure 5a and 5b show the normalized average number movements on weekdays and weekends. Bad weather causes a sharp decrease of daily movements on both weekdays and weekends, and the impact is larger on weekends compared with weekdays.

As such, we model the influence of user mobility as follows: a user's group event attendance is proportional to the user's mobility. Here we define user mobility as the total travel distance in the 48-hour period preceding the creation of the event. More specifically, we estimate user mobility by summing up the geographical distance between every two consecutive location points detected from the user's location trace data in this time period.

6.4.2 Day of the Week. Using the detected group location clusters, we can identify a user's daily movement pattern among these location clusters. We find in our data analysis that users' movement patterns can be quite different on weekdays compared with weekends. The average inter-cluster movement distance is 5.3 km on weekends and 4.8 km on weekdays. This pattern seems reasonable, since intuitively people are more willing to travel longer distances for activities on weekends. And the average movement number is 5.2 on weekdays and 5.4 on weekends.

We also find that weather has a significantly different impact on weekdays compared with weekends. We control geographical confounding factors relating to weather by only using a partial dataset collected at our university. When obtaining the weather information during the time of a given movement made by a user, we round the time to the nearest hour. For example, if the starting time of the movement is at 18:15, we collect the weather data at 18:00 that day using the Weathersource API [36]. For the hours with positive precipitation values (including snow), we consider it to be rain. The normalized average movements by hours under different weather conditions on weekdays and weekends are shown in Figure 5. These results demonstrate that bad weather causes a sharp decrease in daily movements on both weekdays and weekends. We also find that the movement difference between rain and no-rain hours is larger on weekends, indicating that weather has a higher impact on weekend activities compared with weekdays.

In summary, users' movement patterns are considerably different on weekdays and weekends, which also impact users' group meeting preferences. Therefore, it would be beneficial to consider the day of the week information in our prediction model. We also considered including weather information as a contextual feature in our prediction model. However, since we are trying to recommend venues when a group event is created, only weather forecast information can be used, which can be inaccurate but also difficult to collect as our users are spread throughout the US. To ensure that our GEVR system is applicable in real-life cases, we decided not include weather forecast data in our recommendation system in this paper. However, incorporating the impact of weather forecasts on group event decisions could be an interesting topic, which we save for future work.

6.4.3 Population Density. As shown in Figure 1a, our users are spread all over the US, across 40 states and 117 cities. Population density can play a significant role in modeling group members' location preferences. Previous research has found that geography and social relationships are inextricably intertwined. People living in rural and urban places exhibit significant differences in using social technologies, such as Facebook and Myspace [4, 17]. These findings inspire us to study the effect of population density on group event scheduling. Therefore, we also incorporate the population density information into our prediction model.

6.5 Individual Location Preference Prediction

Taking all of the features we have discussed thus far into consideration, we integrate all of the contextual features into a Classification and Regression Tree (CART) model [24]. We predict a user's location preference with four types of features: *location familiarity*, *user mobility*, *day of the week*, and *population density*. Location familiarity is calculated using Equation 7 introduced in Section 6.3. User mobility models a user's average travel distance on a daily basis. For users with high mobility, traveling longer distances can be easier compared with low-mobility users, due to the fact that high-mobility users generally travel by car or public transportation, and they are accustomed to longer-distance travel. Our analysis further illustrates that users' movement patterns vary significantly on weekdays versus weekends, and in high population density areas versus low population density areas; thus, we also add these two features into our regression model. We also evaluated other popular regression methods, such as Linear Regression, Support Vector Regression, and Gradient Boosting Regression. CART outperforms all of these regression models for our application. We present detailed results in Section 8.

One major advantage of our individual location preference prediction model is its flexibility in leveraging new features. Higher-level mobility features, which are related to users' group event attendance activity, can be easily added to our current model. For example, Stenneth et al. [37] and Hemminki et al. [18] have found that a user's mode of transportation, including car, bus, train, walking, and bike, can be accurately inferred using GPS location traces. As future work, we plan to investigate how users' modes of transportation can impact group event attendance.

6.6 Social-based Group Location Prediction Model

Thus far, we have constructed a prediction model that integrates all the features we have identified in an intelligent way to model an individual's location preferences. Now, we propose a new *"social-based"* group location prediction model to aggregate the predicted location preferences for individuals and predict the final location cluster where the group will meet.

Given every group member's location preferences, how does the group arrive at a final decision of where to meet? In previous group recommendation work, three main strategies for aggregating individual preferences have been widely explored [8, 15, 20]: *"least misery," "average satisfaction,"* and *"maximum satisfaction."*

- *Least misery* minimizes the dissatisfaction of the least satisfied member: $\min_{u \in G} \mathbf{p}(u, c)$.
- *Average satisfaction* is a straightforward strategy that assumes every group member carries the same weight and computes the average satisfaction: $\frac{1}{|G|} \sum_{u \in G} \mathbf{p}(u, c)$.
- *Maximum satisfaction* maximizes the enjoyment of the most satisfied group member: $\max_{u \in G} \mathbf{p}(u, c)$.

Earlier studies have also argued that groups are diverse and none of these group decision strategies are dominant across all groups [15, 27]. Gartrell *et al.* found that the social relationship strength of a group plays a significant

