Exploiting Twitter for Next-Place Prediction

Carmela Comito
National Research Council of Italy (CNR)
Institute for High Performance Computing and Networking (ICAR)
Via Pietro Bucci, 7-11C, 87036 Rende (CS), Italy
Email:comito@icar.cnr.it

***Abstract*—The time- and geo-coordinates associated with a sequence of tweets manifest the spatial-temporal movements of people in real life. This paper aims to analyze such movements to predict the next location of an individual based on the observations of his mobility behavior over some period of time and the recent locations that he has visited. To this end, we defined a prediction methodology based on a set of spatio-temporal features characterizing locations and movements among them. We then combined the features in a supervised learning approach based on M5 model trees. The experimental results obtained by using a real-world dataset show that the supervised method is effective in predicting the users next places achieving a remarkable accuracy.**

***Keywords*-Twitter, next-place prediction, spatio-temporal patterns.**

I. Introduction

Social media become very popular in recent years and is receiving an always increasing attention from the research community as through the user-generated data it embeds precious information concerning human dynamics and behaviors within urban context. The ability to associate spatial context to social posts is a popular feature of the most used on line social networks. Facebook and Twitter exploit the GPS readings of users phones to tag user posts, photos and videos with geographical coordinates.

According to this view, people travelling and visiting a sequence of locations generate many trajectories in the form of geo-tagged tweets. The time- and geo-references associated with a sequence of tweets manifest the spatial-temporal movements of Twitter users. This paper aims to analyze such movements to predict the next location visited from a target location based on the observations of users' mobility behavior over some period of time and the recent locations that have been visited from there.

We address the next place prediction problem as a *ranking task*. The aim is to rank the set of locations so that the next location to be visited will be ranked at the highest possible position in the list. To rank locations we propose a methodology based on a set of spatio-temporal prediction features characterizing the locations and the interactions among them. Specifically, two models are proposed. One model exploits the individual features for the prediction. The other one combines the features in a supervised learning framework based on decision tree models to predict future locations. The experimental evaluation performed on a real-word dataset of tweets shows the effectiveness of the proposed approach, with accuracy values of up to 0.90 for the supervised model.

The rest of the paper is organized as follows. Section II overviews related works. Section III describes the reference dataset and the proposed data model. Section IV presents the proposed approach to next-place prediction, introducing the prediction features. The experimental evaluation performed on a real-word dataset of tweets collected in London city is reported in Section V. Finally, Section VI concludes the paper.

II. Related work

Researchers analyze human mobility patterns to improve location prediction services, and therefore exploit their potential power on various applications such as mobile marketing, traffic planning, and disaster relief [4].

Many studies tackle the problems of predicting the next location where a mobile user moves to. Personal-based prediction [15], [7], [16] and general-based prediction [12] are two approaches often adopted in this problem domain. The personal-based prediction approach considers movement behavior of each individual as independent and thus uses only the movements of an individual user to predict his/her next location. On the contrary, the general-based prediction makes a prediction based on the common movement behavior of general mobile users.

Researchers have investigated many explanatory variables for next-place prediction, including cell phone data usage visiting frequency and contextual information from smart phone sensors. Also, with the rapid growth of location-based social networks (LBSN), researchers have used check-in patterns to predict the next check-in. Current research on location prediction in LBSNs mainly focuses on two problems: 1) predicting a users location at any time, including predicting a users home location [2], [13], [1], [14]; 2) predicting the location of each tweet [8], [10].

A variety of algorithms for next place prediction have been proposed, but most focus on classification [13], Markov-based models [3], [5], and extraction of both raw trajectories [9] and semantic trajectories [17], [18].

[11] presents an algorithm for predicting the home location of Twitter users. It builds a set of different classifiers, such as statistical classifiers using words, hashtags or place names of tweets and heuristics classifiers using the frequency of place names or Foursquare check-ins, and then creates an

978-1-908320/86/5/$31.00 © 2017 IEEE

ensemble of the classifiers to improve the prediction accuracy. In contrast, the goal of our work is quite different and as we aim to predict the next location and not just home locations considering more about a users moving trajectories and introducing a set of different spatio-temporal features about the locations and without using the text in the tweets.

