Learning Sequential Mobility and User Preference for new Location Recommendation in Online Social Networks

Carmela Comito

National Research Council of Italy (CNR)
Institute for High Performance Computing and Networking (ICAR)
Via Pietro Bucci, 8-9C, 87036 Rende (CS), Italy
Email:carmela.comito@icar.cnr.it

***Abstract*—The fast expansion during the recent years of online social networks, such as Twitter, Facebook, or Foursquare, is making available an enormous and continuous stream of user-generated contents including information on human mobility within urban context. In particular, online social networks allows for the collection of geo-tagged data obtained through the GPS readings of phones through which users have the possibility to tag posts, photos and videos with geographical coordinates. In this context, recommending the future position of a mobile object is key for the implementations of several applications aiming at improving mobility within urban areas.**

The paper proposes a location recommendation approach that exploits geo-tagged data on social networks. The approach integrates user preference, sequential mobility and geographic constraints. The recommendation task is formulated as a similarity problem among the visiting and mobility profiles of users, accounting the mobility sequentiality in the patterns. Two ranking metrics are introduced to predict places the user could like. The metrics are then combined into an overall recommendation ranking function. The candidate locations are then ranked according to the two similarity measures. The experimental results obtained by using a real-world dataset of tweets show that the proposed method is effective in recommending unseen locations, outperforming representative state-of-the-art approaches.

***Keywords*-Location Recommendation, Online Social Networks, Sequential Mobility**

I. Introduction

The extensive use of location-based social networks (LBSNs) allows for the collection of huge amount of geo-tagged data about people activities and costumes within urban context, including human mobility regularities. According to this view, social network users traveling and visiting a set of locations can produce a huge amount of geo-location data that embed extensive knowledge about human dynamics and mobility behaviors within urban context.

The work presented in the paper aims to analyze the time and geo-referenced information associated with online posts to detect hot destinations and typical travel sequences and discover common behavior, i.e. patterns, rules and regularities in moving trajectories. In this context, recommending the future location to a mobile user is key for the implementations of several applications aiming at improving mobility within urban areas (e.g., traffic congestion, location-based advertisements, tourist recommendation paths). Accordingly, the paper focuses on detecting mobility information from LBSN and analyze such data to recommend new unseen locations that could be interesting for a target user based on the recent venues that she have been visited and on observations of users' mobility behavior over some period of time. The aim of location recommendations is to suggest a list of venues fitting user personal interests within a geographic area. In addition to its value for users, this information is valuable for third-party companies to advertise products, hotels, places, and to forecast service demand such as the number of taxis needed in a part of a city.

LBSN-based recommendations are not only based on preference and geography, but also on social relationships: unvisited venues that friends have checked in may be useful recommendations. Hence, recommender systems of new places using social data try to improve traditional recommender systems by considering two additional dimensions beyond the usual preference dimension: social and geographical dimensions. Accordingly, importance of POI recommendations has attracted a significant amount of research interest and a number of approaches have been proposed in literature. However, all these studies do not consider the influence of sequential patterns of check-in locations on users' check-in behaviors, called mobility sequential influence hereafter, although in reality human movement exhibits sequential patterns.

To overcome the limitation of current state-of-the-art approaches, this paper proposes an approach exploiting both sequential mobility and user preference in the recommendation task. The recommendation problem is formulated as a similarity problem among the visiting and mobility profiles of users. Actually, the recommendation function is a linear combination of user preference similarity and sequential mobility similarity. The candidate locations are then ranked according to the similarity measure. In particular, given a user u, a ranking score is computed for u across all

IEEE/ACM ASONAM 2020, December 7-10, 2020
978-1-7281-1056-1/20/$31.00

the unvisited locations. A recommendation list for user u is produced by selecting the top K locations with the highest ranks. The experimental evaluation performed on a real-world dataset of tweets shows the effectiveness of the proposed approach that outperforms some of the most commonly used baseline recommendation policies.

