A Novel Spatiotemporal Behavior-Enabled Random Walk Strategy on Online Social Platforms

Chenbo Fu, Yinan Xia, Xinchen Yue, Shanqing Yu, Yong Min, Qingpeng Zhang, *Senior Member, IEEE*, and Yan Leng

***Abstract*— Location-Based Social Networks have been widely studied in recent years; new approaches constantly developed to solve individuals' trajectory prediction tasks. However, most of these methods require sufficient data to learn individual features, which is not always satisfied in real situations, especially for online data. The digital data on human behavior typically follows a power-law distribution, indicating that only a few people have rich activities recorded while most people's behavioral data are limited. In order to overcome this hurdle, our work constructs the user behavior proximity network (UBPN) and proposes a new walking strategy based on this network that extracts the hidden information from the social contacts to substitute the unobserved behavioral information of an individual. Specifically, our proposed walking strategy has two walking paths, accounting for the temporal and social information on the ego users' and their alters' mobility activities. This walking strategy is model-agnostic and can be integrated with many existing walk-based deep learning methods. Our work applies the methods on two real-world datasets with rich spatiotemporal information and shows that the performances of the existing prediction methods improve significantly by integrating the proposed walking strategy.**

***Index Terms*— Mobility prediction, network science, social activity, user behavior proximity network (UBPN), walking strategy.**

I. Introduction

The world is connected by integrated social systems, and social networks are now denser than ever. For example, the "small-world" network [1] suggests the average distance between two real-world members is around six, while this distance is further reduced to around five in the online social network [2]. Driven by the internet and social media, new phenomena emerge (e.g., online and offline interaction) and generate great interest in social network analysis. The location-based social network (LBSN) is amongst these systems, which integrates social relationships and human mobility. Therefore, it plays an important role between the impact of social relationships and mobility on each other. Meanwhile, the ongoing process of the LBSN applications, such as Yelp, Gowalla, Foursquare, turns many aspects of our lives into digital data, including geographic locations, friend relationships, time stamps, interests, and hobbies. These data are being collected and become the main data source for studying human dynamics. Furthermore, the various socio-spatial properties of LBSN imply that the social network has a subtle influence on human mobility activities, thus promote the development of two topics, i.e., friendship and human mobility prediction [3]–[9]. These tasks can help the government better design the urban region, controlling the diseases and optimizing transportation strategies [3], [10]–[12], and also provide potential applications in the e-commerce platforms, for instance, discovering the potential interests of users and gaining economic impacts [13]–[16].

Human mobility prediction has attracted lots of attention from urban planning to dining recommendation [17]–[20]. The traditional models are mainly referred to as recommendation systems or sequential models, such as matrix factorization (MF) [4], [21] and Markov chain (MC) [22], [23]. The MF model decomposes the user location matrix to learn the general features of the crowd. The MC method models human behavior by learning the location transfer matrix. Later, Rendle *et al.* [24] proposed the model of factorizing personalized Markov chains (FPMCs) to combine the matrix decomposition model with the Markov chain model and use it for the next basket of recommendations. However, these methods ignore the spatiotemporal factors of human mobility. For example, the mobility patterns for one day and one month may differ substantially. In addition, one person's short-distance trajectory may be affected by his (her) work location, and a long-distance trajectory may be affected by the vacation location. Therefore, Cheng *et al.* [25] extended FPMC by embedding movement constraints and personalized Markov chains. Feng *et al.* [26] proposed a personalized ranking metric embedding method (PRME), which embeds geographic features based on the modeling of personalized check-in sequences. Xie *et al.* [27] captured the contextual preferences and dynamic interests by embedding the information in the movement sequence into the low-dimensional latent space. A recurrent neural network (RNN) is a powerful tool to capture the sequential features and achieve better performance in word embedding [28] and the sequential click prediction tasks [29]. These characteristics naturally introduce the RNN methods and their extensions to the prediction tasks. Spatial-temporal recurrent neural networks (ST-RNNs) proposed by

Manuscript received January 23, 2021; revised May 9, 2021 and July 5, 2021; accepted August 8, 2021. Date of publication August 27, 2021; date of current version May 30, 2022. This work was supported in part by the Zhejiang Provincial Natural Science Foundation of China under Grant LGF21G010003 and Grant LGF20F020016 and in part by the National Natural Science Foundation of China (NSFC) under Grant 71972164. *(Corresponding author: Chenbo Fu.)*

Chenbo Fu, Yinan Xia, Xinchen Yue, Shanqing Yu, and Yong Min are with the Institute of Cyberspace Security, College of Information Engineering, Zhejiang University of Technology, Hangzhou 310023, China (e-mail: cbfu@zjut.edu.cn; ynxiazj@yeah.net; 1138420318@qq.com; yushanqing@zjut.edu.cn; myong@zjut.edu.cn).

Qingpeng Zhang is with the School of Data Science, City University of Hong Kong, Hong Kong (e-mail: qingpeng.zhang@cityu.edu.hk).

Yan Leng is with the McCombs School of Business, The University of Texas at Austin, Austin, TX 78712 USA (e-mail: yan.leng@mccombs.utexas.edu).

Digital Object Identifier 10.1109/TCSS.2021.3105381

Liu *et al.* [30] predicts the next POI by modeling the spatiotemporal context. The model trains individual data into several segments based on a time window, allowing it to learn short-term interests. Leng *et al.* [12] integrate RNN into a recommendation system to predicts next locations using cell phone traces. Li *et al.* [31] proposed temporal and multi-level context attention (TMCA) using the LSTM-based framework to predict the next POI. The contextual attention recurrent architecture (CARA) model proposed by Manotumruksa *et al.* [32] uses the spatiotemporal context of check-ins information to capture the human dynamic preferences through a gating mechanism. In general, most RNN/LSTM based methods can model short-term preferences, but they are not as effective for long-term preferences. However, if the data are sparse, the inter-event time between two consecutive activities experiences a long gap, and the model based on long-term preference will perform better. Therefore, DEEPMOVE [33] uses multi-modal RNN to capture multiple factors when humans are moving, and two attention mechanisms are integrated to learn long-period features. The spatiotemporal gated network (STGN) [34] improved the ST-RNN model and introduced a gating mechanism to enhance long- and short-term memory.

Although these methods have achieved encouraging results, social information, including social relationship (e.g., friends) and their online social activities (e.g., post, share, and like) are still be ignored in the literature. Recent research finds that human behaviors are closely related to their social relationship [5], [6], [35]. For example, people are more likely to travel to places where their friends have visited [36], [37], even the running exercises are affected by social networks [38]. Moreover, Cho *et al.* [5] found that in LBSN, social relations can explain about 10% to 30% of human movements and Gao *et al.* [8] found that the human mobility behavior is determined by the social network and historical behavior of the individual. Furthermore, there are two reasons that integrating social information is effective. The first reason is that although the enormous mobility data can be collected, it suffers from data sparsity issue due to the high cost of data collection and privacy concerns. Unlike the traffic datasets, the mobility data collected from LBSN apps are sparse and heterogeneous. For example, in the Yelp dataset, most users visit restaurants with a few number of reviews [19], making the models difficult to capture effective features and leading to poor performance. Second, there is an upper limit for user behavior prediction [39]. Without introducing new information, the model cannot exceed the prediction limit. Thus, introducing additional easy-collected information, such as social information, become a possible solution. Based on this, recent studies introduce social information to improve the performance of the models. For example, Yang *et al.* [40] designed a method for embedding user mobility and social relations simultaneously and mapping user movement behavior to a low-dimensional space based on traditional embedding. Lian *et al.* [41] proposed a model called collaborative exploration and periodically returning (CEPR), which learns the characteristics of users' social networks and geographic factors and predicts the probability of people exploring new locations. Luceri *et al.* [42] proposed a framework named social influence deep learning (SIDL), which combines deep learning and network science to model social influences from the local, community, and global perspectives to predict human dynamics. Although these models have achieved good results after introducing social information, to our knowledge there are few models that utilize social information are used for user location prediction.

Compared to the mobility data, the social information collected from the internet is relatively dense [9], [43], which is proven to improve the model's performance. However, on the one side, dense data may contain lots of irrelevant information. On the other side, using online social information to supplement sparse location information effectively has become a key direction. Recently, Bagrow and Mitchell [44] found that in the online social platform, e.g., Twitter, part of individual information is stored in the online social networks and utilizing a few friends' online information, e.g., 8 and 9 of an individual's contacts, the prediction accuracy can reach the same level as that of the individual alone. Being inspired by this work and considering that human mobility is also one type of sequence data, our study integrates the social information and proposes a new walking strategy. Specifically, by constructing a user behavior proximity network (UBPN) and considering the time and space proximity constraint, our walking strategy can filter the noise information and dig out the hidden information in the social networks. In more detail, the main contributions of our work are as follows.

1) First, this study shows that social information provides a substitute for missing information of individuals in two real datasets.
2) Second, the UBPN is constructed, and based on this network, a model-agnostic new walking strategy is proposed. This walking strategy does not change the structure of the existing prediction model and can be integrated with many existing walk-based deep learning methods.
3) Finally, our walking strategy is integrated into six baseline methods. The experimental results show that our method outperforms the other baseline methods. Furthermore, our method also compares with other social information involved models and the results also validated the effectiveness of our method in extracting social information.