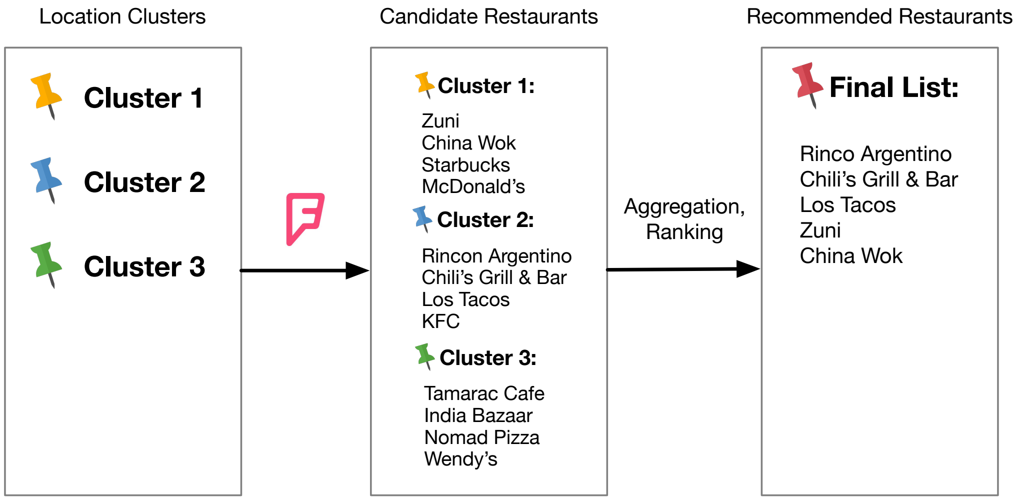


Fig. 6. The main flow of the group event venue recommendation framework: For each detected group location cluster, we call the Foursquare Venue Recommendation API to get a list of nearby restaurants. We aggregate the restaurants of each cluster to generate a candidate restaurant pool. Prediction results from the *"social-based"* group location cluster prediction model are then used to rank the restaurant pool and generate the final Top-K restaurant list.

role in the group decision-making process. When the relationship strength is high, the final decision tends to reflect the maximum satisfaction strategy. When the relationship strength is weak, the final decision tends to reflect average satisfaction or least misery [15]. The intuition here is clear: If the group members don't know each other, the main guidance for making the group decision is to avoid dissatisfying the other group members. If the group has high social relationship strength, some people who are knowledgeable about the topic may lead the discussion and others tend to follow their suggestions. Taking these factors into consideration, we propose the following *"social-based"* group consensus function:

$$\mathbf{P}(G,c) = \begin{cases} min_{u \in G}\mathbf{p}(u,c) & \text{if} \quad \text{Social}(G) < \beta & \text{(least misery)} \\ \frac{1}{|G|}\sum_{u \in G}\mathbf{p}(u,c) & \text{if} \quad \beta \leq \text{Social}(G) \leq \alpha & \text{(average satisfaction)} \\ max_{u \in G}\mathbf{p}(u,c) & \text{if} \quad \alpha < \text{Social(G)} & \text{(maximum satisfaction),} \end{cases} \tag{8}$$

where Social(G) is defined in Equation 3, and parameters α and β are thresholds for the group's social relationship strength. We choose $\alpha = 0.6$ and $\beta = 0.2$ in our evaluation based on a grid search targeting the best performance for predicting groups' final meeting location clusters using our dataset. We will present the details of our experiments in Section 8.

7 VENUE RECOMMENDATION FRAMEWORK

In this section, we present our proposed group event venue recommendation framework. The intuition behind our framework is to re-rank group's nearby venues, returned by Foursquare's API, based on the group location cluster prediction results calculated by Equation 8 to provide event venue recommendation for groups. The main flow of the framework is shown in Figure 6. Aiming to recommend event venues for groups of mobile users in real-life, our framework is composed of three parts:

(1) The group location clusters generated by grouping location traces' points described in Section 5;
(2) The *"social-based"* group location prediction model introduced in Section 6;

(3) The venue recommendation system provided by Foursquare's API [14].

Given a combination of location coordinates, a radius, and a limit (maximum number of results to return), the Foursquare API will return a list of recommended venues near the specified location within the given radius. In our setting, given a group G and its group location clusters $\{c_1, c_2, \cdots, c_n\}$, we call the Foursquare API to return a list of m recommended venues $\{v_{i1}, v_{i2}, \cdots, v_{im}\}$ for each cluster c_i. Here $\{v_{i1}, v_{i2}, \cdots, v_{im}\}$ follows the order returned by the Foursquare API. All recommended venues are combined to generate a candidate venue pool. Two factors are leveraged to re-rank the venue pool: the venue rank in the returned list provided by the Foursquare API and probabilities of corresponding location clusters being selected as the final one calculated by our *"social-based"* group location prediction model. More specifically, the rank score of a venue v_{ij} in the final recommendation list is defined as

$$\text{rankscore}(v_{ij}) = \frac{m-(j-1)}{m} \cdot \mathbf{P}(G, c_i), \tag{9}$$

where $\frac{m-(j-1)}{m}$ models the rank weight returned by the Foursquare API and $\mathbf{P}(G, c_i)$ is the probability of location cluster c_i being selected as the final location cluster computed by Equation 8. The final venue recommendation list created by our framework is ranked by rankscore(v_{ij}) among all the venues in the candidate venue pool.

8 EVALUATION

In this section, we evaluate the effectiveness of GEVR's proposed individual location preference prediction model, social-based group location cluster prediction model, and group event venue recommendation framework.

8.1 Performance of Predicting Individual Location Preference

We first evaluate the performance of predicting individual location preference. Specifically, we evaluate "whether a group member will vote for a suggested venue or not," and the answer is a binary label with 1 representing *yes* and 0 representing *no*. The evaluation metrics we use are AUC (the area under the ROC curve) and F1 score. Since our model quantifies a group member's location cluster preference by outputting a probability value, AUC can find the best threshold for the binary classification. And F1 score is computed to evaluate whether the model provides comparatively robust performance. We use a stratified 5-fold cross validation with respect to the users. All the features are standardized.

As introduced in Section 6, our prediction model utilizes four main types of features: location familiarity, user activeness, day of the week, and population density. By comparing the performance of the popular supervised models shown in Table 2, we find that *"Edge-RWR+CART"* achieves the best results. The other supervised regression models include Linear Regression (LR), Support Vector Regression (SVR), and Gradient Boosting Regression (GR). *"Edge-RWR+CART"* achieves the best overall performance in estimating users' location cluster preference (AUC = 0.851, F1 = 0.831) and outperforms all the other supervised models by a clear margin (beating the second best one *"Edge-RWR+GR"* by 8.9% in AUC and 7.1% in F1 Score). More importantly, our newly proposed *"Edge-RWR"* beats the traditional *"RWR"* consistently when combined with each supervised learning model that we tested with an average improvement of 13%, indicating that our edge-specified network is superior in estimating the location cluster preference between a group member and a location cluster.