The rest of the paper is organized as follows. Section II overviews related works. Section III formulates the addressed problem and describes the data model. The methodology proposed to recommend new locations is introduced in Section IV. The experimental evaluation performed on a real-world dataset of tweets collected in London city is reported in Section V. Finally, Section VI concludes the paper.

II. Related work

A branch of recent research starts learning a user's interests from the user's location history and incorporates the social environment of the user to make recommendations. Specifically, [4], [7], [8], [9], [11] deposit people's location histories into a user-location matrix where a row corresponds to a users' location history and each column denotes a venue like a restaurant. Each entry in the matrix represents the number of visits of a particular user to a physical venue. Then, a user-based collaborative filtering (CF) model is employed to infer a user's interest to an unvisited venue. However, the similarity between two users is simply represented by the Cosine similarity between the two users' rows, overlooking the features of human mobility in geographic spaces, such as sequential and hierarchical properties of locations. To better estimate the similarity between users, Zheng et al. [11] proposed a hierarchical-graph-based similarity measurement taking the human mobility features into account. The location recommendation system using the user similarity outperforms those using the Cosine similarity.

Pham et al. [6] find user and item clusters in social networks and use such information to enhance CF methods. Random walk has also been exploited for recommendation. Yildirim and Krishnamoorthy [10] build a graph of items, in which each link is weighed by the similarity of its two owner items. Given a user u who has rated a set of items I, random walks on this item graph are adopted to find the items similar to I. In order to guarantee the connectivity of the item graph, two items are linked with a small weight even if the similarity is not computable. This approach makes the method unsuitable for large datasets. Konstas et al. study CF methods on a music social network to predict music playcounts [3]. They create a heterogeneous graph containing users, music tracks, and tags, and show that a random walk on the graph outperforms traditional CF methods. Their technique is tailored for music data and it is not applicable in our problem.

III. Problem formulation and data model

A recommendation problem is generally defined as the problem aiming at predicting values for unrated content, and using those predictions to rank items as recommendations.

The aim of *location recommendations* is to suggest a list of venues fitting user personal interests within a geographic area.

LBSNs present unprecedented large-scale check-in data to describe users' mobile behavior in spatial, temporal, and social aspects. Previous research exploited check-in preferences and social friendships on LBSN for location recommendation. Among existing work, the mobility paths typically travelled by users have not been explored for recommending locations.

Based on the above observation, the work proposed in this paper introduces a novel approach for location recommendation exploiting typical patterns usually travelled by users also considering spatio-temporal features about the movements among locations. Therefore, investigating the features embedded in daily patterns enables us to better understand human mobile behavior, providing a potential opportunity to design more advanced location recommender systems on LBSNs.

We express the problem in terms of a generic location-based social network (could be Facebook, Gowalla, Twitter) with geo-tagged posts to which we refer with the generic term check-ins, defined as follows:

Definition *1:* **Check-in**. A check-in c is defined as a triple $\mathrm{c} = (\mathrm{u}, \mathrm{l}, \mathrm{t})$ where u is the user that checked-in, l a location from where c has been posted, and t, is the time at which c has been published.

Accordingly, we represent a LBSN as a set of check-ins $\mathcal{C}$ posted by a set of users $\mathcal{U}$ from a set of locations $\mathcal{L}$ within a given geographic region.

In a day d a user u might visit one or more locations within a given geographic region. We refer to such movements as paths, defined as temporally ordered sequences of places visited by users. In the following we refer to the terms paths or travel routes interchangeably. For each user we compute his daily paths, formalized as follows:

Definition *2:* **Path**. A path is a spatio-temporal sequence of visited locations by a user u according to temporal order during the same day d.

$$\mathrm{p_u} = \mathrm{v_{u,l_0}} \longrightarrow \mathrm{v_{u,l_1}} \longrightarrow \ldots \longrightarrow \mathrm{v_{u,l_n}}$$

A visit $\mathrm{v_{u,l}}$ to a location l is characterized by: (i) a user u who visits the location l, and (ii) a sequence of check-ins $\{\mathrm{c_0}, \ldots, \mathrm{c_n}\}$ that u posts in l before moving to another place.