Chang et al. [1] utilized logistic regression model to combine a set of features extracted from Facebook data. The features include a users previous check-ins, users friends check-ins, demographic data, distance of place to users usual location, etc. Their results demonstrated that the number of previous check-ins by the user is a strong predictor, and also previous check-ins made by friends and the age of the user are good features for prediction.

In [7], Jeung et al. propose an approach which predicts future locations of a user by combining predefined motion functions, i.e., linear or non-linear models that capture object movements as sophisticated mathematical formulas, with the movement patterns of the user, extracted by a modified version of the Apriori algorithm. In [15], Yavas et al. mine the movement patterns of an individual user to form association rules and use these rules to make location prediction. Additionally, they consider the support and confidence in selecting the association rules for making predictions. In [16], Ye et al. propose a novel pattern, called Individual Life Pattern, which is mined form individual trajectory data, and they uses such pattern to describe and model the mobile users periodic behaviors.

Ying et al.[17] integrate semantic information about the places visited by an individual in addition to its location data to enhance the accuracy of the prediction about his future location. The approach relies on the notion of semantic trajectories, which represents the mobility of an individual as a sequence of visited places tagged with semantic information. To predict the next location based on semantic trajectories, the authors have developed a framework called SemanPredict. The online prediction module is responsible for matching the current trajectory of a user with the closest trajectory in the database by relying on the geographical and semantic features.

The work more similar to our approach is the one of Noulas et al. [13]. This work predicts the next place in a city that the user will visit, proposing a set of features that exploit information on transitions between types of places, mobility flows between venues, and spatio-temporal characteristics of user check-in patterns. They combining all features in two supervised learning models, based on linear regression and M5 model trees, resulting in a higher overall prediction accuracy.

Summarizing, the proposed algorithm differs from the approaches of this category as we aim to infer, given a location, the next location that could be visited from there, whereas in the above approaches the focus is on the user, that is given a user what will be his next location.

III. Twitter dataset and data model

The geo-located data mined in this work is a dataset of tweets tagged with GPS location within the boundaries of the city of London, one of the top three cities by number of tweets[1]. Numerically speaking, we consider a Twitter dataset of 7,424,112 tweets issued by 292,195 mobile users in 6,098,148 distinct locations, during a period of six month started in June 2013 and ended in November 2013. We built a multi-threaded crawler to access the Twitter Streaming API. The crawler collects the tweets filtered by location and processes the results to obtain a dataset in which each entry is a tweet that includes the ID of the user who created the tweet, the timestamp and the GPS coordinates of the tweet. The dataset represents a sequence of daily snapshots, with an average number of tweets per day greater than 40,000.

***Definition** 1:* **Geo-tagged tweet**. A geo-tagged tweet $\mathrm{tw} \in \mathrm{TW}$ is characterized by the user u who tweeted, a location l from where tw has been posted and a timestamp, t, that is the time at which it has been posted. The location is identified by a pair of geographic coordinates $\mathrm{l} = (\mathrm{x}, \mathrm{y})$, latitude and longitude, respectively. Accordingly, a geo-tagged tweet can be defined as a triple $\mathrm{tw} = (\mathrm{u}, \mathrm{l}, \mathrm{t})$.

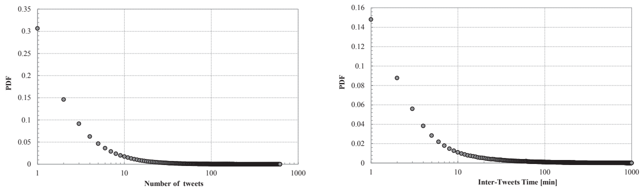
The data analysis reveals that the behaviour of the users is very heterogeneous: note the long tail of the probability distribution functions (PDF) both of the number of tweets and of the time interval that elapses between successive users tweets. Figure 1 (a) shows the PDF of the number of tweets per user in a month. Even if the volume of tweets per month is very high, most of the users, the 78% post less than 10 tweets per month. This may depend on the fact that many users are tourists and then occasionally visit the city. 21% of users are more active making more than 10 tweets but less than 100, finally very few users, just 1%, post more than 100 tweets per month.