8.2 Performance of Predicting Where Groups Meet

The next question is whether we can accurately predict where a group meets by using our newly designed *"social-based"* group decision strategy. We formulate the problem as a **multi-class prediction task**: Given all the group location clusters, can we accurately predict which location cluster the group will decide to meet at. Note that the design of the mobile application allows a host to override the voting results by selecting a final venue that is not necessarily the winning one (based on group members' votes). This may happen when group

Table 2. Performance for predicting individual location preference. Standard error values are shown in parentheses. *"Edge-RWR+CART"* outperforms all the other supervised models by a clear margin. Moreover, our newly proposed *"Edge-RWR"* beats the traditional *"RWR"* consistently when combined with each supervised learning model that we tested, indicating that our edge-specified network is superior in estimating the familiarity between a group member and a location cluster.

Model	AUC	F1 Score
RWR+LR	0.702 (0.020)	0.675 (0.021)
Edge-RWR+LR	0.705 (0.017)	0.669 (0.022)
RWR+SVR	0.745 (0.010)	0.722 (0.008)
Edge-RWR+SVR	0.763 (0.011)	0.757 (0.009)
RWR+GR	0.788 (0.015)	0.747 (0.013)
Edge-RWR+GR	0.795 (0.011)	0.766 (0.008)
RWR+CART	0.838 (0.011)	0.812 (0.009)
Edge-RWR+CART	**0.851 (0.011)**	**0.831 (0.010)**

members discussed where to hang out and made an agreement to disregard the poll results. In our dataset, the host decided to override the majority voting results in roughly 20% of the group events. Therefore, we could measure the effectiveness as either predicting the most popular location cluster according to the poll results, i.e., the "winning location cluster," or predicting the final location cluster actually visited by the group including overrides, namely the "final location cluster." To evaluate the performance of our model, we report the prediction accuracy of both the winning location cluster and the final location cluster. We compare our *"social-based"* group location prediction model with the three commonly used group decision models introduced in Section 6.6, and three other state-of-the-art group recommendation models that have been shown to be effective in their specific group settings in prior literature:

- *Logsitic regression-based* [8]: A logistic regression-based approach that determines the probability of a group view (p_G) from the individual probabilities:

$$\log \frac{p_G}{1-p_G} = \alpha_0 + \alpha_1 p_1 + \alpha_2 p_2 + \alpha_3 p_3 + \ldots + \alpha_n p_n. \quad (10)$$

- *Rule-based* [15]: A rule-based group consensus framework that constructs associative classification rules by mining the training dataset.
- *Expertise-based* [32]: Estimating group decision by averaging the individual preference with weighted expertise. The expertise is defined as the number of times this user participated before.

The results of predicting the winning location cluster and the final location cluster are shown in Table 3. As can be observed in the table, our newly proposed social-based group decision strategy achieves the best performance in both predicting winning location cluster (Accuracy=0.826) and final location cluster (Accuracy=0.801). It is also worth noting that our *"social-based"* group location prediction model achieves better prediction accuracy than all other three widely-used single group decision strategies, with an average improvement of 12%. This echoes the insights illustrated in [15, 27]: groups are diverse and none of the single group decision strategies are dominant across all groups. The accuracy achieved by the *"logistic regression-based"* approach also shows that a linear combination of individual preferences without considering the social relationship strength among group members does not perform well. Moreover, our *"social-based"* model performs much better than the *"expertise-based"* model. One possible explanation is that our dataset currently does not have enough events created by the same group. As a result, the expertise of group members cannot be accurately modeled in such cold-start scenarios. We would like to explore if the *"expertise-based"* model will perform better when we have more group events data. We save this topic for future study.

Table 3. Accuracy of predicting the winning location cluster and final location cluster. Individual location preferences are estimated using *"Edge-RWR+CART"* as it achieves the highest performance in modeling individual location preference. Standard error values are shown in parentheses. Our *"social-based"* group location prediction model achieves better prediction accuracy than all the other comparative models, with an average improvement of 12%.

Model	Winning location cluster	Final location cluster
Least misery [32]	0.689 (0.019)	0.668 (0.018)
Average satisfaction [8, 15, 32]	0.771 (0.016)	0.753 (0.017)
Most pleasure [8, 32]	0.760 (0.019)	0.739 (0.022)
Logistic regression-based [8]	0.767 (0.016)	0.741 (0.019)
Rule-based [15]	0.686 (0.022)	0.657 (0.024)
Expertise-based [32]	0.758 (0.018)	0.729 (0.016)
Social-based	**0.826 (0.013)**	**0.801 (0.013)**

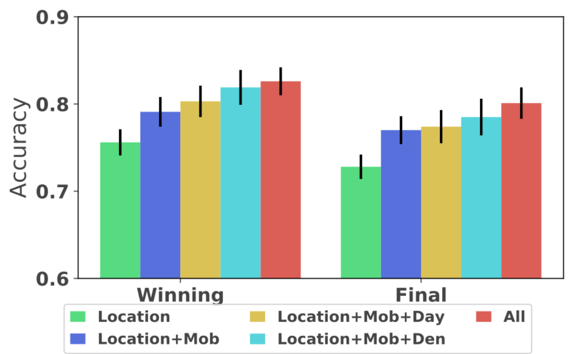
Performance for groups of different sizes. To better understand the impact of group size on recommendation performance, we computed the prediction accuracy for different group sizes using our *"social-based"* group location cluster prediction model. Our model achieves an average accuracy of 0.818 for groups with three members (0.826 for the winning location cluster and 0.810 for the final location cluster), 0.807 for groups with four members (0.833 for the winning location cluster and 0.781 for the final location cluster), and 0.806 for groups with 5–6 members (0.823 for the winning location cluster and 0.788 for the final location cluster). These results show that our model performs reasonably well for groups with sizes between 3–6. However, the trend or the change in prediction performance is not statistically significant. As described in Section 3, only 17% of the groups in our dataset have more than four members. This would be an interesting problem to explore when we collect more data for larger groups.