IV. The recommendation approach

Recommender systems are widely used and they have been studied in research quite extensively. The most popular approach in recommender systems is that of collaborative filtering, where recommendations are created based on whether a user has purchased a product in the past and on whether she liked it or not. Using the past behavior of a user, new recommendations are created based on the similarity of users or the similarity of products (items). While these algorithms can be adjusted to the problem of recommending new locations to users, by taking into account previous user check-ins, significant information like the distance of the proposed location to the user neighborhood or the social interaction between the user and those users that have visited this location are ignored.

The rich data about past user behavior that is traced by the LBSN differentiates the problem significantly from its traditional settings. The spatial and temporal nature in the past user behavior and also the information about the user social interaction with other users, provide a richer background to build a more accurate recommendation model.

There are two main approaches that have been proposed in the literature: similarity-based and graph-based approaches. In this work we propose a similarity-based approach that incorporates user preferences, social influence, sequentiality mobility information and geographical influence to generate recommendations. The approach is based on a recommendation policy that integrates sequential mobility, as expressed by the movements of Twitter users, with the preference and attitudes of users towards a location. The strategy, referred to as *User Preference and path Similarity Recommendation (UPSR)*, combines users similarity in the way they visit locations and modeled in the *User Preference Recommendation (UPR)* policy, with the similarity in the way they move among such locations, modeled in the *Path Similarity Recommendation (PSR)* policy.

A ranking function is then used to determine the extend to which a given user u could be interested to visit an unseen location l. The general formulation of the ranking function is as follows:

$$rank(u,l) = \frac{\sum_{j=1}^{n_u} sim(u,u_j).s(u_j,l)}{\sum_{j=1}^{n_u} sim(u,u_j)} \qquad (1)$$

where the predicted score of u for the location l is defined by the average rating of other users u_j on the location l, denoted as $s(u_j,l)$, weighted by their usage similarity with u, denoted as $sim(u,u_j)$

To measure the similarity between the profiles of two users, different similarity measures can be used. Cosine similarity can be effectively used to measure the similarity between two users. Therefore, the ranking function can be rewritten as follows:

$$rank(u,l) = \frac{\sum_{j=1}^{n_u} cos(u,u_j).s(u_j,l)}{\sum_{j=1}^{n_u} cos(u,u_j)} \qquad (2)$$

Given a user u, a ranking score is computed for u across all the unvisited locations. A recommendation list for user u is produced by selecting the top K locations with the highest ranks.

According to Equation 2, the overall recommendation score is then computed as the linear function of the user preference similarity score and the path similarity score:

$$UPSR(u,l) = rank_{UPR}(u,l) + rank_{PSR}(u,l) \qquad (3)$$

The following of the section describes the two policies PSR and UPR that compose the proposed recommendation strategy.

A. Path similarity-based recommendation

The user current location is key for the recommendation process: it indicates a spatial constraint for generating recommendations as people are more likely to visit nearby locations than distant ones. Furthermore, the current location could also influence user preferences. For example, travelling to Rome venues like museums or archaeological sites could have a high recommendation rank, even though the user typically prefers sports events. Another key aspect to consider when designing a recommender is that, due to the sequential property of locations, a user's current location affects future travel decisions. For instance, the majority of people visiting Trevi Fountain will subsequently travel to Pantheon, or a restaurant recommendation may be appropriate after being at the theatre. Discovering these sequential relations and incorporating them into recommendations is key.