The rest of the article is organized as follows. In Section II, introduces the dataset. Section III presents the theoretical analysis. In Section IV introduces the method. The experiments and results are shown in Section V. Finally, Section VI concludes with some discussions.

II. Datasets

The present study uses two publicly available datasets, i.e., Gowalla data[1] and Yelp (Las Vegas) data.[2] Gowalla is a location-based social network, this dataset contains global-scale check-ins and social networks from

[1] http://snap.stanford.edu/data/loc-gowalla.html
[2] https://www.yelp.com/dataset/challenge

TABLE I
DESCRIPTIVE STATISTICS OF OUR DATASETS

	#Review	#Location	#User	Time Frame
Gowalla	6,442,892	1,280,969	107,092	02/2009-10/2010
Yelp (Las Vegas)	685,336	23,143	284,859	01/2016-01/2018

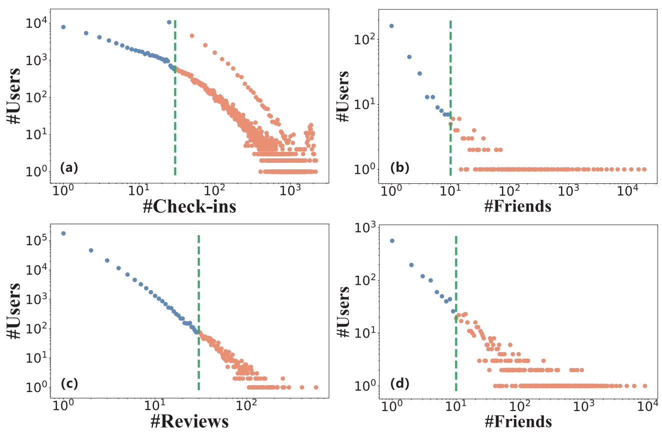


Fig. 1. Statistical distribution of (a) Gowalla check-ins; (b) Gowalla friends; (c) Yelp (Las Vegas) restaurant reviews; and (d) Yelp (Las Vegas) friends. The blue dots mean most active users, i.e., the users who have more than ten friends and submit at least thirty check-ins (reviews).

February 2009 to October 2010. Yelp is a popular crowd-sourced business reviewing website. This dataset contains the restaurant and user information. The restaurant information includes the latitude and longitude, city name, the overall score, etc. The user information contains the comment, comment time, and social relationship. For comparability, our work study two years of data (i.e., from January 2016 to January 2018) and the city with the largest amount of data (i.e., Las Vegas) in the Yelp dataset. The overview of these two datasets is presented in Table I. The user distribution of the Gowalla and Yelp (Las Vegas) dataset is shown in Fig. 1. Fig. 1(a) and (b) shows the Gowalla datasets statistical distribution, and Fig. 1(c) and (d) shows the statistical distribution of the Yelp (Las Vegas) dataset. It can be seen that both datasets are subject to a power-law distribution, which means that most people have less than 30 historical activities and less than ten friends.

These geotagged data with temporal information allow us to quantitatively examine human mobility patterns at a city-scale automatically. However, the Gowalla check-in data lacks information about the cities. To solve this, our study extracts information about the city through the Geocoder API.[3] Furthermore, although the Yelp data contains check-in data, the total amount is not sufficient. Therefore, our work uses the review data to replace check-in data and extract geographic information from the restaurant data. Considering that the users rarely transfer between cities, this study only focuses on the largest city in the dataset, i.e., Las Vegas. And to get statistically meaningful results, our study focuses on the most active users in two datasets. More specifically, the users who have more than ten friends and submit at least thirty check-ins (reviews) are selected. Finally, there are a total of 3 30 487 check-ins from 2299 users on 19 520 POIs in Gowalla data and 7031 reviews from 1592 users on 62 976 restaurants in Yelp (Las Vegas) data.

[3]https://geocoder.readthedocs.io/providers/ArcGIS.html

III. THEORETICAL ANALYSIS

Although data collection from online social platforms is convenient, the available data for most individuals is insufficient [45], [46]. For instance, as shown in Fig. 1, most individuals only have 1~2 check-ins (reviews) and friends. Thus, the resulting latent representations lead to poor performance of traditional machine learning methods to predict individuals' trajectory [33], [47]. One way to mitigate this problem is to utilize social network information. It has been demonstrated that social network information can help the relevant model to obtain a substantial performance improvement [40]–[42]. Thus in this section, the focus is on the feasibility of utilizing social information and the effect of data proximity from an information entropy perspective.

The amount of information can be represented by the information entropy, which is usually used in the natural language process, e.g., measuring the information contained in the text [48], [49]. Similar to the text, human mobility patterns are also in the form of sequences; therefore, this study uses the information entropy [49] to estimate the latent trajectory which hides in the friend's mobility data. Defining human mobility entropy as $\hat{h}$, which implies the amount of the information needed to predict the next location, then the $\hat{h}$ can be calculated as follows:

$$\hat{h} = \frac{L \log_2 L}{\sum_{i=1}^{L} m_i} \tag{1}$$

where L represents the length of the location sequence of the individual. m_i represents the matching length of the individual i's location, which represents the shortest length of a series of subsequent mobility behaviors that never appear in the historical data starting from the i-th location.

When considering the social relations, it can easily expand the mobility entropy to cross mobility entropy, which takes the users (the ego) and their contacts (the alter contacts) into consideration. Specifically, the calculation for cross mobility entropy is as follows:

$$\hat{h}_\times(A|B) = \frac{L_A \log_2 L_B}{\sum_{i=1}^{L_A} m_i(A|B)}. \tag{2}$$

Denoting A as the ego user and B as the alter contact, where $m_i(A|B)$ means starting at the i location, the shortest matching length of subsequent mobility behaviors for ego user A that never appeared in alter contact B's historical mobility data. Here, the historical mobility data represents all mobility behaviors of the alter contact B occurring earlier than location i of ego user A. L_A and L_B represent the mobility behavior sequence length of user A and B, respectively. Compared with the mobility entropy, the cross mobility entropy replaces the L with L_B in log term of molecular because the cross mobility entropy depends on the historical mobility data of alter contact B. When the ego user has more than one alter contact, the information provided by the alter contacts can

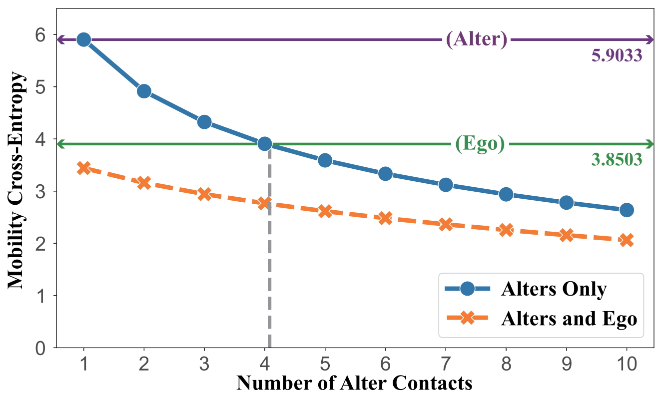


Fig. 2. Mobility entropy as functions of the number of alter contacts on Gowalla dataset. The green line represents ego user mobility entropy and the purple line represents a single alter contact's cross mobility entropy. The blue line represents cumulative cross mobility entropy of alter contacts and the orange line represents cumulative cross mobility entropy of both ego user and alter contacts.

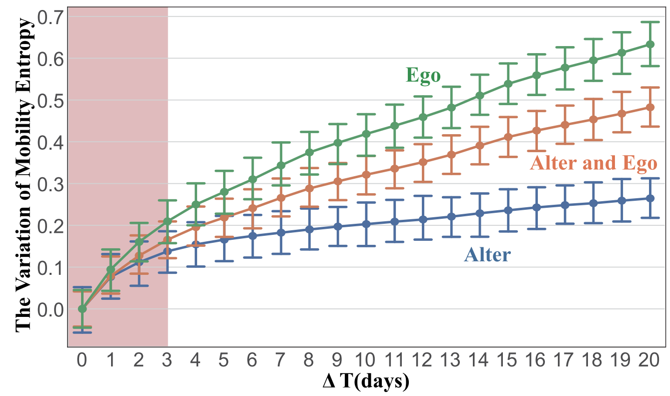


Fig. 3. Variation of mobility entropy as functions of removing the activities for recent ΔT days. The largest part of the variation of mobility entropy occurred in the first three days (shaded red region), with a change of nearly 2.1%.

be calculated by cumulative cross mobility entropy, which represents the average amount of information needed in the friends' mobility data

$$\hat{h}_{\times}(A|\mathcal{B}) = \frac{L_A \log_2 L_{A\mathcal{B}}}{\sum_{i=1}^{L_A} \max\{m_i(A|B), B \in \mathcal{B}\}} \tag{3}$$

where $\mathcal{B}$ is the social contacts of ego user A, $\max\{m_i(A|B), B \in \mathcal{B}\}$ corresponds to the longest match length over any of the sequences of the alter contacts in $\mathcal{B}$. $L_{A\mathcal{B}} = \sum_{B\in\mathcal{B}} \omega_B L_B / \sum_{B\in\mathcal{B}} \omega_B$ is the average weighted mobility behavior sequence length, where the weight ω_B means the number of match time for the alter contact B. Thus, the information contained in the alter contacts is helpful for the ego users, which can be calculated by the cross mobility entropy or the cumulative cross mobility entropy.