Feature analysis. Our *"social-based"* prediction model utilizes four types of features: Location familiarity (Location), user mobility (Mob), day of the week (Day), and population density (Den). To understand the effectiveness of these features, we evaluate the prediction performance using different feature combinations. As shown in Figure 7a, using location familiarity alone achieves reasonable performance, and adding user mobility further increases the accuracy (from 0.756 to 0.791 for predicting the winning location cluster, and from 0.728 to 0.770 for predicting the final location cluster). One possible explanation is that users with high mobility may travel by car or use public transportation. Reaching far-away places are thus easier for them. It is also likely that users with high mobility are inherently active event attendees, as they are used to frequent meetings with friends after school or work [41]. Additionally, both day of the week and population density features contribute to the overall improvement.

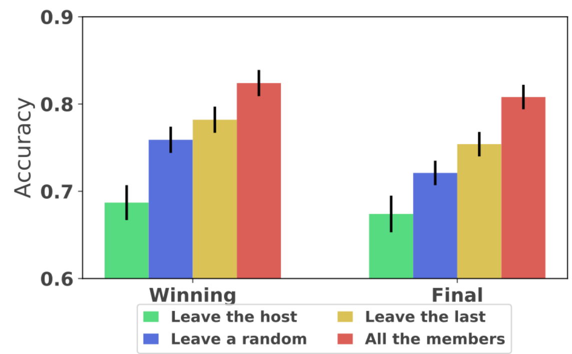
Performance on Groups with Cold-Start Users. As we discussed in previous sections, it is common for groups to have newly joined members without (sufficient) historical information. To evaluate our *"social-based"* group location prediction model in such cold-start scenarios, we set up two experiments:

- *Leave-one-out*: Assuming one group member is a new user of the app, we do not have any historical information about him/her.
- *Effect of data sparsity*: We further study how our model deals with the cold-start problem by using only a proportion of each user's data (20%, 40%, 60%, 80%, and 100%).

Figure 7b shows the performance of the "leave-one-out" experiment. We simulate three "leave-one-out" situations by leaving (a) the group host, (b) a random group member (except the host and the last joined group member), or (c) the last joined group member as the new user with no historical information available. As we can see in Figure 7b, the group host's location information has the largest impact on the overall prediction performance



(a) Feature analysis.



(b) Cold-start analysis.

Fig. 7. Feature analysis (Figure 7a) and cold-start analysis (Figure 7b). In Figure 7a, using the location familiarity and user mobility features already provide reliable performance. Both day of the week and population density features also contribute to the overall improvement. In Figure 7b, the group host's location information has the most significant impact on the overall prediction performance when predicting both the "winning location cluster" and "final location cluster," while the last joined group member has the least impact.

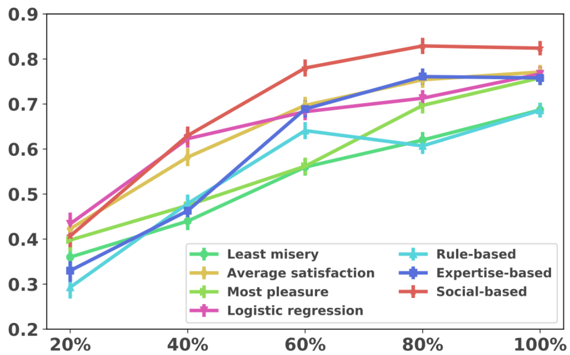
when predicting both the "winning location cluster" and the "final location cluster," while the last joined group member has the least impact. This finding echoes an observation made in [41]: "The group host will have more influence on the group decision-making process. The late coming voters tend to vote for options that align with existing voting results, thus, having smaller impacts on the final decisions." More importantly, our *"social-based"* group location prediction model achieves promising results with cold-start users in the group, with an average accuracy of 0.768 when assuming the last group member' data is unknown and an average accuracy of 0.740 when leaving a random group member's data out.

We further study how our proposed model deals with the data sparsity problem by using different proportions of the users' location traces and compare it with all the other models shown in Table 3. The results are shown in Figure 8. We gradually increase the proportion of users' location traces used for testing the performance from the first 20%, first 40%, till 100% (temporally ordered). As can be seen in the figure, our proposed *"social-based"* model consistently outperforms all the other models when the proportion is above 40%. It does not perform very well when only 20% of the location traces are available for prediction. It is likely that the social relationship strength among group members cannot be precisely estimated with very limited data, which affects the overall prediction accuracy.

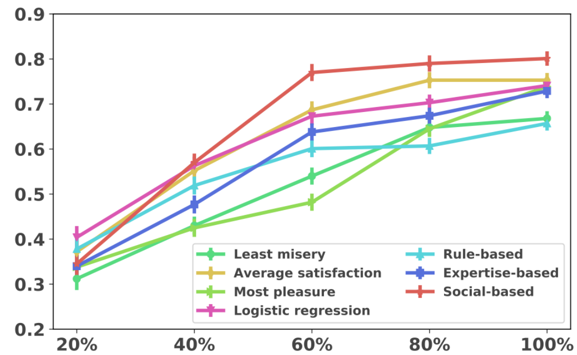
8.3 Performance of Event Venue Recommendation for Groups of Mobile Users

To evaluate whether our proposed GEVR system is effective for providing event venue recommendations for groups of mobile users, we compare it with the following baselines:

- *Foursquare*: Get 50 recommended restaurants provided by the Foursquare Venue Recommendation API by searching the city name where the group is located at and retain the restaurants' Foursquare rankings without any modification.
- *Most popular*: Same as the Foursquare model but re-ranks the restaurants by Foursquare number of check-ins (popularity).
- *Highest rating*: Same as the Foursquare model, but re-ranks the restaurants by Foursquare user review star rating. If there is a tie, the restaurant with more check-ins is ranked higher.



(a) Accuracy of predicting the winning location cluster.



(b) Accuracy of predicting the final location cluster.