Accordingly, we introduced a policy referred to as *Path similarity based recommendations* (PSR) [2] in which we model the sequential relation in visiting locations by mining the paths travelled by users and exploiting spatio-temporal dynamics in the flows between venues so as to capture the factors that may drive users' movements. The rationale of the proposed approach is to recommend the next location visited by exploiting historic data, like the sequence of locations visited in the past as expressed in the geo-coordinates of tweets. These sequences to which we refer to as mobility traces or trajectories or paths are not explicitly available from Twitter posts, thus, ad-hoc methods have to be defined. In the faced setting, since there may be millions of distinct check-in locations in an LBSN, this task is even more complicated as the location prediction space is rather big. As a consequence, to guarantee satisfactory performance we focus on a subset of data that is particularly relevant for the analyses we want to perform. We constraint the candidate venues to the set of locations in the city from where a significant number of tweets have been posted. To this aim we extract and analyze frequent mobility patterns to improve location prediction. To determine the mobility patterns, different mobility models can be used. In the following a set of definitions characterizing the different typology of mobility regularities are introduced.

***Definition** 3:* **Mobility Pattern of a user**. A mobility pattern or a frequent travel route (or frequent path) of a user u is a sequence of locations frequently visited by u in a consecutive temporal order, with a frequency no smaller than a minimum support s_{min}:

$$\mathrm{MP_u = v_{l_0,u,t_0} \longrightarrow v_{l_1,u,t_1} \longrightarrow \ldots \longrightarrow v_{l_k,u,t_k}(s)}$$

with $\mathrm{MP_u} \subset \mathcal{P}_\mathrm{u}$, $t_i < t_{i+1}$, and where $\mathcal{P}_u$ is the set of paths travelled by u, s is the percentage of paths of users u that contain the patterns $\mathrm{MP_u}$, with $\mathrm{s \geq s_{min}}$.

We adopt a two-phases approach for mining popular travel routes: (i) the first phase consists of applying sequential pattern mining on the location sequences; (ii) the second one consists of extracting the maximal frequent sub-sequences from all the frequent sequences mined. This second step is necessary in order to ensure that trajectories with large segments in common are not reported simultaneously.

To this aim we propose an algorithm that extends the well-known PrefixSpan [5] algorithm to obtain only maximal frequent patterns.

Different formulations of mobility patterns are introduced, based on whether one is interested in collective or individual mobility analysis, exploiting, respectively all the trajectories or a subset of them for the study.

***Definition** 4:* **Personal Mobility Patterns** ($\mathcal{PMP}$): are mobility regularities of a specific individual obtained by considering only its past travel routes. The set of mobility patterns of a given user u, is the union of all its mobility patterns:
$\mathcal{PMP}_u = \bigcup_{i=1 \ldots N_P} MP_u(i)$ where N_P is the overall number of mobility patterns of u.

***Definition** 5:* **Cumulative Mobility Patterns** ($\mathcal{CMP}$): are the union of the personal mobility patterns of all the users. First mobility behaviour for each user are mined by exploiting only its mobility history, then all the individual models are merged and exploited for the prediction.
$\mathcal{CMP} = \bigcup_{u=1 \ldots N_U} \mathcal{PMP}_u$ where N_U is the number of distinct users.

***Definition** 6:* **Mass (Crowd) Mobility Patterns** ($\mathcal{MMP}$): are obtained by considering the trajectories of all the available users to detect global behaviors based on the assumption that people often follow similar movement patterns. Thus, are patterns that are classified as frequent because are common to several people. In fact, a global routine, instead of representing the systematic movement of an individual, represents a common behavior of the crowd.

$$\mathrm{MMP = v_{l_0,u,t_0} \longrightarrow v_{l_1,u,t_1} \longrightarrow \ldots \longrightarrow v_{l_k,u,t_k}(s)}$$

with $\mathrm{MMP} \subset \mathcal{P}$, $t_i < t_{i+1}$, and where s is the percentage of overall travel routes that contain $\mathrm{MP_u}$, with $\mathrm{s \geq s_{min}}$. The overall mass mobility patterns are as follows: $\mathcal{MMP} = \bigcup_{i=1 \ldots N_M} MMP_i$ where N_M is the overall number of mass mobility patterns.