Fig. 2 shows the experimental results of mobility entropies on the Gowalla dataset, including user mobility entropy, cross mobility entropy, and cumulative cross mobility entropy. The lower value of the entropy implies high predictability for users. From Fig. 2, it can be seen that the mobility entropy of the ego user (green line) is about 3.85, which is lower than cross mobility entropy (purple line) of the alter contact ($\approx$ 5.90). Thus, the ego information contained in one alter contact is limit. It should be noted that although a single alter contact contains less information about the ego user, when the number of alter contacts increases, the information about the ego user also increases, as shown in Fig. 2 (blue line). When the number of alter contacts is more than 4, the cumulative cross mobility entropy is lower than ego user's mobility entropy, which implies as long as there are enough alter contacts, the ego user information can be speculated from alter contacts. Furthermore, the value of the cumulative cross mobility entropy can further decrease after adding more alter contacts, thus implies the alter contacts may contain the missing information of the ego users and suggests that the information of alter contacts can be utilized to predict individual trajectory. Notably, Wang *et al.* [50] pointed out that human attention is limited; in other words, people will only pay attention to a limited number of friends. And as the amount of alter contacts' information increases, the noise information irrelevant to the ego users will also be introduced. Therefore, to tackle this problem, it should choose the right number of alter contacts to balance the amount of extra information and the performance of the model. Moreover, the results also show that the mobility entropy can further decrease (orange line) when considering both ego and alter information.

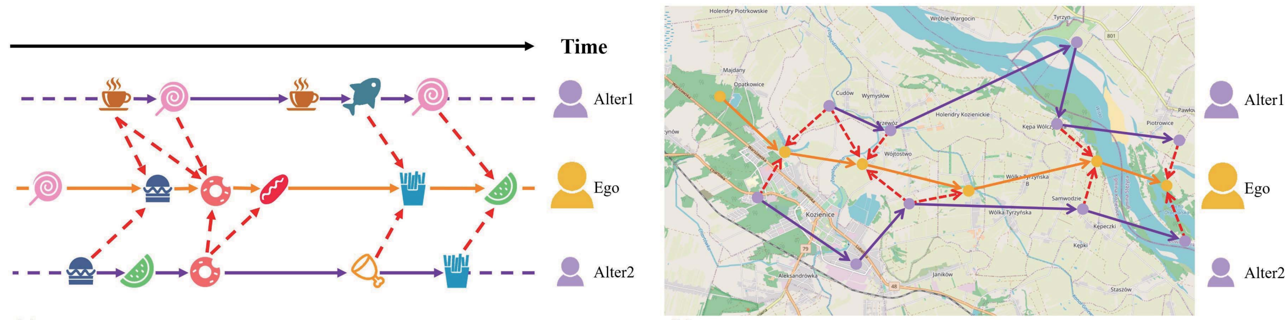
The recent mobility patterns will infer the individual's future mobility trajectory [5], [49]. To estimate the recent effects, the variation of mobility entropy on the Gowalla dataset is plotted in Fig. 3. As shown in Fig. 3, after removing the most activities of recent ΔT days, all the entropies increase gradually and the most important ascending area is in the first three days (shaded red region, with a change of nearly 2.1%). This finding suggests that the time window of individual activities can be chosen as one to three days. Additionally, the increase in cross mobility entropy (blue line) also implies that ego's mobility behavior refers to alter contacts behavior.

IV. Methodology

Most of the existing position prediction models utilize single-source information only and ignoring the social information that may contain lots of hidden information of ego users. In this section, the hidden information is extracted and integrated into the ego user's mobility pattern to help improve the prediction model's performance. To achieve this, the UBPN is constructed and the hidden information of ego users is sampled on this network through the random walk.

A. User Behavior Proximity Network

The LBSN contains both the activity data and the social relationship of ego users. However, this network ignores the mutual influence among friends. As mentioned in Section II, these recent mobility activities of alter contacts have a significant impact on ego users. Thus, to improve the current LBSN, the UBPN is constructed by introducing the temporal social inference into user behavior activity. A simple illustration is shown in Fig. 4(a), in this network, the node represents the mobility activity of the ego or the alter contact, which is represented as the food icon, and the link represents the impact between the activities. Where the solid arrow means



(a) (b)

Fig. 4. Simple illustration of UBPN in (a) mobility activity space and (b) real space. Where the food icons represent mobility activities of individuals, the yellow and purple solid arrow represent direct influence by the individual (the ego or alter contact) self and the red dashed arrow represents indirect influence by individual's alter contacts.

direct influence by the individual self (historical activities) and the dashed arrow represents the indirect influence by the individual's alter contacts. Additionally, Fig. 4(b) is the same network in real space. In this study, the temporal social inference mainly focuses on user proximity, including both ego and alter contacts. More specifically, the ego user proximity implies that the ego users are more likely to travel to the locations that were recently visited in the mobility sequences. This impact is marked as the solid line (yellow and purple) in Fig. 4. Similarly, the ego users may be affected by alter contacts, which is marked as the red dotted line in the figure. These recency features are closer to reality, for example, the ego users may often visit their favorite restaurants or accept a friend's suggestion to visit another restaurant in a limited time window. Therefore, the alter proximity should follow the following conditions:

$$\delta = \begin{cases} 1, & T_{\text{ego}} - T_{\text{alter}} < \Delta T \,\&\, C_{\text{ego}} = C_{\text{alter}} \\ 0, & \text{else} \end{cases} \tag{4}$$

where $T_{\text{ego}}, T_{\text{alter}}$ represent the visited time of ego and alter contacts respectively and ΔT means time window. $C_{\text{ego}}, C_{\text{alter}}$ represent the visited area of ego and alter contacts, respectively. The reason for using the visited areas instead of the specific visited locations because those locations close to each other exhibit similar characteristics [25] and induce similar foraging costs on users [19], i.e., restaurants around the commercial district. Here, the visited locations are clustered to areas by density-based spatial clustering of applications with noise (DBSCAN), which can effectively find clusters of arbitrary shapes and do not need the number of clusters [19], [51], [52]. There are two parameters for DBSCAN, i.e., neighborhood density threshold MinPts and the clustering radius ε. After specifying each location must belong to a single cluster, the neighborhood density threshold MinPts $= 1$. And by adopting the method in [53], the clustering radius ε for city size can be automatically determined. As an example, Fig. 5 shows the clustering of the restaurants in Las Vegas city by using DBSCAN, in which different colors represent different areas.

B. Random Walk Strategy

To extract the hidden information from the alter contacts, a random walk strategy is designed based on UBPN. A more

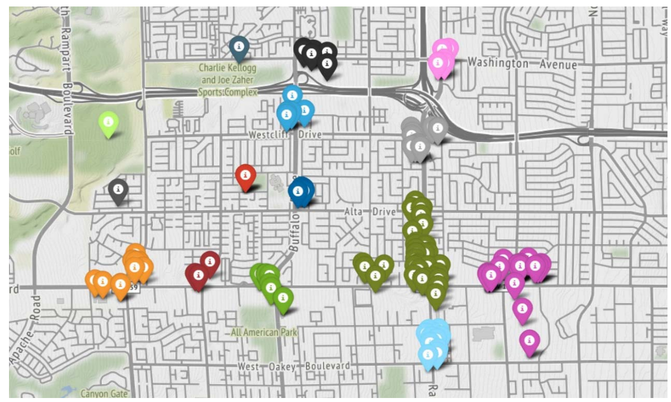


Fig. 5. Clustering of the restaurants in Las Vegas city by using DBSCAN.

detailed illustration is shown in Fig. 6. Our walking strategy has two walking paths, the main walking path sample the data of the ego users, and the auxiliary walking path sample the data of the alter contacts. Where the blue solid arrow represents the walking direction. It can be seen that, in our model, the walking direction is in the reverse order of time series. Formally, choosing an ego user, the shaded yellow region is the data that needs to be supplemented. Assume that the random walk starts from the mobility activity M1, if the in-degree of activity M1 is greater than 1 ($k_i^{\text{in}} > 1$), which implies the ego user is affected by alter contacts, then the walking strategy has a chance to go through the auxiliary walking path which is based on the data of the alter contacts. Otherwise, it will sample the ego user's history data only, i.e., the main walking path. The auxiliary walking path supplements the missing data of the ego users. The start activity of the auxiliary walking path will be chosen by the most recent activity of the alter contact (e.g., M2) in the relative time window (the shaded green region). And the roaming depth of the auxiliary walking path is defined by $dp = k$, which is used to limit the sampling depth of the alter contacts. To merge the auxiliary walking path to the main walking path, the return activity (e.g., M3) must occur before the end activity of the auxiliary walking path. Hence, in our example, the return activity node is M4. Specifically, the transition probability can be written as

$$\text{Pro} = \begin{cases} 1, & k_i^{\text{in}} = 1 \\ \delta * [\alpha * T + \beta * \Pi^{\max} + (1 - \alpha - \beta) * D], & k_i^{\text{in}} > 1 \end{cases} \tag{5}$$