Fig. 8. Accuracy of predicting the winning location cluster (Figure 8a) and the final location cluster (Figure 8b) using all the models in Table 3 with different proportions of users' location trace data. Our proposed *"social-based"* model consistently outperforms all the other models when the proportion of data is over 40%.

- *Equal-weighted*: For each detected group location cluster, getting 50 recommended restaurants by searching the nearby area within 0.5 km using the Foursquare Venue Recommendation API and rank them using Equation 9, assuming every location cluster is weighted equally. If there is a tie, restaurants with more check-ins are ranked higher.
- *Least misery* [32]: Using the *"least misery"* group location prediction model to recommend venues.
- *Average satisfaction* [8, 15, 32]: Using the *"average satisfaction"* group location prediction model to recommend venues.
- *Most pleasure* [8, 32]: Using the *"most pleasure"* group location prediction model to recommend venues.
- *Logistic regression-based* [8]: Using the *"logistic regression-based"* group location prediction model to recommend venues.
- *Rule-based* [15]: Using the *"rule-based"* group location prediction model to recommend venues.
- *Expertise-based* [32]: Using the *"expertise-based"* group location prediction model to recommend venues.

Here, getting 50 recommended restaurants means setting the field *limit* = 50 as the number of results to return when using the Foursquare API. We choose 50 because it is the maximum number allowed. The returned list could be shorter than 50 if there are not enough restaurants registered nearby.

The three basic baselines *"most popular," "highest rating,"* and *"Foursquare"* are intended to demonstrate the recommendation performance without leveraging the detected group location clusters. The *"equal-weighted"* baseline uses group location clusters information, but assumes no prediction results are provided, so all detected group location clusters are equally weighted. The next six baselines combine individual location preference with different group decision strategies, aiming to determine the feasibility of our social-based group decision strategy for event venue recommendation. We use a widely used metric, **hit rate**, to evaluate the recommendation performance of these models. For each group, we recommend top-N (N=5,10,15,20 in our experiments). Hit rate is defined as the proportion of groups' top-N recommendation lists that includes the final venue decided by the group in our dataset. For example, suppose there are 100 group events; then we would generate 100 top-N lists. For a given event, if the final venue chosen by the participants is included in the top-N recommendation list generated for that event, then we consider that the recommendation to be a success. If out of 100 events, 30 events' final venues are included in their corresponding top-N lists, then the hit rate (final venue) is 0.3.

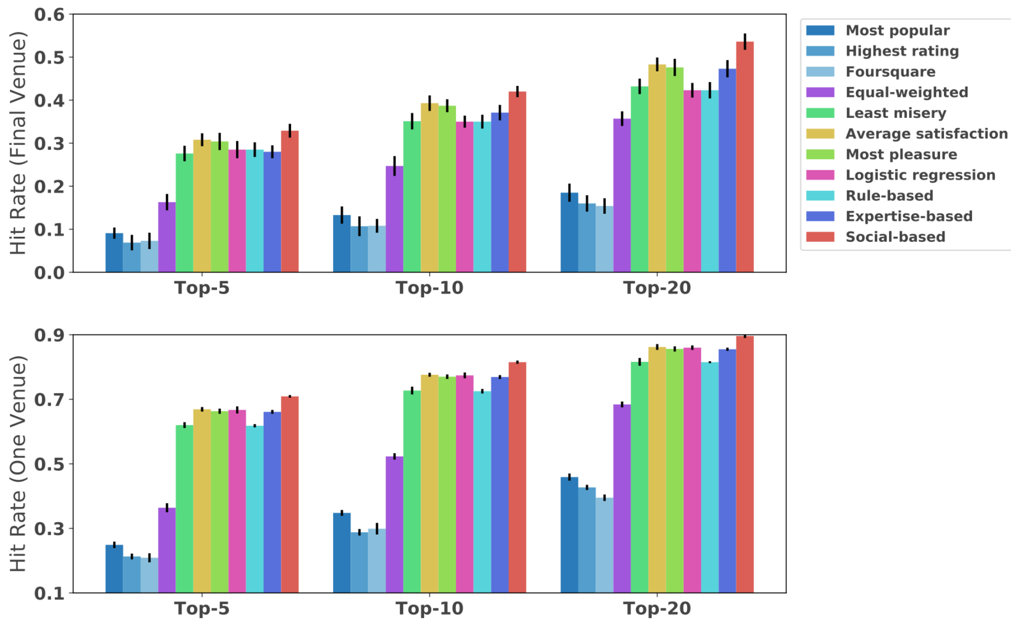


Fig. 9. Recommendation performance of different models on hitting a group's final venue (top subfigure) and hitting at least one of the group's suggested venues (bottom subfigure). Our *"social-based"* group location prediction model outperforms all the other comparative models consistently in both cases.

Results of hitting the final group meeting venue for each of these baseline models, as well as our GEVR, are shown in Figure 9, top subfigure. We can see that our baselines gradually build towards the performance of our final recommendation system GEVR, indicating that each of the components (**location clusters**, **individual location preference model**, and **social-based group decision strategy**) is a meaningful contribution to the event venue recommendation for groups of mobile users.

The three basic baselines are clearly ineffective, achieving an average hit rate for Top 5, Top 10, and Top 20, with only 0.08, 0.12, and 0.17, respectively. In comparison, the *"equal-weighted"* model, which considers group location cluster information, achieves better performance. But it is not as effective as the remaining ones that model the group decision-making process to help re-rank the recommendation list. Our *"social-based"* group location prediction model outperforms the other six comparative models with an average improvement of 14% for Top-5, 15% for Top-10, and 19% for Top-20. This fits our expectation, as the social-based group decision strategy provides the most promising accuracy in predicting the groups' final location cluster, as shown in Table 3.