***Definition** 7:* **Hybrid Mobility Patterns** ($\mathcal{HMP}$): exploits personal and crowd mobility patterns.
$\mathcal{HMP} = \mathcal{PMP} \bigcup \mathcal{MMP}$

***Definition** 8:* **Region Mobility Patterns** ($\mathcal{RMP}$): the overall mobility patterns of a geographic region R are the union of the personal personal mobility patterns of all the users (the cumulative patterns) and the crowd mobility patters. We refer to such patterns also as *collective* mobility patterns.
$\mathcal{MP} = \mathcal{CMP} \bigcup \mathcal{MMP}$

The PSR policy can exploit one of the above introduced mobility patterns model to account sequential mobility; clearly, its accuracy can vary according to the mobility model chosen.

To estimate the interest of a user u in a given unseen location l we consider a set of paths the user u follows not including location l. We then consider the other users that followed at least a path including location l. If the path history profile of user u is similar to most of the path history of such other users, then the probability that user u could be interested in l is high.

For each user u we consider her path history based on the paths that u has travelled along with the corresponding normalized frequencies of travels:
$\mathrm{ph_u = (f_{u,p_1}, f_{u,p_2}, ..., f_{u,p_p})}$

The normalized frequency of taking a path by a user is expressed as the number of times that a given path p is travelled out of the travels that have been done along all the paths:

$$\mathrm{f_{u,p}} = \frac{\mathrm{t_{u,p}}}{\sum_{\mathrm{j=1}}^{\mathrm{n_p}} \mathrm{t_{u,p_j}}} \tag{4}$$

The ranking function to measure the similarity between the visiting profiles of two users is expressed as follows:

$$rank_{PSR}(u,l) = \frac{\sum_{j=1}^{n_u} cos(ph_u, ph_{u_j}).f_{u_j,p}}{\sum_{j=1}^{n_u} cos(ph_u, ph_{u_j})} \tag{5}$$

B. User preference for a location

User preference towards a location l is measured according to the visit similarity with other users.

The metrics referred to as *User Preference similarity Recommendations* (UPR), accounts users similarity according to the way they visit locations. A basic assumption in location recommendation is that similar users have similar preferences on locations. This assumption is actually taken from the collaborative filtering world, which applies the collaborative filtering method directly over the venues.

To estimate the interest of a user u in a given location l we consider the set of other users that visited the same location l and compute the similarity of u to such users. If u is similar to most of these users, then there is a high chance that u will be interested in location l too. For each user u we extract her location visiting history based on the locations that u has visited and their visiting normalized frequencies. Accordingly, the visiting history of user u is represented as the following vector:
$\mathrm{vh_u = (f_{u,l}, f_{u,l_i}, ..., f_{u,l_n})}$

The visiting normalized frequency of a user to a location is expressed as the number of times that a given location l at time t is visited out of the visits that have been done in all the locations in the data set.

$$\mathrm{f_{u,l}} = \frac{\mathrm{v_{u,l}}}{\sum_{\mathrm{j=1}}^{\mathrm{n_l}} \mathrm{v_{u,l_j}}} \tag{6}$$

The ranking function for this policy is as follows:

$$rank_{UPR}(u,l) = \frac{\sum_{j=1}^{n_u} cos(vh_u, vh_{u_j}).f_{u_j,l}}{\sum_{j=1}^{n_u} cos(vh_u, vh_{u_j})} \tag{7}$$

V. Experimental evaluation

In this section, we report the main results of the experimental evaluation performed to assess the effectiveness of the proposed method.

A. Twitter dataset

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]. The dataset consists 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.

B. Evaluation methodology

This section describes the evaluation methodology used by introducing the performance metrics and the baseline models against which the proposed approach is validated.

1) Baseline recommendation policies: To show the effectiveness of the proposed approach, a set of state-of-the-art recommendation methods have been used for performance evaluation comparison. Since the main factors contributing to the unified recommendation model UPSR are sequential mobility, user preference and geographic constraints, a number of reference related approaches are used as baseline models for evaluation comparison. The baseline models are listed below.