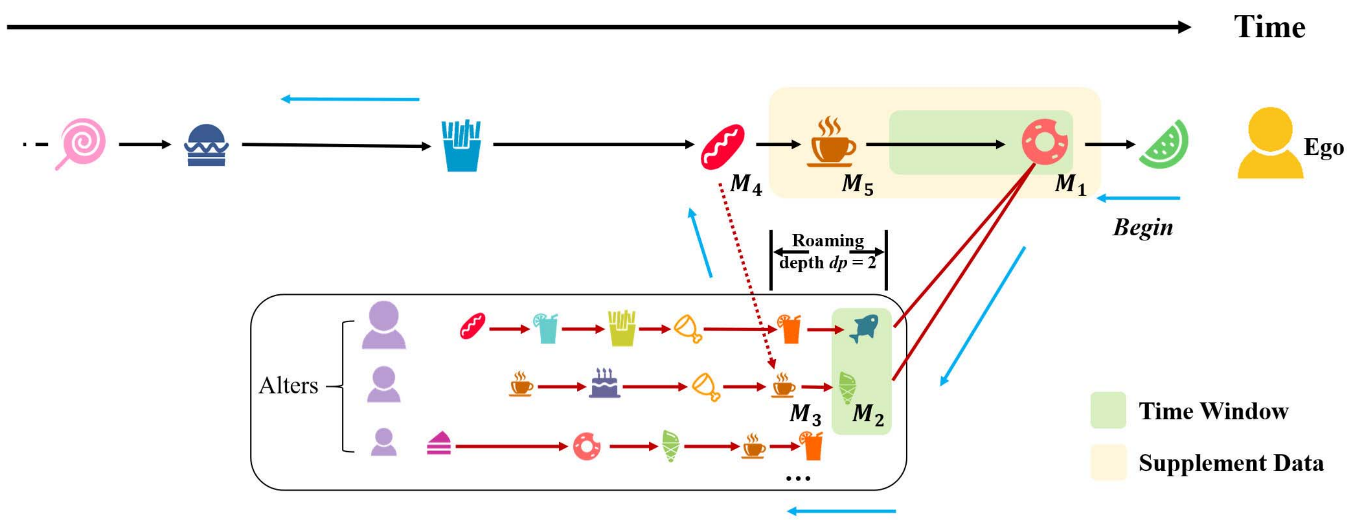


Fig. 6. Illustration demonstrating the proposed walking strategy on the UBPN. The blue solid arrow indicates the direction of walking.

where $T = (1 - 0.56 * \Delta T^{0.06})$ represents the time effect between two activities, which is based on Ebbinghaus's forgetting curve [54]. It is intuitive to see that the longer the time interval between two activities, the lower the transition probability. $\Pi^{\max}$ represents the highest prediction limit based on historical data of the individuals [39], alternatively, how much useful information is hidden in the alter contacts. The $\Pi^{\max}$ can be calculated as follows:

$$\begin{cases} \Pi \leq \Pi^{\max}(\hat{h}_\times, L) \\ \hat{h}_\times = H(\Pi^{\max}) + (1 - \Pi^{\max}) \log_2 L - 1 \\ H(\Pi^{\max}) = -\Pi^{\max} \log_2(\Pi^{\max}) \\ -(1 - \Pi^{\max}) \log_2(1 - \Pi^{\max}) \end{cases} \tag{6}$$

where $H(\Pi^{\max})$ is the binary entropy function. The high value of $\Pi^{\max}$ implies that lots of ego user's information is hidden in alter user's mobility patterns, resulting in a higher transfer probability between the activities of ego and alter contacts. Thus, $\Pi^{\max}$ provides an easy-to-explain way to find out which friends provide more information according to prediction limits. The distance parameter is denoted as $D = e^{-\Delta \text{dist}}$ in (5), and the Δdist means euclidean distance between two adjacent activities related locations. α, β are the hyperparameters that adjust the time and social effect. It should be noted that, for different ego users, there may have large variations in the number of alter contacts, from a few to a hundred. However, human attention is limited, people will only focus on those friends their care about most [50]. For this reason, the alter contacts we chose for ego users is defined as $N'_f = \min[N_f, S_N]$, where N_f is the actual number of friends of the ego users and S_N is the sample number which is fixed as five in our experiments.

The advantage of this walking strategy is that the strategy not only samples the direct mobility activities of ego users but also samples the indirect mobility activities of alter contacts, i.e., the activities that may have large impacts on ego user's current activities. And these impacts are constrained by time and space factors. Furthermore, by adopting this strategy, the noise contained in the information and computation cost is reduced, which creates a tradeoff between balancing computational efficiency and prediction performance. Compared to the traditional prediction methods which only considering the ego user's information but ignoring the social impacts. Our walking strategy makes up for this shortcoming and extracts the hidden information of the ego users by considering the temporal and social impact.

V. Experiments and Results

Our method is performed in the two datasets mentioned in Section II, i.e., Gowalla and Yelp (Las Vegas). For the experiments, 80% of historical user mobility data is used for training, and the remaining 20% for testing.

A. Baselines

To validate the effectiveness of our method, the proposed walking strategy is integrated into six prediction methods for comparison, i.e., FPMC [24], PRME [26], RNN [29], GRU [55], LSTM [56] and DEEPMOVE [33]. These methods are widely used for behavior prediction and briefly introduced as follows.

1) *FPMC [24]:* This is a classical sequence prediction method based on the Markov chain.
2) *PRME [26]:* This method takes into account the spatial distance interval of neighboring behaviors when learning embeddings.
3) *RNN [29]:* This is an efficient method of temporal prediction widely used for word embedding and ad click prediction.
4) *GRU [55]:* This is a variation of the RNN model equipped with two gates to control the flow of information.
5) *LSTM [56]:* This is another variation of the RNN model, which has shown a strong ability to handle sequential data and learn to remember features over time.
6) *DEEPMOVE [33]:* This method uses the attention mechanism to learn the user's long-term preference from the user's behavioral history and then uses RNN to learn the user's short-term preference.

TABLE II

OVERALL PERFORMANCE ON GOWALLA DATASET OF DIFFERENT MODEL IN TERMS OF ACC, RECALL AND F1

Model	isWalk	acc@1	acc@5	acc@10	recall@1	recall@5	recall@10	f1@1	f1@5	f1@10
RNN	-	0.12438	0.24467	0.29380	0.11840	0.22862	0.26970	0.11840	0.07621	0.04904
	Yes	**0.15026**	**0.29933**	**0.36393**	**0.15182**	**0.29354**	**0.34274**	**0.15182**	**0.09785**	**0.06232**
GRU	-	0.12965	0.25593	0.30791	0.12499	0.23750	0.28219	0.12499	0.07917	0.05131
	Yes	**0.15767**	**0.31315**	**0.38028**	**0.15529**	**0.29675**	**0.35215**	**0.15529**	**0.09892**	**0.06403**
LSTM	-	0.12556	0.24764	0.30049	0.11563	0.22554	0.26854	0.11563	0.07518	0.04882
	Yes	**0.14440**	**0.28610**	**0.34674**	**0.14380**	**0.27526**	**0.32433**	**0.14380**	**0.09175**	**0.05897**
FPMC	-	0.10149	0.24365	0.31041	0.10433	0.23964	0.30065	0.10433	0.07988	0.05466
	Yes	**0.10419**	**0.24817**	**0.31984**	**0.11044**	**0.24874**	**0.31907**	**0.11044**	**0.08291**	**0.05801**
PRME	-	0.08313	0.22193	0.28838	0.10462	0.23682	0.28856	0.10462	0.07894	0.05247
	Yes	**0.08962**	**0.23046**	**0.29267**	**0.10832**	**0.23692**	**0.29284**	**0.10832**	**0.07897**	**0.05324**
DEEPMOVE	-	0.13841	0.28299	0.34680	0.10469	0.21955	0.26977	0.10469	0.07318	0.04905
	Yes	**0.15903**	**0.32124**	**0.38401**	**0.13313**	**0.27396**	**0.32284**	**0.13755**	**0.09004**	**0.05846**