We further calculate the success rate of hitting at least one of the group's suggested venues, i.e., venues that they voted on, using all these models. This is a different and broader metric than the prior hit rate (final venue), so we define this as hit rate (one venue). For example, suppose for a given event that the group members have suggested three venues that they are interested in visiting. Then if any one of those three appear in the top-N list, we consider that recommendation to be a success. If out of 100 group events, at least one of the voted-upon venues appear in a top-N list for 30 of those events, then the hit rate (one venue) is 0.3. The results of this metric are shown in Figure 2, bottom subfigure. As can be seen, the relative recommendation performance of the different methods follows a similar pattern as the previous metric. The overall absolute value of this hit rate (one venue) metric is higher than the final venue hit rate because there is a higher chance of any one of the suggested venues appearing in a top-N list compared with just the final venue appearing.

9 CONCLUDING DISCUSSION

In this work, we present GEVR, the first event venue recommendation system for groups of mobile users. The system tackles the recommendation challenges by splitting the problem into three separate steps: (1) detecting group location clusters using group members' location traces; (2) predicting where the group will meet; and (3) aggregating and re-ranking venues in nearby areas, crawled through the Foursquare Venue Recommendation API, to create a final venue recommendation list based on the prediction results. When predicting a group's gathering decision, we first model individual group members' location preferences based on four types of features: location familiarity, user activeness, day of the week, and population density. We then design a novel social-based group location prediction model to aggregate individual preferences and predict the group's final decisions based on the group's social relationship strength.

To evaluate the performance of GEVR, we have collected a unique dataset from a group event scheduling mobile application, OutWithFriendz. In total, 625 users participated in this study and created 502 group events. Our users are widespread across the US, covering 40 states and 117 cities. Evaluation results show that GEVR can provide over 80% accuracy for predicting which location cluster the group will meet at, and our newly designed social-based group location prediction model outperforms all the other state-of-the-art group location prediction models with an average improvement of 16% on hit rate.

9.1 Significant Implications

One of the significant implications of this work is that it is essential to integrate location considerations into an event venue recommendation framework for groups of mobile users. As demonstrated clearly in Figure 9, approaches that do not consider group members' location behavior in the recommendation process performed much worse than our GEVR approach. We expect that the research community will build upon this finding, and future solutions for group recommendation in other contexts will also incorporate mobility into their recommendation strategies.

Similarly, our work clearly demonstrates that methods for individual venue recommendation such as *"most Popular," "highest rating,"* and *"Foursquare"* are considerably less effective when applied to group venue recommendation than algorithms that incorporate knowledge of group characteristics. Group properties, such as mobility spread, membership fluidity, and diversity of interests, introduce a higher level of complexity that impacts the final venue chosen by the group. Unlike individual recommendation, which mostly relies on personal preferences, negotiation and coordination are necessary for group members to reach a final agreement. Understanding the group negotiation and coordination process is critical for understanding group event scheduling behaviors and providing effective group event venue recommendation. Our work also indicates that one simple group decision strategy, e.g., average satisfaction, cannot apply to all the groups. Groups with strong social relationships usually apply different strategies compared with groups with weak social relationships.

Our results are uniquely validated based on a large number of real-world group venue events that have been quite challenging to obtain. For privacy reasons, we are unable to release the dataset of group events, because the location traces are difficult to fully anonymize. However, we can lower the barrier to entry for future researchers and have obtained permission from the authors of the software used in this study to release the applications as open source. This will enable other researchers to obtain valuable ground truth data to study the fine-grained behavior of groups of mobile users, and thus drive the growth in the research community for developing practically validated recommendation solutions for groups of mobile users.

Our work builds upon commercial location-based social services (LBSNs) such as Foursquare. This strategy opens up opportunities for more in-depth partnerships with the industry on joint development of novel group recommendation algorithms and systems that can be validated in the real world. This research strengthens the capabilities of academia in approaching the industry for collaborations, such as Google, Apple, and mobile social networks like Facebook, Instagram, and Snapchat, in hopes of introducing mobile tools that can facilitate useful group event coordination and venue recommendation.

There are a variety of strong use cases for our GEVR in real life. First, online group event organization services could use GEVR to find venues in their databases that are highly likely to be suitable for most group members. As previous research has pointed out, for many social groups, how to organize an event and attract more participants continues to be a difficult task. Even experienced organizers still feel stressed when planning a group event [2]. Our GEVR could to some extent reduce the group event host's burden of deciding where to meet. Additionally, although future research is needed to verify this hypothesis, it is possible that our GEVR could help target potential participants who would be interested in attending this group event. With the suggested venues and ongoing voting results of a new group event, we may be able to find group members' friends who fit this proposal perfectly. Validating this hypothesis could also be an exciting topic for future research.

9.2 Future Work

We plan to investigate the following dimensions to further improve the performance of GEVR. We would like to explore richer contextual clues in hopes of improving the accuracy of group recommendation. Location/mobility has proven to be quite helpful in improving the performance of the algorithm. More detailed social connection information may also prove fruitful. Weather forecast information could also be leveraged in our system. In addition, more historical information, such as voting behaviors within the application and venue preferences, would be useful to explore. Our dataset is comprised primarily of younger people, and we would like to incorporate a broader age demographic among our groups. We are also interested in how to incorporate serendipity into the group event venue recommendation process. Our current recommendation system is dependent on historical behavior in terms of preferences and locations. However, we also find that roughly 10% of the group events occurred in venues that do not fall into any of the detected group location clusters, indicating that groups may sometimes explore new venues that have not been visited before. To improve our system, we can leverage collaborative filtering, which has been shown to be very effective for individual recommendation. Specifically, when we have many groups in the same city, we can measure the similarity of different groups based on the context information, such as location traces and venue preferences. We can then recommend venues that are frequently visited by similar groups. Another strategy is to add a few "most popular" or "highest rated" venues in nearby areas that the group has never visited before into the recommendation list. How to balance the level of diversity between familiar and unfamiliar recommended places is an important research question for future exploration.

During our user study, we do not restrict the types of events that groups can organize. There are different types of events in our dataset, such as dining, outdoor hiking, and movies. However, we find that dining is the dominant event type. As described in Section 3, dining events account for 81% of all events in our dataset. As such, we decided to focus on dining recommendation in this paper. This is actually more challenging than recommending other types of event venues, since the number of restaurants in the same location cluster is much higher than other types of venues, such as movie theaters or hiking trails. We expect the proposed recommendation strategy would also help for other types of events, and would like to investigate it further when we have more data for other types of events.