- *User similarity based recommendations (USR).* With this model only user preference is used for recommendation as expressed in Equation 7.
- *Location similarity based recommendations (LSR)* This policy is based on the way a location is typically visited by users. Given a location l not already visited by user u, if the way user u visits other locations is similar to the ones of the rest of the users that already visited also location l, then there is a high probability that u will visit in the same way location l. To this purpose, each location is represented by its visit history, a vector that maintains for each user the normalized frequency of visits to the location l: $\mathrm{vh_l} = (\mathrm{f_{u_1,l}}, \mathrm{f_{u_2,l}}, ..., \mathrm{f_{u_j,l}})$. In this case the scoring function is as follows:
  $rank_{LSR}(u,l) = \frac{\sum_{j=1}^{n_l} cos(vh_l, vh_{l_j}).f_{u,l}}{\sum_{j=1}^{n_l} cos(vh_l, vh_{l_j})}$
  the K new locations with the highest scores are returned as the recommended ones.
- *Spatial Distance based recommendation (SDR).* This policy is based only on the spatial distance of an unseen location l to the locations that have been previously visited by a user denoted as $\mathcal{L}_{v_u}$. The ranking function is as follows:
  $rank_{SDR}(u,l) = 1 - min_{l_i \in \mathcal{L}_{v_u}} dist(l, l_i)$
  with $\mathrm{l} \in \mathcal{L} \setminus \mathcal{L}_{v_u}$.
  We determine the distance among two GPS points, expressed as latitude and longitude coordinates, by using the Haversine distance. The Haversine distance formula is an equation giving great-circle distances (that is, the shortest distance over the earth's surface) between two points on a sphere from their longitudes and latitudes:
  $d_h(l, l_i) = 2r \arcsin d \quad \text{with}$
  $d = \left(\sqrt{\sin^2(\frac{lat_{l_i} - lat_l}{2}) + \cos(lat_l)\cos(lat_{l_i})\sin^2(\frac{lon_{l_i} - long_l}{2}} \right)$
  and where $r = 6,371$ Km is the earth's mean radius.
- *First-order Markov Chain (FMC).* This method, proposed in many works like [1], is the state-of-the-art personalized successive POI recommendation method. It is based on a matrix factorization method to embed the personalized Markov chains and the localized regions. However the approaches implementing such method derive the sequential probability of user u to new location based on only the latest visited location in the sequence.
- *PSR* [2]. This metrics accounts only the sequential mobility influence for location recommendation.

2) Performance metrics: To perform the evaluation of the proposed approach, we divide a user's location history into two parts: we select the location history generated till a given time t_{past} as a test set and we use the rest of the users' location history as a training set to learn the users' preferences. We regard the venues that a user has visited till the time t_{past} as the ground truth and match the recommended locations against these venues. The more recommended locations truly visited by a user in the test city, the more effective the recommendation method is.

We consider the minimum bounding box of all the visited locations in the ground truth to simulate the geospatial range that would be specified in the users' recommendation request. Based on the ground truth and recommendations,

[1] http://semiocast.com/

we are able to asses the effectiveness of the approach by computing accuracy and prediction rate metrics:

- *Prediction Rate*: it is the percentage of successful predictions.
  Prediction Rate= $\frac{PredictionsDone}{PredictionsRequested}$.
- *Accuracy*: it is the fraction of the correct predictions over the total number of successful predictions.
  Accuracy= $\frac{CorrectPredictions}{PredictionsDone}$

In particular, we compare the different ranking strategies when using different recommendation list sizes K. In this case, we successfully predict the recommendation if we rank a location in the top-K places.

Each of these metrics is then averaged over all target locations in our evaluation set to measure the performance of the different recommending policies.