A similar pattern arises when considering the time elapsed between successive tweets. Figure 1 (b) shows that about 40% of tweets are posted with high frequency (i.e. with an inter time of 10 minutes). However the other 60% of inter tweet time intervals have a length that varies in a very large range of values . On te other hand, only 28% of tweets are posted with a frequency greater than 3 hours.

We also observe the tweets frequency during the course of the week. Figure 2 (a) shows that the tweets rate for each day of the week has a periodic behaviour. Days exhibit a peak in the evening and a dip at night time. The figure also highlights some differences between week days and the weekend. In particular, in weekends the volume of tweets is higher, mainly during the morning and there is a peak at lunchtime. This is more evident on Saturdays.

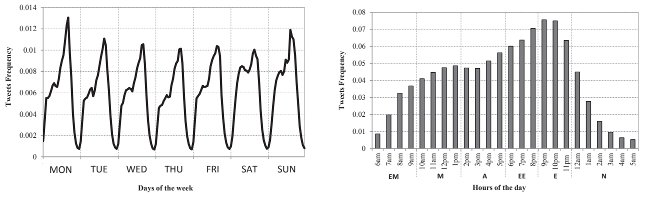
These patterns seem to mirror user behaviour: for instance, during a week day, a user might spend morning and afternoon at the workplace, taking a lunch break in a restaurant, while in the evening he might go to the gym, to the cinema or stay at home. In order to exploit these temporal patterns for our classification task we divide the day into six different time slots, as shown in Figure 2 (b). The time slots, are formally specified as follows:

[1] http://semiocast.com/



(a) (b)

Figure 1. Probability Distribution Function of number of monthly tweets per user (a) and of the time elapsed between consecutive tweets (b).



(a) (b)

Figure 2. Tweets frequency during a week (a) and Temporal evolution of tweets frequency during the daily time slots(b).

Definition 2: TS is a finite set of timeslots with $|\mathrm{TS}| = 6$. Each $\mathrm{ts} \in \mathrm{TS}$ is a time-object of varying time duration belonging to a day. $\mathrm{TS} = \{\mathrm{N}, \mathrm{EM}, \mathrm{M}, \mathrm{A}, \mathrm{EE}, \mathrm{E}\}$ where:
$\mathrm{N} = \mathrm{Night}[12:00\mathrm{am} - 05:59\mathrm{am}]$;
$\mathrm{EM} = \mathrm{EarlyMorning}[06:00\mathrm{am} - 09:59\mathrm{am}]$;
$\mathrm{M} = \mathrm{Morning}[10:00\mathrm{am} - 01:59\mathrm{pm}]$;
$\mathrm{A} = \mathrm{Afternoon}[02:00 - 05:59\mathrm{pm}]$;
$\mathrm{EE} = \mathrm{EarlyEvening}[06:00\mathrm{pm} - 08:59\mathrm{pm}]$;
$\mathrm{E} = \mathrm{Evening}[09:00\mathrm{pm} - 11:59\mathrm{pm}]$.
On the basis of this definition, we specify a mapping function $\mathrm{TS}(\mathrm{tw})$ that associates the corresponding time slot to the timestamp of a tweet.

IV. Ranking locations to place prediction

In this section we first formulate the next place prediction problem and then introduce the prediction features.

A. Problem Formulation

We formalize the problem of predicting the next location visited by users who move from a given place, based on the observations of historic visits. Our goal is to identify the most likely location that will be visited considering thousands of candidate venues.

In this work we focus on the check-ins posted by twitter users in the city of London. In particular, we refer only to the geo-spatial references of such posts without considering any other information contained in the tweets like the text. Accordingly, our data model is defined as follows.

Given a set of geo-tagged tweets $\mathcal{TW}$, we extract the set of locations L visited from Twitter users U by only exploiting the geographical coordinates. After having extracted the locations, our aim is to predict the next location that users will visit from a specific location by exploiting the spatial-temporal patterns exhibited by users in that location.

Given the set of tweets $\mathcal{TW}$ currently posted in a city, for each location l in the city, we aim to infer the next location visited from there among a finite set of locations L.

We constraints the candidate venues to the set of locations L in the city from where a significant number of tweets have been posted.