TABLE III

OVERALL PERFORMANCE ON YELP (LAS VEGAS) DATASET OF DIFFERENT MODEL IN TERMS OF ACC, RECALL AND F1

Model	isWalk	acc@1	acc@5	acc@10	recall@1	recall@5	recall@10	f1@1	f1@5	f1@10
RNN	-	0.00266	0.00852	0.01385	0.00254	0.00833	0.01631	0.00254	0.00319	0.00257
	Yes	**0.05278**	**0.08645**	**0.10679**	**0.03330**	**0.05913**	**0.07292**	**0.03330**	**0.01971**	**0.01326**
GRU	-	0.00346	0.01012	0.01571	0.00284	0.01077	0.01699	0.00284	0.00359	0.00309
	Yes	**0.06193**	**0.11958**	**0.14704**	**0.03017**	**0.08033**	**0.10413**	**0.03017**	**0.02678**	**0.01893**
LSTM	-	0.00241	0.00799	0.01465	0.00258	0.00889	0.01615	0.00258	0.00296	0.00294
	Yes	**0.05323**	**0.09801**	**0.11151**	**0.03268**	**0.06372**	**0.07889**	**0.03268**	**0.02124**	**0.01434**
FPMC	-	**0.00257**	0.00973	0.02047	**0.00271**	0.01008	0.02075	**0.00271**	0.00336	0.00377
	Yes	0.00212	**0.01007**	**0.02058**	0.00200	**0.01066**	**0.02144**	0.00200	**0.00355**	**0.00390**
PRME	-	0.00232	0.00980	0.01998	0.00256	0.01093	0.02267	0.00256	0.00364	0.00412
	Yes	**0.00252**	**0.01245**	**0.02189**	**0.00275**	**0.01310**	**0.02380**	**0.00275**	**0.00437**	**0.00433**
DEEPMOVE	-	0.00239	0.01145	0.01997	0.00297	0.00920	0.01708	0.00297	0.00307	0.00310
	Yes	**0.06208**	**0.12054**	**0.14943**	**0.05266**	**0.07888**	**0.09120**	**0.05266**	**0.02790**	**0.01936**

B. Evaluation Metrics

In this work, three common evaluation metrics are used to evaluate our method and baseline methods: mean accuracy@N, mean recall@N, and mean F1-score@N. These indicators are widely used to evaluate prediction results [27], [57]–[59]. Let R denote the activities in the test set.

1) *Mean Accuracy@N:* Mean accuracy@N is defined as the proportion of the number of correct predicted locations in the test set, that is

$$\text{Accuracy@N} = \frac{\sum_{r\in R} \text{isCorrect}(r, \text{Top@N})}{|R|} \tag{7}$$

where is Correct$(r, \text{Top@N})$ returns 1 if the location r of the user visited is included in the top N prediction list, otherwise returns 0.

2) *Mean Recall@N:* For all the behavioral data R and all users U in the test set, let R_i represents the behavior sequence of user i, then the mean recall@N can be calculated as

$$\text{Recall@N} = \frac{1}{|U|}\sum_{U_i\in U}\frac{\sum_{r\in R_i} \text{isCorrect}(r, \text{Top@N})}{|R_i|}. \tag{8}$$

3) *Mean F1@N:* Similarly, the mean precision@N can be calculated as (9), then the mean F1@N can be calculated as (10)

$$\text{Precision@N} = \frac{1}{|U|}\sum_{U_i\in U}\frac{\sum_{r\in R_i} \text{isCorrect}(r, \text{Top@N})}{|R_i| \times N} \tag{9}$$

$$\text{F1@N} = \frac{2 \times \text{Precision@N} \times \text{Recall@N}}{\text{Precision@N} + \text{Recall@N}}. \tag{10}$$

The performance of our proposed walking strategy in combination with six baselines on two datasets are shown in Tables II and III, respectively, where $N \in \{1, 5, 10\}$. The results show that all the performance of the prediction models improved after integrated the proposed walking strategy, especially for these deep learning-based methods, e.g., RNN, GRU, LSTM, and DEEPMOVE. For example, in the deep learning-based methods over the Yelp (Las Vegas) dataset, all performance indicators have improved nearly nine times, and on the Gowalla dataset can enhance performance by about 22%. This can be explained by the primary role of the proposed walking strategy, which is to encode the general social interaction information in the mobility sequence of users. Specifically, the stronger the model's ability to learn the sequence, the more obvious the effect of prediction performance improvement. For example, compared with the Markov-based model, the RNN-based model has a stronger ability to learn the long-term features in the sequence data, which leads to a better performance in results. Furthermore, the quality of the dataset affects performance improvement. For example, the density (check-ins/users) of the Yelp (Las Vegas) dataset is 39.56, which is much sparser than the Gowalla dataset (143.75). After integrating the proposed walking strategy, the performance improvement on the Yelp (Las Vegas) is significantly higher than that on the Gowalla dataset. All performance indicators on the Yelp dataset are nearly eighteen times better than those on the Gowalla dataset since more missing data can be supplemented by our method on sparse datasets.

C. Parameter Sensitivity

The proposed walking strategy involves a number of parameters, so it is necessary to examine the parameter sensitivity. Choosing the GRU model as the benchmark

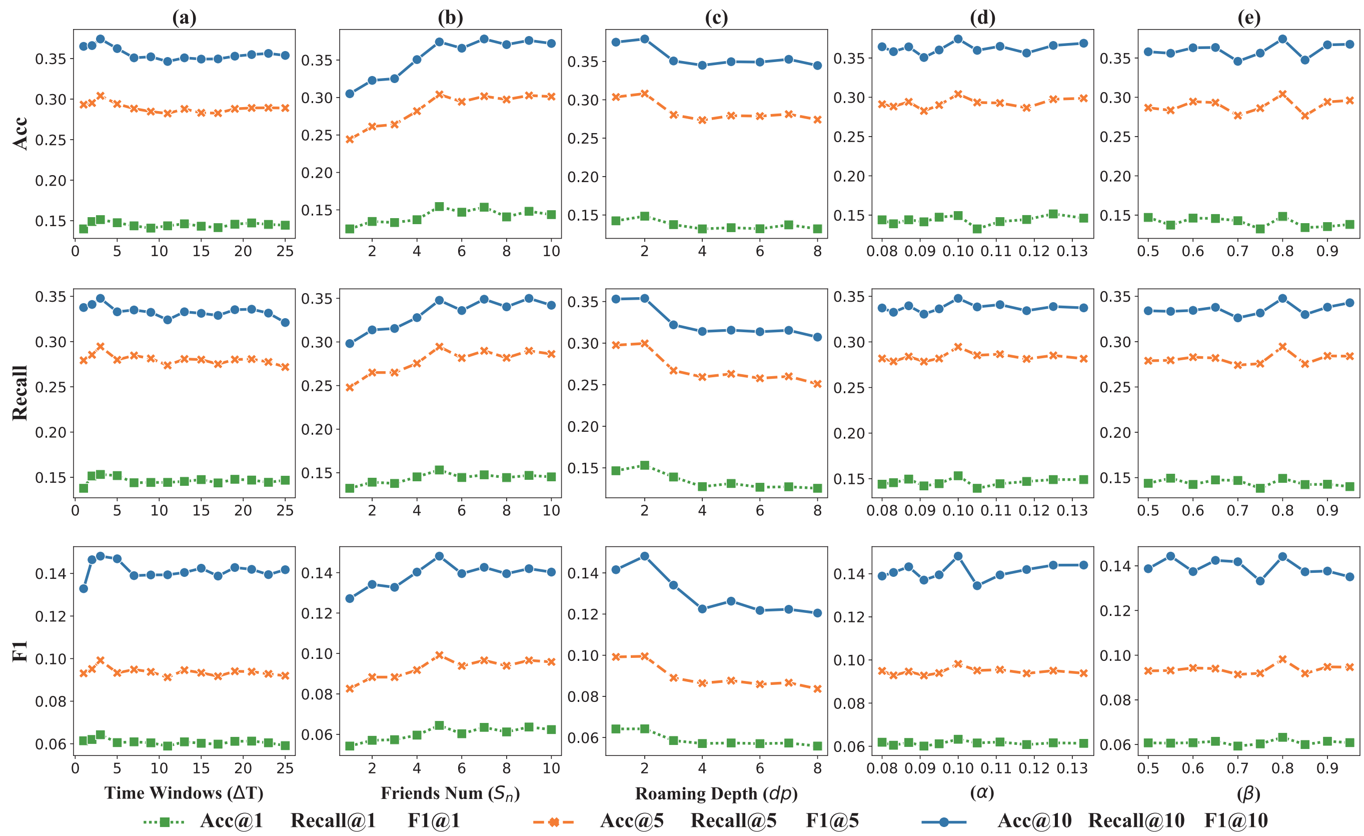


Fig. 7. Performance (Acc, Recall, and F1) comparison over the different walking parameters on the Gowalla dataset. (a) shows the performance as the time window change, (b) shows the performance change of the number of friends, (c) shows the performance change of roaming depth on alter, and (d) and (e) show the influence of the hyperparameters in (5) on performance.