Privacy concerns are becoming increasingly important, and we are interested in exploring ways to protect group members' location privacy. Protecting members' location privacy in group events may be more challenging than ensuring the privacy of individual users in other applications, as users' sensitive location information may be exposed to other group members. There are well-known individual privacy protection mechanisms such as k-anonymity and differential privacy. Whether those methods can offer adequate privacy protection for groups in a dynamic mobile setting is an open question. To address this problem, we would like to investigate several possible solutions. For example, for users who have privacy concerns and do not want their locations to be tracked, we could ask them to explicitly specify a list of their favorite venues when registering to use our application. Location preferences can be modeled using these specified venues, which may not be as effective as using location traces directly, but offers a good comprise between accuracy and privacy. In addition, adding a few "most popular" or "highest rated" venues in infrequently visited and unknown areas may be a good strategy to minimize disclosure of location information among group members.

ACKNOWLEDGMENTS

We thank anonymous reviewers for helpful comments and discussions. We thank Scott Fredrick Holman from the CU Boulder Writing Center for his feedback and support during the writing process. This work is supported in part by the US National Science Foundation (NSF) through grant CNS 1528138.

REFERENCES

[1] Niels Agatz, Alan Erera, Martin Savelsbergh, and Xing Wang. 2012. Optimization for dynamic ride-sharing: A review. *European Journal of Operational Research* 223, 2 (2012), 295–303.
[2] Judy Allen. 2008. *Event planning: The ultimate guide to successful meetings, corporate events, fundraising galas, conferences, conventions, incentives and other special events.* John Wiley & Sons.
[3] Daniel Ashbrook and Thad Starner. 2003. Using GPS to learn significant locations and predict movement across multiple users. *Personal and Ubiquitous Computing* 7, 5 (2003), 275–286.
[4] Lars Backstrom, Eric Sun, and Cameron Marlow. 2010. Find me if you can: improving geographical prediction with social and spatial proximity. In *Proceedings of WWW*. 61–70.
[5] Paul Baumann, Wilhelm Kleiminger, and Silvia Santini. 2013. The influence of temporal and spatial features on the performance of next-place prediction algorithms. In *Proceedings of UbiComp*. 449–458.
[6] Christoph Beckmann and Tom Gross. 2011. AGReMo: providing ad-hoc groups with on-demand recommendations on mobile devices. In *Proceedings of the 29th Annual European Conference on Cognitive Ergonomics*. ACM, 179–182.
[7] US Census Bureau. 2016. US Population Density Dataset 2010-2016. https://www2.census.gov/programs-surveys/popest/datasets/2010-2016/counties/totals/co-est2016-alldata.csv. [Online; accessed 01-November-2018].
[8] Allison JB Chaney, Mike Gartrell, Jake M Hofman, John Guiver, Noam Koenigstein, Pushmeet Kohli, and Ulrich Paquet. 2014. A large-scale exploration of group viewing patterns. In *Proceedings of TVX*. 31–38.
[9] Longbiao Chen, Daqing Zhang, Leye Wang, Dingqi Yang, Xiaojuan Ma, Shijian Li, Zhaohui Wu, Gang Pan, Thi-Mai-Trang Nguyen, and Jérémie Jakubowicz. 2016. Dynamic cluster-based over-demand prediction in bike sharing systems. In *Proceedings of UbiComp*. 841–852.
[10] Blerim Cici, Athina Markopoulou, Enrique Frias-Martinez, and Nikolaos Laoutaris. 2014. Assessing the potential of ride-sharing using mobile and social data: a tale of four cities. In *Proceedings of the 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing*. ACM, 201–211.
[11] Rong Du, Zhiwen Yu, Tao Mei, Zhitao Wang, Zhu Wang, and Bin Guo. 2014. Predicting activity attendance in event-based social networks: Content, context and social influence. In *Proceedings of UbiComp*. ACM, 425–434.
[12] Alon Elgar and Shlomo Bekhor. 2004. Car-rider segmentation according to riding status and investment in car mobility. *Transportation Research Record: Journal of the Transportation Research Board* 1894 (2004), 109–116.
[13] Martin Ester, Hans-Peter Kriegel, Jörg Sander, and Xiaowei Xu. 1996. A Density-based Algorithm for Discovering Clusters a Density-based Algorithm for Discovering Clusters in Large Spatial Databases with Noise. In *Proceedings of KDD*.
[14] Foursquare. 2018. Foursquare Venue Recomendation API. https://developer.foursquare.com/docs/api/venues/explore. [Online; accessed 01-November-2018].
[15] Mike Gartrell, Xinyu Xing, Qin Lv, Aaron Beach, Richard Han, Shivakant Mishra, and Karim Seada. 2010. Enhancing group recommendation by incorporating social relationship interactions. In *Proceedings of ACM Group*. 97–106.
[16] Petko Ivanov Georgiev, Anastasios Noulas, and Cecilia Mascolo. 2014. The Call of the Crowd: Event Participation in Location-Based Social Services. In *Proceedings of ICWSM*. 141–150.

[17] Eric Gilbert, Karrie Karahalios, and Christian Sandvig. 2008. The network in the garden: an empirical analysis of social media in rural life. In *Proceedings of CHI*. 1603–1612.

[18] Samuli Hemminki, Petteri Nurmi, and Sasu Tarkoma. 2013. Accelerometer-based transportation mode detection on smartphones. In *Proceedings of SenSys*. 13.

[19] Cheng-Kang Hsieh, Hongsuda Tangmunarunkit, Faisal Alquaddoomi, John Jenkins, Jinha Kang, Cameron Ketcham, Brent Longstaff, Joshua Selsky, Betta Dawson, Dallas Swendeman, et al. 2013. Lifestreams: A modular sense-making toolset for identifying important patterns from everyday life. In *Proceedings of SenSys*.

[20] Anthony Jameson and Barry Smyth. 2007. Recommendation to groups. In *The Adaptive Web*. Springer, 596–627.