We formulate the next-location prediction problem as a *ranking task*. Our aim is to rank the set of locations so that the next location to be visited will be ranked at the highest possible position in the list. Specifically, given a location $\mathrm{l} \in \mathrm{L}$, we aim to predict the most likely location $\mathrm{k} \in \mathrm{L}$ that will be visited next. We address this problem as a *ranking problem* where k is chosen from a set of candidate places ranked based on the prediction features. Each feature computes a score for each candidate venue k, which is used to obtain a ranked list of locations R. Each ranked list is sorted in decreasing order and the position of location k in R is denoted with $\mathrm{rank}(\mathrm{k})$. Our purpose is that the future venue that will be visited is highly ranked by the prediction algorithms. We also compute the ground truth ranked list $\dot{\mathrm{R}}$ on the basis of the historical movements between the locations under predictions.

B. Prediction features

The proposed features characterize locations in terms of spatial-temporal patterns, and also in terms of mobility interactions among them.

The features identified are listed below.

Number of tweets. This feature represents the number of tweets posted by users in location k.

$$\mathrm{TW_k} = |\mathrm{tw} = (\mathrm{u}, \mathrm{k}, \mathrm{t}) \in \mathrm{TW}| \tag{1}$$

Number of users. Knowing the number of people who visit a place is indicative of its popularity. The number of visitors of a location k, is the number of distinct users who have posted at least one tweet while at the location.

$$\mathrm{U_k} = |\mathrm{u} \in \mathrm{U} : \mathrm{TW_{k,u}} \neq \emptyset| \tag{2}$$

where $\mathrm{TW_{k,u}}$ is the set of tweets post in k by u.

Night Location. This feature is proposed to separate the locations visited mostly at night time from those visited during the day. This is formally achieved by measuring the ratio of night tweets versus the total number of tweets.

$$\mathrm{night_k} = \frac{|\mathrm{tw} = (\mathrm{u},\mathrm{k},\mathrm{t}) \in \mathrm{TW} : \mathrm{t} \in \mathrm{EE} \vee \mathrm{t} \in \mathrm{E} \vee \mathrm{t} \in \mathrm{N}|}{|\mathrm{TW_k}|} \tag{3}$$

Weekend Location. Likewise, we aim to partition the locations visited mostly during the weekend from those visited during weekdays. This is formally achieved by measuring the ratio of weekend tweets versus the total number of tweets.

$$\mathrm{weekend_k} = \frac{|\mathrm{tw} = (\mathrm{u},\mathrm{k},\mathrm{t}) \in \mathrm{TW} : \mathrm{t} \in \mathrm{Sat} \vee \mathrm{t} \in \mathrm{Sun}|}{\mathrm{TW_k}} \tag{4}$$

Tweets Entropy. This feature tells whether users tend to tweet regularly in a location k, describing the distribution of its tweets across the users. For this purpose we use the Shannon Entropy:

$$\begin{gathered} h(x_u) = -\sum_{u=1}^{n} p(x_u) log\ p(x_u), \\ where\ \ p(x_u) = f(k,u) = \frac{|TW_{k,u}|}{|TW_k|} \end{gathered} \tag{5}$$

The features listed in the following exploit past user movements between each location under prediction k and the target venue l. Exploiting this class of features we can answer the following questions: (1) how many movements are there from l to k (2) how far is l from k? (3) how long does it take from l to k?

To model interactions among couple of locations, we introduced the following definition of path:

***Definition** 3:* A path $\mathrm{p_{l,k}}$ is a movement that starts in location l and ends in location k as expressed by the sequence of tweets posted from such locations:

$$\mathrm{p_{l,k}} = \mathrm{tw_{l_1,t_1}} \longrightarrow \mathrm{tw_{l_2,t_2}} : \mathrm{l_1} = \mathrm{l} \wedge \mathrm{l_2} = \mathrm{k} \wedge \mathrm{t_1} < \mathrm{t_2}$$

Number of travels. The feature measures the number of travels along a path between two locations. It is formally defined as:

$$\mathrm{TR_{l,k}} = |\mathrm{p_{l,k}}| \tag{6}$$

Space. The feature represents the geographic distance between l and k, computed by the Haversine formula.

$$\mathrm{space_{l,k}} = \mathrm{Hdist(l,k)} \tag{7}$$

Time. The feature represents the average time to move directly from l to k.

$$\mathrm{time_{l,k}} = \frac{\sum_{\forall \mathrm{p_{l,k}}} \mathrm{t_{k,first}} - \mathrm{t_{l,last}}}{|\mathrm{p_{l,k}}|} \tag{8}$$

where $\mathrm{t_{l,last}}$ is the timestamp of the last tweet posted in l, and $\mathrm{t_{k,first}}$ is the timestamp of the first tweet posted in k.