model, Figs. 7 and 8 show the performance of the main parameters in our walking strategy on Gowalla and Yelp (Las Vegas) datasets, respectively, i.e., time window (ΔT), alter number (S_N), roaming depth (dp) and hyperparameters (α, β). The sensitivity of different time windows is shown in Figs. 7(a) and 8(a), respectively. It can be found that the performance of the model increases first and then gradually decreases, especially in Yelp (Las Vegas) dataset (the peak value is about thirteen days). This implies the impact of alter contacts on ego users will not happen immediately, which needs to take some time to reach the optimum. It is understandable, for example, if one friend posted a visited location on the web, then it may take some time for the ego user to see the post and be affected according to her situation. The reason why the optimal time window of the Yelp (Las Vegas) dataset ($\approx$ 13 days) is longer than that of the Gowalla dataset ($\approx$ 3 days) may be caused by the attributes of the dataset. The Gowalla dataset mainly consists of check-in data, which have high real-time performance. Comparing to the Gowalla dataset, the Yelp (Las Vegas) dataset consists of review data, which leads to low real-time performance. The performance decreases as the time window continues to increase, which is mainly because a time window being too long introduces a lot of noise. For example, the time window equals one year in the extreme. The sensitivity of different alter numbers is shown in Figs. 7(b) and 8(b), respectively. It can be seen that the performance of the model improves as the alter number increases, and then reaches saturation, thus

TABLE IV

COMPARISON OF DIFFERENT MODELS ON YELP (LAS VEGAS) DATASET

Model	recall@1	f1@1
DGRec	0.00539	0.01011
FM	0.03062	0.05251
NFM	0.03049	0.05213
CKE	0.03144	0.05262
DEEPMOVE	0.00297	0.00296
DEEPMOVE (ours)	**0.05266**	**0.05267**

most because the attention of users is limited. The saturation number of the Yelp (Las Vegas) dataset is larger than that of the Gowalla dataset is also because the Yelp (Las Vegas) dataset is more sparse and contains more missing data. The results in Figs. 7(c) and 8(c) also imply that too deep roaming depth introduces the noises of the alter contacts, which is not useful for supplementing the ego user data. The sensitivity of hyperparameters is shown in Fig. 7(d) and (e) (see Fig. 8), it can be found that the performance of the model is not affected by the hyperparameters. This is mostly because the walking strategy has already been restricted to UBPN, so the hyperparameters α and β can only play a fine-tuning role, and cannot affect the model performance greatly.

D. Effectiveness in Extracting Social Information

To validate the effectiveness of our method in extracting social information, our study compares our method with the social information involved method (DGRec [60]) and three traditional models which including latent social information as follows.

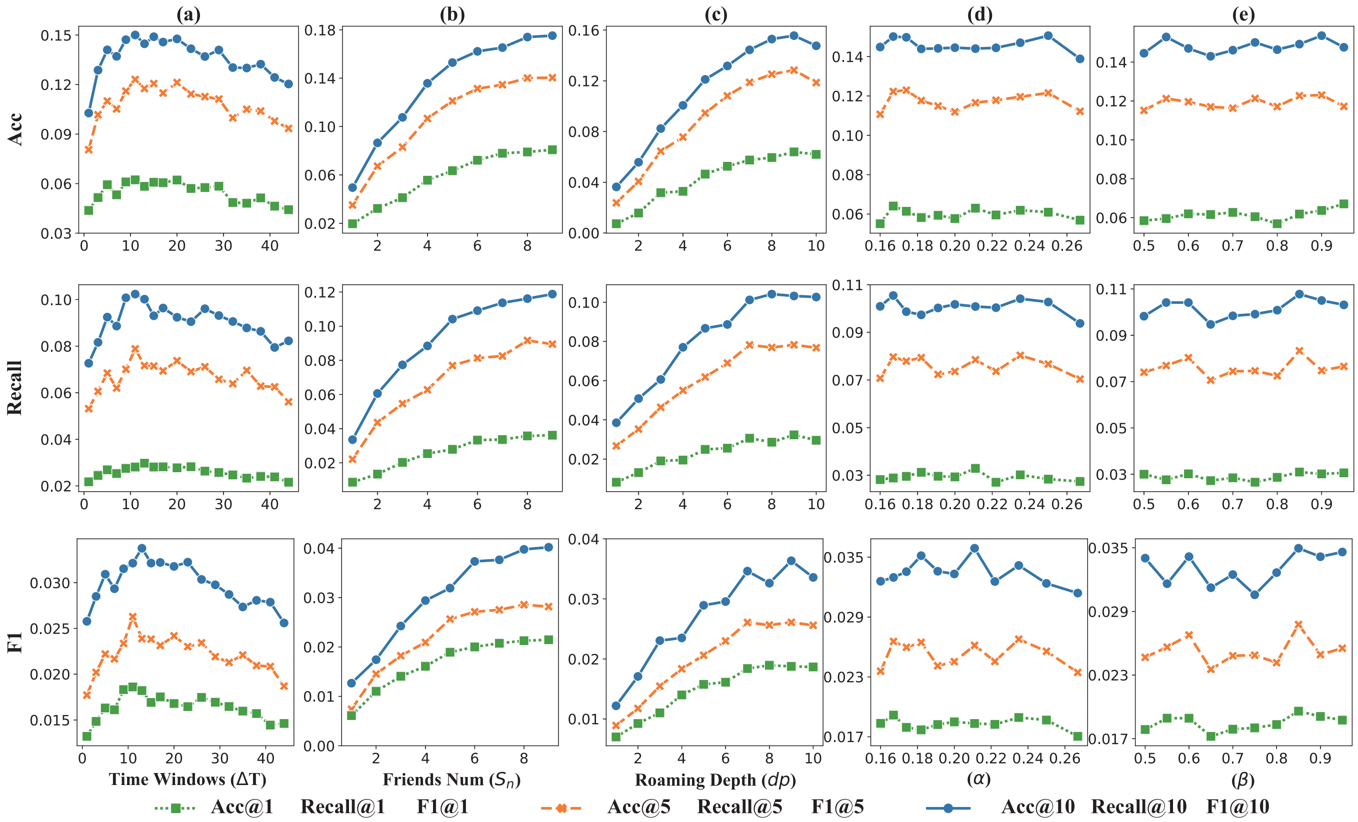


Fig. 8. Performance (Acc, Recall, and F1) comparison over the different walking parameters on the Yelp (Las Vegas) dataset. (a) shows the performance as the time window change, (b) shows the performance change of the number of friends, (c) shows the performance change of roaming depth on alter, and (d) and (e) show the influence of the hyperparameters in (5) on performance.

1) *DGRec [60]:* This is a model based on graph convolutional networks and social information for session-based recommendation in online communities.
2) *FM [61]:* This is a model combines the advantages of support vector machines (SVMs) and factorization models.
3) *NFM [62]:* This is a model that combines the linearity of factorization machines (FMs) in modeling second-order feature interactions and the non-linearity of neural network in modeling higher-order feature interactions.
4) *CKE [63]:* This is a model that integrates collaborative filtering and knowledge base.

The comparison results of these models and our method on the Yelp (Las Vegas) dataset are shown in Table IV. The results show that the prediction result of DGRec is better than DEEPMOVE (no social information), which implies that the social information do help models to improve the prediction performance. However, the performance of DGRec is less than other traditional latent social information involved models, i.e., FM, NFM, and CKE, also implies that although social information is helpful, it is challenging to apply social information on the model. Furthermore, when our method is applied to DEEPMOVE (ours), the prediction result is better than all other methods, proving that the way our research uses social information is more effective than other models. It also should be noted that these comparison methods are deeply integrated with social information, that is, it is impossible to separate the social information utilization methods from the models, while our method can be applied to most existing models.

VI. Conclusion and Discussion

In conclusion, in order to compensate for the poor performance caused by the sparsity and heterogeneity of online user data, our work proposes a UBPN, which considering the impact of spatiotemporal proximity. Correspondingly, a new walking strategy is proposed to extract the hidden information from the alter contacts to supplement ego user's data. The walking strategy takes into account temporal and social impact, and can be directly integrated into the existing prediction models. Experimental results on Yelp (Las Vegas) and Gowalla dataset show that our method can perform better than the baseline methods. Moreover, our work also studies the sensitivity of the parameters and finds that there exists an optimal value in the time window (ΔT) and roaming depth (dp), and a saturation value in the number of alter contacts S_n. It should be noted that the present UBPN is constructed based on correlation, not causality. Thus we may replace correlation with causality in future work.

The present study highlights that the proposed walking strategy can effectively extract the alter contact's hidden information, thus being used as a substitute for missing information of the ego user and improving the performance of the prediction model. Our results also highlight that personal information is strongly embedded in the social network. Furthermore, our work may help the relevant departments predict the patient's movement trajectory during an epidemic and also may support the entertainment website (e.g., Yelp) managers in their ambition to promote the recommendation system.

ACKNOWLEDGMENT

The authors would like to thank all the members in the Web Research Group, Zhejiang University of Technology, Hangzhou, China, for the valuable discussion about the ideas and technical details presented in this article.