[21] Kasthuri Jayarajah, Youngki Lee, Archan Misra, and Rajesh Krishna Balan. 2015. Need accurate user behaviour?: pay attention to groups!. In *Proceedings of UbiComp*. ACM, 855–866.

[22] Nicholas D Lane, Li Pengyu, Lin Zhou, and Feng Zhao. 2014. Connecting personal-scale sensing and networked community behavior to infer human activities. In *Proceedings of the 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing*. ACM, 595–606.

[23] Defu Lian, Cong Zhao, Xing Xie, Guangzhong Sun, Enhong Chen, and Yong Rui. 2014. GeoMF: joint geographical modeling and matrix factorization for point-of-interest recommendation. In *Proceedings of KDD*. 831–840.

[24] Wei-Yin Loh. 2011. Classification and regression trees. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery* 1, 1 (2011), 14–23.

[25] Xin Lu, Changhu Wang, Jiang-Ming Yang, Yanwei Pang, and Lei Zhang. 2010. Photo2trip: generating travel routes from geo-tagged photos for trip planning. In *Proceedings of ICMI*. 143–152.

[26] Augusto Q Macedo, Leandro B Marinho, and Rodrygo LT Santos. 2015. Context-aware event recommendation in event-based social networks. In *Proceedings of RecSys*. 123–130.

[27] Judith Masthoff. 2004. Group modeling: Selecting a sequence of television items to suit a group of viewers. In *Personalized digital television*. Springer, 93–141.

[28] Anastasios Noulas, Salvatore Scellato, Neal Lathia, and Cecilia Mascolo. 2012. Mining user mobility features for next place prediction in location-based services. In *Proceedings of ICDM*. 1038–1043.

[29] Anastasios Noulas, Salvatore Scellato, Neal Lathia, and Cecilia Mascolo. 2012. A random walk around the city: New venue recommendation in location-based social networks. In *Privacy, Security, Risk and Trust (PASSAT), 2012 International Conference on Social Computing*. IEEE, 144–153.

[30] Chunjong Park, Junsung Lim, Juho Kim, Sung-Ju Lee, and Dongman Lee. 2017. Don't Bother Me. I'm Socializing!: A Breakpoint-Based Smartphone Notification System. In *Proceedings of CSCW*. 541–554.

[31] Moon-Hee Park, Han-Saem Park, and Sung-Bae Cho. 2008. Restaurant recommendation for group of people in mobile environments using probabilistic multi-criteria decision making. In *Asia-Pacific Conference on Computer Human Interaction*. Springer, 114–122.

[32] Elisa Quintarelli, Emanuele Rabosio, and Letizia Tanca. 2016. Recommending new items to ephemeral groups using contextual user influence. In *Proceedings of RecSys*. 285–292.

[33] Juan Ramos et al. 2003. Using tf-idf to determine word relevance in document queries. In *Proceedings of the first instructional conference on machine learning*, Vol. 242. 133–142.

[34] Katharina Reinecke, Minh Khoa Nguyen, Abraham Bernstein, Michael Näf, and Krzysztof Z Gajos. 2013. Doodle around the world: Online scheduling behavior reflects cultural differences in time perception and group decision-making. In *Proceedings of CSCW*. 45–54.

[35] Daniel M Romero, Katharina Reinecke, and Lionel P Robert Jr. 2017. The Influence of Early Respondents: Information Cascade Effects in Online Event Scheduling. In *Proceedings of WSDM*. 101–110.

[36] Weather Source. 2017. Weather Source API. https://weathersource.com/products/onpoint-api/. [Online; accessed 01-November-2018].

[37] Leon Stenneth, Ouri Wolfson, Philip S Yu, and Bo Xu. 2011. Transportation mode detection using mobile phones and GIS information. In *Proceedings of SIGSPATIAL*. 54–63.

[38] Hanghang Tong, Christos Faloutsos, and Jia-yu Pan. 2006. Fast Random Walk with Restart and Its Applications. In *Proceedings of ICDM*. 613–622.

[39] Wikipedia. 2018. Craigslist — Wikipedia, The Free Encyclopedia. http://en.wikipedia.org/w/index.php?title=Craigslist&oldid=836044084. [Online; accessed 01-November-2018].

[40] Wikipedia. 2018. Microwork — Wikipedia, The Free Encyclopedia. http://en.wikipedia.org/w/index.php?title=Microwork&oldid=828297165. [Online; accessed 01-November-2018].

[41] Jason Shuo Zhang, Khaled Alanezi, Mike Gartrell, Richard Han, Qin Lv, and Shivakant Mishra. 2018. Understanding Group Event Scheduling via the OutWithFriendz Mobile Application. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 1, 4 (2018), 175.

[42] Jason Shuo Zhang and Qin Lv. 2018. Hybrid EGU-based group event participation prediction in event-based social networks. *Knowledge-Based Systems* 143 (2018), 19–29.

[43] Jason Shuo Zhang and Qin Lv. 2019. Understanding Event Organization at Scale in Event-Based Social Networks. *ACM Trans. Intell. Syst. Technol.* 10, 2, Article 16 (Jan. 2019), 23 pages.
[44] Jason Shuo Zhang, Chenhao Tan, and Qin Lv. 2018. This is why we play: Characterizing Online Fan Communities of the NBA Teams. *Proceedings of the ACM on Human-Computer Interaction* 2, CSCW (2018), 197.
[45] Yu Zheng, Like Liu, Longhao Wang, and Xing Xie. 2008. Learning transportation mode from raw GPS data for geographic applications on the web. In *Proceedings of WWW*. 247–256.
[46] Yu Zheng, Lizhu Zhang, Xing Xie, and Wei-Ying Ma. 2009. Mining interesting locations and travel sequences from GPS trajectories. In *Proceedings of WWW*. 791–800.
[47] James Zou, Reshef Meir, and David Parkes. 2015. Strategic voting behavior in doodle polls. In *Proceedings CSCW*. 464–472.

Received August 2018; revised November 2018; accepted January 2019