V. Experimental Evaluation

In this section we present the experimental evaluation performed to assess the effectiveness and accuracy of the proposed prediction strategy. Specifically, we first evaluate the predictive power of each feature, then we combined the most performing features in a supervised classifier to asses if prediction accuracy can be improved.

A. Metrics

We use two metrics for the evaluation of the ranking functions, the Percentile Rank (PR), and the prediction accuracy (Accuracy).

The Percentile Rank (PR) [6] of a location l is defined as follows:

$$\mathrm{PR(l)} = \frac{|\mathrm{L}| - \mathrm{rank(l)} + 1}{|\mathrm{L}|} \tag{9}$$

where $rank(l)$ is the ranking obtained for location l by exploiting one of the prediction features.

The PR score is equal to 1 when the location that will be visited next is ranked first and it linearly decreases to 0 as the correct location is positioned at the bottom of the list. We evaluate a scenario where we generate for each location an ordered list of candidate locations, sorted from the one predicted to be most likely next visited till the least visited one. Specifically, we generate a ranking list for each of the considered features.

The Average Percentile Rank (APR) is obtained by averaging across all user check-in predictions.

The **Accuracy@X** allows to compare the different ranking strategies in terms of their prediction accuracy when using different prediction list sizes X. In this case, we successfully predict the next location if we rank a location in the top-X places. The average accuracy (Accuracy@X) measures the fraction of times that the future next location in the predicted list R is at the top-X of the the ground truth ranked list $\dot{\mathrm{R}}$.

B. Individual Feature Performance

We studied the predictive power of each feature: we first compute predictive scores for every pair of locations in the test set and then we rank these candidate locations according to their score.

Figure 3 shows the prediction accuracy achieved by each feature with the prediction list size. As can be noted from the graph, the *number of travels* feature outperforms all the others achieving an accuracy of about 0.66 for a list size of 40. The features accounting location popularity as *number of tweets* and the *number of users* achieve a good accuracy reaching 0.56 and 0.58, respectively. The *entropy*, *space* and *time* features are performing very similarly reaching a maximun value of about 0.30. Finally, the *weekend* and *night* features achieve a very low accuracy resulting, thus, to not be effective in terms of prediction.

Figure 4 shows the APR results for all the prediction features. For all the features the score is significantly better than the one achieved in terms of accuracy. Anyhow, for most of the features such scores are slightly higher than the Random Baseline (which would achieve 0.50), ranging from 0.50 to 0.65. They are an exception the features *number of travels*, *number of users*, and *number of tweets*, with *number of travels* achieving the highest score of about 0.80.

C. Supervised Learning Approach

In order to detect if a combination of factors improve predictions, we combine the individual features in a supervised

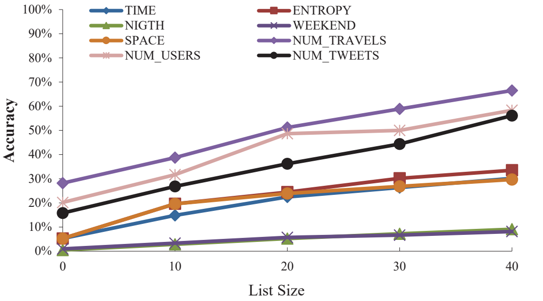


Figure 3. Prediction Accuracy of the different features with list size.