REFERENCES

[1] D. J. Watts and S. Strogatz, "Collective dynamics of 'small-world' networks," *Nature*, vol. 393, no. 6684, pp. 440–442, 1998.
[2] L. Backstrom, P. Boldi, M. Rosa, J. Ugander, and S. Vigna, "Four degrees of separation," in *Proc. 4th Annu. ACM Web Sci. Conf.*, 2012, pp. 33–42.
[3] Q. Hao *et al.*, "Equip tourists with knowledge mined from travelogues," in *Proc. 19th Int. Conf. World Wide Web (WWW)*, 2010, pp. 401–410.
[4] Y. Hu, Y. Koren, and C. Volinsky, "Collaborative filtering for implicit feedback datasets," in *Proc. 8th IEEE Int. Conf. Data Mining*, Dec. 2008, pp. 263–272.
[5] E. Cho, S. A. Myers, and J. Leskovec, "Friendship and mobility: User movement in location-based social networks," in *Proc. 17th ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining (KDD)*, 2011, pp. 1082–1090.
[6] S. Scellato, A. Noulas, and C. Mascolo, "Exploiting place features in link prediction on location-based social networks," in *Proc. 17th ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining (KDD)*, 2011, pp. 1046–1054.
[7] C. Fu *et al.*, "Link weight prediction using supervised learning methods and its application to yelp layered network," *IEEE Trans. Knowl. Data Eng.*, vol. 30, no. 8, pp. 1507–1518, Aug. 2018.
[8] H. Gao, J. Tang, and H. Liu, "Exploring social-historical ties on location-based social networks," in *Proc. 6th Int. AAAI Conf. Weblogs Social Media*, 2012, pp. 114–121.
[9] A. Sadilek, H. Kautz, and J. P. Bigham, "Finding your friends and following them to where you are," in *Proc. 5th ACM Int. Conf. Web Search Data Mining (WSDM)*, 2012, pp. 723–732.
[10] V. W. Zheng, Y. Zheng, X. Xie, and Q. Yang, "Collaborative location and activity recommendations with GPS history data," in *Proc. 19th Int. Conf. World Wide Web (WWW)*, 2010, pp. 1029–1038.
[11] Y. Zheng, L. Zhang, X. Xie, and W.-Y. Ma, "Mining interesting locations and travel sequences from GPS trajectories," in *Proc. 18th Int. Conf. World Wide Web (WWW)*, 2009, pp. 791–800.
[12] Y. Leng, L. Rudolph, J. Zhao, and H. N. Koutsopolous, "Synergistic data-driven travel demand management based on phone records," in *Proc. ACM SIGKDD Int. Workshop Urban Comput.*, 2017, pp. 1–9.
[13] X. Chen, S. Li, H. Li, S. Jiang, Y. Qi, and L. Song, "Generative adversarial user model for reinforcement learning based recommendation system," in *Proc. 36th Int. Conf. Mach. Learn.*, 2019, pp. 1052–1061.
[14] Y. Feng *et al.*, "Deep session interest network for click-through rate prediction," 2019, *arXiv:1905.06482*. [Online]. Available: http://arxiv.org/abs/1905.06482
[15] Q. Chen, H. Zhao, W. Li, P. Huang, and W. Ou, "Behavior sequence transformer for e-commerce recommendation in Alibaba," in *Proc. 1st Int. Workshop Deep Learn. Pract. High-Dimensional Sparse Data*, Aug. 2019, pp. 1–4.
[16] Y. Leng, R. Ruiz, X. Dong, and A. Pentland, "Interpretable recommender system with heterogeneous information: A geometric deep learning perspective," 2020. [Online]. Available: https://ssrn.com/abstract=3696092
[17] M. Batty *et al.*, "Smart cities of the future," *Eur. Phys. J. Special Topics*, vol. 214, no. 1, pp. 481–518, Nov. 2012.
[18] Y. Leng, X. Dong, and E. Moro, "The rippling effect of social influence via phone communication network," in *Complex Spreading Phenomena in Social Systems*. Cham, Switzerland: Springer, 2018, pp. 323–333.
[19] Q. Xuan *et al.*, "Modern food foraging patterns: Geography and cuisine choices of restaurant patrons on yelp," *IEEE Trans. Computat. Social Syst.*, vol. 5, no. 2, pp. 508–517, Jun. 2018.
[20] Y. Leng *et al.*, "Collective behavior over social networks with data-driven and machine learning models," Ph.D. dissertation, Massachusetts Inst. Technol., Cambridge, MA, USA, 2020.
[21] R. Pan and M. Scholz, "Mind the gaps: Weighting the unknown in large-scale one-class collaborative filtering," in *Proc. 15th ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining*, 2009, pp. 667–676.
[22] S. Gambs, M.-O. Killijian, and M. N. del Prado Cortez, "Next place prediction using mobility Markov chains," in *Proc. 1st Workshop Meas., Privacy, Mobility (MPM)*, 2012, pp. 1–6.
[23] W. Mathew, R. Raposo, and B. Martins, "Predicting future locations with hidden Markov models," in *Proc. ACM Conf. Ubiquitous Comput. (UbiComp)*, 2012, pp. 911–918.
[24] S. Rendle, C. Freudenthaler, and L. Schmidt-Thieme, "Factorizing personalized Markov chains for next-basket recommendation," in *Proc. 19th Int. Conf. World Wide Web (WWW)*, 2010, pp. 811–820.
[25] C. Cheng, H. Yang, M. R. Lyu, and I. King, "Where you like to go next: Successive point-of-interest recommendation," in *Proc. 23rd Int. Joint Conf. Artif. Intell.*, 2013, pp. 2605–2611.
[26] S. Feng, X. Li, Y. Zeng, G. Cong, Y. M. Chee, and Q. Yuan, "Personalized ranking metric embedding for next new POI recommendation," in *Proc. 24th Int. Joint Conf. Artif. Intell.*, 2015, pp. 1–7.
[27] M. Xie, H. Yin, H. Wang, F. Xu, W. Chen, and S. Wang, "Learning graph-based POI embedding for location-based recommendation," in *Proc. 25th ACM Int. Conf. Inf. Knowl. Manage.*, Oct. 2016, pp. 15–24.
[28] T. Mikolov, S. Kombrink, L. Burget, J. Černocký, and S. Khudanpur, "Extensions of recurrent neural network language model," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP)*, May 2011, pp. 5528–5531.
[29] Y. Zhang *et al.*, "Sequential click prediction for sponsored search with recurrent neural networks," in *Proc. 28th AAAI Conf. Artif. Intell.*, 2014, pp. 1369–1375.
[30] Q. Liu, S. Wu, L. Wang, and T. Tan, "Predicting the next location: A recurrent model with spatial and temporal contexts," in *Proc. 13th AAAI Conf. Artif. Intell.*, 2016, pp. 1–7.
[31] R. Li, Y. Shen, and Y. Zhu, "Next point-of-interest recommendation with temporal and multi-level context attention," in *Proc. IEEE Int. Conf. Data Mining (ICDM)*, Nov. 2018, pp. 1110–1115.
[32] J. Manotumruksa, C. Macdonald, and I. Ounis, "A contextual attention recurrent architecture for context-aware venue recommendation," in *Proc. 41st Int. ACM SIGIR Conf. Res. Develop. Inf. Retr.*, Jun. 2018, pp. 555–564.
[33] J. Feng *et al.*, "DeepMove: Predicting human mobility with attentional recurrent networks," in *Proc. World Wide Web Conf. (WWW)*, 2018, pp. 1459–1468.
[34] P. Zhao *et al.*, "Where to go next: A spatio-temporal gated network for next POI recommendation," in *Proc. 33rd AAAI Conf. Artif. Intell.*, vol. 33, 2019, pp. 5877–5884.
[35] S. Scellato, C. Mascolo, M. Musolesi, and V. Latora, "Distance matters: Geo-social metrics for online social networks," in *Proc. 3rd Wonference Online Social Netw.*, 2010, pp. 1–9.
[36] J. Goldenberg and M. Levy, "Distance is not dead: Social interaction and geographical distance in the internet era," 2009, *arXiv:0906.3202*. [Online]. Available: http://arxiv.org/abs/0906.3202
[37] D. Mok, B. Wellman, and J. Carrasco, "Does distance matter in the age of the internet?" *Urban Stud.*, vol. 47, no. 13, pp. 2747–2783, 2010.
[38] S. Aral and C. Nicolaides, "Exercise contagion in a global social network," *Nature Commun.*, vol. 8, no. 1, pp. 1–8, Apr. 2017.
[39] C. Song, Z. Qu, N. Blumm, and A.-L. Barabási, "Limits of predictability in human mobility," *Science*, vol. 327, no. 5968, pp. 1018–1021, 2010.
[40] D. Yang, B. Qu, J. Yang, and P. Cudre-Mauroux, "Revisiting user mobility and social relationships in LBSNs: A hypergraph embedding approach," in *Proc. World Wide Web Conf. (WWW)*, 2019, pp. 2147–2157.
[41] D. Lian, X. Xie, V. W. Zheng, N. J. Yuan, F. Zhang, and E. Chen, "CEPR: A collaborative exploration and periodically returning model for location prediction," *ACM Trans. Intell. Syst. Technol.*, vol. 6, no. 1, pp. 1–27, Mar. 2015.
[42] L. Luceri, T. Braun, and S. Giordano, "Analyzing and inferring human real-life behavior through online social networks with social influence deep learning," *Appl. Netw. Sci.*, vol. 4, no. 1, p. 34, Dec. 2019.
[43] M. Backes, M. Humbert, J. Pang, and Y. Zhang, "Walk2friends: Inferring social links from mobility profiles," in *Proc. 24th ACM SIGSAC Conf. Comput. Commun. Secur.*, 2017, pp. 1943–1957.
[44] J. P. Bagrow and L. Mitchell, "The quoter model: A paradigmatic model of the social flow of written information," *Chaos, Interdiscipl. J. Nonlinear Sci.*, vol. 28, no. 7, Jul. 2018, Art. no. 075304.
[45] C. Fu, Y. Zheng, Y. Liu, Q. Xuan, and G. Chen, "NES-TL: Network embedding similarity-based transfer learning," *IEEE Trans. Netw. Sci. Eng.*, vol. 7, no. 3, pp. 1607–1618, Jul. 2020.