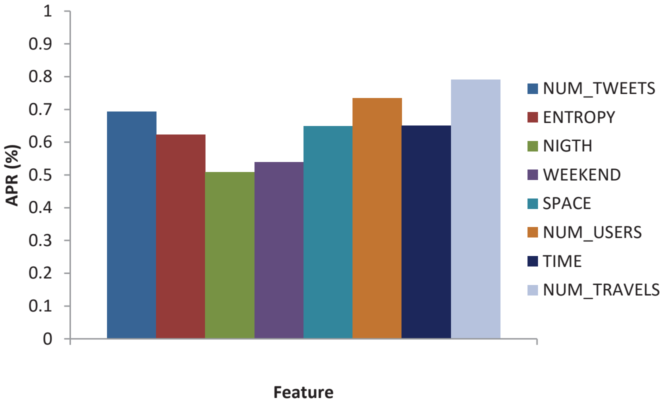


Figure 4. APR of the prediction features with list size

learning framework. We aim to obtain at least the same prediction accuracy of the best performing individual features. We use two algorithms, the M5 decision trees [14] and Linear Regression. We employ the implementations available in the popular machine learning framework WEKA.

To predict the next location we build supervised models for each target location on the basis of historic trajectories. More specifically, for each candidate location we build a training example which encodes the values of the features. This feature vector corresponds to a positive label if there is at least a correspondence recorded in the historic data between that candidate location and the target one, otherwise the label is negative. In this way, we train the model to distinguish between places that could be visited from a location, and places that will not be visited.

The supervised learning algorithms compute a regression scores, on the basis of which we rank the candidate venues, reducing the regression problem to a ranking one.

We compared the supervised approach against the individual prediction features. For this comparison we used the most performing prediction features as resulting from the above experimental evaluation. Figure 5 shows that supervised approaches outperform single features predictions. Specifically, M5P exhibits the best performance, achieving an accuracy that is largely higher compared to all the feature prediction approaches reaching a maximum of about 0.90. From the graph one can also note the linear regression is outperformed by all the other approaches till a list size of 5, while for larger list size it achieves a good accuracy that reaches a considerable peak of 0.83 when the list size is 40. According to the previous results, *number of travels* is the most performing among individual features.

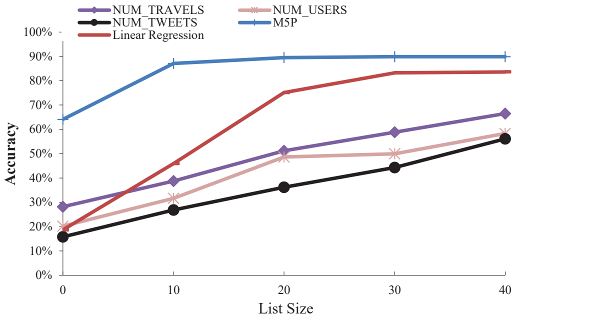


Figure 5. Supervised approach vs single-feature - Prediction Accuracy with list size

Figure 6 shows that also in terms of APR scores, M5P outperforms the other models with an APR of 0.90 versus the 0.79 achieved by the linear regression model which achieves the same score of the *number of travels* feature that again is the best performing among the individual prediction feature strategies.

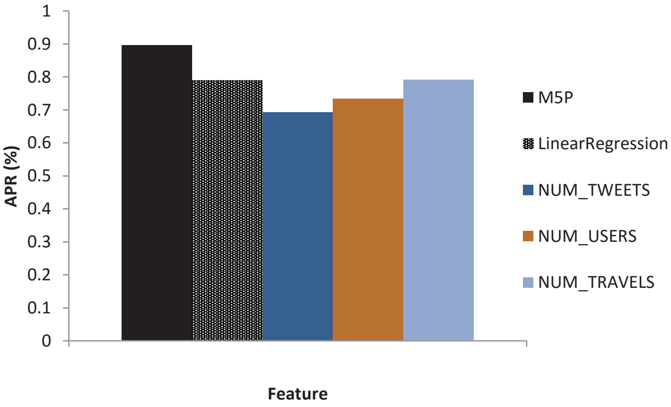


Fig. 6. Supervised approach vs single-feature - APR with list size.

VI. Conclusion

The paper proposed a methodology to predict next location by exploiting Twitter data. The prediction methodology is based on a set of spatio-temporal features characterizing locations and movements among them such as historical visits to locations, geographic distance between them, their popularity. The features have been also combined in a supervised learning approach based on M5 model trees. We have analysed a dataset of tweets collected in London city during January 2013. The experimental results show that the supervised method is effective in predicting the users next places achieving a remarkable accuracy.