[46] C. Fu, Y. Zheng, S. Li, Q. Xuan, and Z. Ruan, "Predicting the popularity of tags in StackExchange QA communities," in *Proc. Int. Workshop Complex Syst. Netw. (IWCSN)*, Dec. 2017, pp. 90–95.

[47] M. Grbovic and H. Cheng, "Real-time personalization using embeddings for search ranking at airbnb," in *Proc. 24th ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining*, Jul. 2018, pp. 311–320.

[48] I. Kontoyiannis, P. H. Algoet, YU. M. Suhov, and A. J. Wyner, "Nonparametric entropy estimation for stationary processes and random fields, with applications to English text," *IEEE Trans. Inf. Theory*, vol. 44, no. 3, pp. 1319–1327, May 1998.

[49] J. P. Bagrow, X. Liu, and L. Mitchell, "Information flow reveals prediction limits in online social activity," *Nature Hum. Behav.*, vol. 3, no. 2, pp. 122–128, Feb. 2019.

[50] X. Wang, W. Zhu, and C. Liu, "Social recommendation with optimal limited attention," in *Proc. 25th ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining*, Jul. 2019, pp. 1518–1527.

[51] M. Ester *et al.*, "A density-based algorithm for discovering clusters in large spatial databases with noise," in *Proc. 25th ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining*, 1996, vol. 96, no. 34, pp. 226–231.

[52] A. Noulas, S. Scellato, R. Lambiotte, M. Pontil, and C. Mascolo, "A tale of many cities: Universal patterns in human urban mobility," *PLoS ONE*, vol. 7, no. 5, May 2012, Art. no. e37027.

[53] M. N. Gaonkar and K. Sawant, "AutoEpsDBSCAN: DBSCAN with Eps automatic for large dataset," *Int. J. Adv. Comput. Theory Eng.*, vol. 2, no. 2, pp. 11–16, 2013.

[54] D. E. Trahan and G. J. Larrabee, "Effect of normal aging on rate of forgetting," *Neuropsychology*, vol. 6, no. 2, p. 115, 1992.

[55] K. Cho *et al.*, "Learning phrase representations using RNN encoder-decoder for statistical machine translation," 2014, *arXiv:1406.1078*. [Online]. Available: http://arxiv.org/abs/1406.1078

[56] S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural Comput.*, vol. 9, no. 8, pp. 1735–1780, 1997.

[57] X. Chen *et al.*, "Sequential recommendation with user memory networks," in *Proc. 11th ACM Int. Conf. Web Search Data Mining*, Feb. 2018, pp. 108–116.

[58] G. Karypis, "Evaluation of item-based top-N recommendation algorithms," in *Proc. 10th Int. Conf. Inf. Knowl. Manage. (CIKM)*, 2001, pp. 247–254.

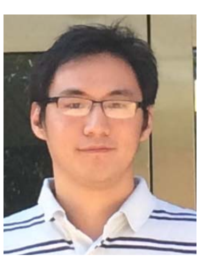
[59] H. Wang, N. Wang, and D.-Y. Yeung, "Collaborative deep learning for recommender systems," in *Proc. 21th ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining*, Aug. 2015, pp. 1235–1244.

[60] W. Song, Z. Xiao, Y. Wang, L. Charlin, M. Zhang, and J. Tang, "Session-based social recommendation via dynamic graph attention networks," in *Proc. 12th ACM Int. Conf. Web Search Data Mining*, Jan. 2019, pp. 555–563.

[61] S. Rendle, Z. Gantner, C. Freudenthaler, and L. Schmidt-Thieme, "Fast context-aware recommendations with factorization machines," in *Proc. 34th Int. ACM SIGIR Conf. Res. Develop. Inf. (SIGIR)*, 2011, pp. 635–644.

[62] X. He and T.-S. Chua, "Neural factorization machines for sparse predictive analytics," in *Proc. 40th Int. ACM SIGIR Conf. Res. Develop. Inf. Retr.*, Aug. 2017, pp. 355–364.

[63] F. Zhang, N. J. Yuan, D. Lian, X. Xie, and W.-Y. Ma, "Collaborative knowledge base embedding for recommender systems," in *Proc. 22nd ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining*, Aug. 2016, pp. 353–362.

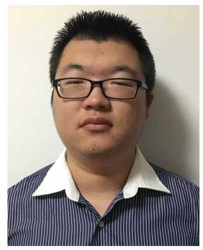


Chenbo Fu received the B.S. degree in physics from Zhejiang University of Technology, Hangzhou, China, in 2007, and the M.S. and Ph.D. degrees in physics from Zhejiang University, Hangzhou, in 2009 and 2013, respectively.

He was a Post-Doctoral Researcher with the College of Information Engineering, Zhejiang University of Technology. He was a Visiting Scholar with the Department of Computer Science, University of California at Davis, Davis, CA, USA, in 2014. In 2017, he was a Research Assistant with the Department of Electronic Engineering, City University of Hong Kong, Hong Kong, SAR, China. He is currently a Lecturer with the Institute of Cyberspace Security, Zhejiang University of Technology. His research interests include network-based algorithms design, social networks, data mining, chaos synchronization, network dynamics, and machine learning.

Yinan Xia received the B.S. degree from Shaoxing University, Shaoxing, China, in 2018. He is currently pursuing the M.S. degree with the College of Information Engineering, Zhejiang University of Technology, Hangzhou, China.

His current research interests include social network analysis and data mining.



Xinchen Yue received the B.S. degree in automation from Shaoxing University, Shaoxing, China, in 2018. He is currently pursuing the M.S. degree in control theory and engineering with the College of Information and Engineering, Zhejiang University of Technology, Hangzhou, China.

His current research interests include science of science and causal inference, especially the intersection of the two fields.

Shanqing Yu received the M.S. degree from the Graduate School of Information, Production and Systems, Waseda University, Tokyo, Japan, in 2008, and the Ph.D. degree from the School of Computer Engineering, Waseda University, in 2011.

She is currently a Lecturer with the Institute of Cyberspace Security, Zhejiang University of Technology, Hangzhou, China, and the College of Information Engineering, Zhejiang University of Technology. Her research interests include intelligent computation and data mining.

Yong Min received the B.S. degree in material engineering and the Ph.D. degree in computer science and technology from Zhejiang University, Hangzhou, China, in 2003 and 2010, respectively.

He was a Post-Doctoral Researcher with the College of Life Sciences, Zhejiang University, in 2011. He was a Visiting Scholar with the Department of Environmental Science, University of Quebec at Montreal, Montreal, QC, Canada, in 2017. He is currently an Associate Professor with the Institute of Cyberspace Security, Zhejiang University of Technology, Hangzhou. His research interests include network-based algorithms design, social network analysis, data mining, computational communication, network dynamics, and machine learning.

Qingpeng Zhang (Senior Member, IEEE) received the B.S. degree in automation from Huazhong University of Science and Technology, Wuhan, China, in 2009, and the Ph.D. degree in systems and industrial engineering from The University of Arizona, Tucson, AZ, USA, in 2012.

He was a Post-Doctoral Research Associate with the Tetherless World Constellation, Department of Computer Science, Rensselaer Polytechnic Institute, Troy, NY, USA. He is currently an Associate Professor with the School of Data Science, City University of Hong Kong, Hong Kong. His research interests include social computing, complex networks, data mining, and semantic web.

Dr. Zhang is also an Associate Editor of the IEEE TRANSACTIONS ON INTELLIGENT TRANSPORTATION SYSTEMS, the IEEE TRANSACTIONS ON COMPUTATIONAL SOCIAL SCIENCE, *Journal of Alzheimer's Disease*, and *PLOS One*.

Yan Leng received the M.S. degree from the Computer Science and Civil Engineering Department, Massachusetts Institute of Technology (MIT), Cambridge, MA, USA, in 2016, and the Ph.D. degree from the Media Lab, MIT, in 2020.

She is currently an Assistant Professor with the McCombs School of Business, The University of Texas at Austin, Austin, TX, USA. She is also a Network Scientist working on social science problems. Her research interests include the intersection of machine learning, network theory, and causal inference. She uses large-scale behavioral data to understand collective human behavior over social networks and builds computational techniques for solving societal and organizational issues.