REFERENCES

[1] J. Chang and E. Sun. Location3: How users share and respond to location-based data on social networking sites. In *Proceedings of ICWSM-11*, 2011.

[2] Z. Cheng, J. Caverlee, and K. Lee. You are where you tweet: A content-based approach to geo-locating twitter users. In *Proceedings of the 19th ACM International Conference on Information and Knowledge Management*, CIKM '10, pages 759–768. ACM, 2010.

[3] T. M. T. Do and D. Gatica-Perez. Contextual conditional models for smartphone-based human mobility prediction. In *Proceedings of the 2012 ACM Conference on Ubiquitous Computing*, pages 163–172. ACM, 2012.

[4] C. Ernst, A. Mladenow, and C. Strauss. Location-based crowdsourcing in disaster response. In *Proceedings of the 14th International Conference on Advances in Mobile Computing and Multi Media*, MoMM '16, pages 28–34. ACM, 2016.

[5] S. Gambs, M.-O. Killijian, and M. N. n. del Prado Cortez. Next place prediction using mobility markov chains. In *Proceedings of the First Workshop on Measurement, Privacy, and Mobility*, 2012.

[6] Y. Hu, Y. Koren, and C. Volinsky. Collaborative filtering for implicit feedback datasets. In *Proceedings of the 2008 Eighth IEEE International Conference on Data Mining*, ICDM '08, pages 263–272. IEEE Computer Society, 2008.

[7] H. Jeung, Q. Liu, H. T. Shen, and X. Zhou. A hybrid prediction model for moving objects. In *Proceedings of the 24th International Conference on Data Engineering, ICDE 2008, April 7-12, 2008, Cancún, México*, pages 70–79, 2008.

[8] M. Kawano and K. Ueda. Where are you talking from?: Estimating the location of tweets using recurrent neural networks. In *Proceedings of the Second International Conference on IoT in Urban Space*, Urb-IoT '16, pages 57–60. ACM, 2016.

[9] J. Krumm and E. Horvitz. Predestination: Inferring destinations from partial trajectories. In *Proceedings of the 8th International Conference on Ubiquitous Computing*, UbiComp'06, pages 243–260. Springer-Verlag, 2006.

[10] K. Lee, R. K. Ganti, M. Srivatsa, and L. Liu. When twitter meets foursquare: Tweet location prediction using foursquare. In *Proceedings of the 11th International Conference on Mobile and Ubiquitous Systems: Computing, Networking and Services*, MOBIQUITOUS '14, pages 198–207, 2014.

[11] J. Mahmud, J. Nichols, and C. Drews. Home location identification of twitter users. *ACM Trans. Intell. Syst. Technol.*, 5(3):47:1–47:21, 2014.

[12] M. Morzy. Prediction of moving object location based on frequent trajectories. In *Computer and Information Sciences - ISCIS 2006, 21th International Symposium*, pages 583–592, 2006.

[13] A. Noulas, S. Scellato, N. Lathia, and C. Mascolo. Mining user mobility features for next place prediction in location-based services. In *ICDM*, pages 1038–1043, 2012.

[14] H. Wang, M. Terrovitis, and N. Mamoulis. Location recommendation in location-based social networks using user check-in data. In *Proceedings of the 21st ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems*, SIGSPATIAL'13, pages 374–383. ACM, 2013.

[15] G. Yavas, D. Katsaros, O. Ulusoy, and Y. Manolopoulos. A data mining approach for location prediction in mobile environments. *Data Knowl. Eng.*, 54(2):121–146, 2005.

[16] Y. Ye, Y. Zheng, Y. Chen, J. Feng, and X. Xie. Mining individual life pattern based on location history. In *MDM*, pages 1–10, 2009.

[17] J. J.-C. Ying, W.-C. Lee, T.-C. Weng, and V. S. Tseng. Semantic trajectory mining for location prediction. In *Proceedings of the 19th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems*, pages 34–43. ACM, 2011.

[18] J. J.-C. Ying, E. H.-C. Lu, W.-C. Lee, T.-C. Weng, and V. S. Tseng. Mining user similarity from semantic trajectories. In *Proceedings of the 2Nd ACM SIGSPATIAL International Workshop on Location Based Social Networks*, pages 19–26. ACM, 2010.