C. Results

1) Effectiveness of the approach: Figure 1 shows the accuracy achieved by the recommendation approach with respect to the prediction list size when using different mobility models. As expected, the personal mobility model (PMP) provides more accurate predictions. Nevertheless, the mobility implemented by the region mobility model (RMP) got a remarkable accuracy particularly evident for larger list size. The collective mobility model (CMP), which exploits all the personal patterns, got the lower accuracy, especially for small list size, on average is about 15% less than the personal and region strategies. The crowd mobility approach (MMP) is on average 5% more accurate than the collective model but its accuracy is substantial lower than the personal model even if it improves with the list size. The hybrid mobility model (HMP), as expected, achieved an accuracy slightly worst than UPSR since it does not take into account the collective mobility routines (CMP), and it improves the accuracy compared to the collective model.

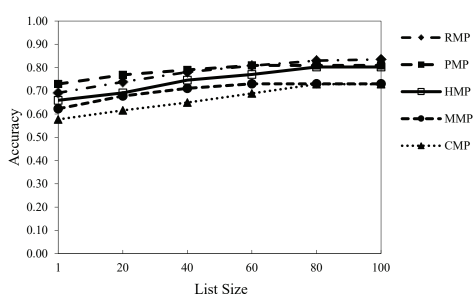

Figure 1. Accuracy w.r.t. List Size, when using different mobility models.

From Figure 2 is evident that the mobility model implemented by the region mobility model (RMP) outperforms the other models in terms of prediction rate: the patterns extracted by RMP represent all the possible mobility routines of a given geographic region, both the ones specific to the individuals (PMP patterns) and the ones of the collectivity (cumulative and crowds patterns). In fact, the use in RMP of collectivity mobility knowledge strongly increases the prediction rate since almost all the requests have a prediction. All the mobility models including collective routines achieved a good prediction rate, like the hybrid and the crowd models, while the models using personal routines achieved the lowest results in terms of prediction rate.

Overall, the region mobility model achieved the best performance: even if its accuracy is slightly lower than the personal mobility model the advantage in terms of prediction rate makes this model largely better than the others. In the rest of the experiments the region mobility model RMP will be used in the proposed UPSR recommendation strategy.

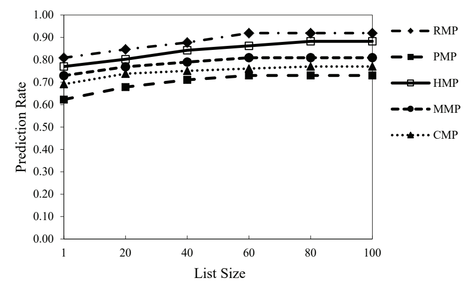

Figure 2. Prediction w.r.t. List Size, when using different mobility models.

2) Comparison with baseline models: Figure 3 and Figure 4 show how the performance of the different approaches varies with the recommendation list size.

The proposed method outperforms baseline approaches significantly. Figure 3 shows that the accuracy increases with the number of recommended locations for all the policies. However, they achieved a quite different trend, as summarized in the following.

LSR drops behind all the other methods. This outcome indicates that item-based CF is not an effective approach since venues in LBSNs may not have been visited by sufficient many users and, thus, the computed similarity between two locations may not provide a good clue to decide whether a user likes a location. The policy LSR got the lowest performance also in terms of prediction rate, showing the benefit of exploiting other than user's location history also visiting similarities with other users and social relations among them.

USR and SDR perform poorly because they do not make use of sequential information. Those results show that the conventional recommendation algorithms, which mainly exploit the user preference, are not effective for recommendation in LBSNs since they don't exploit neither spatio-temporal information nor social relationships. The outcome highlights the advantage obtained by considering into the recommendation model the sequential relation in visiting locations as expressed in the paths travelled by users.

It is woth noting that UPSR exceeds both USR and SDR not only because it accounts sequential mobility but also because is more capable of modeling a user's preferences through the proposed UPR model.

SDR does not exploit user preference neither takes into account sequential mobility, for this reason is outperformed by almost all the other methods.

The relatively high performance of the proposed model, FMC and PSR methods demonstrates that the sequential information plays an important role in location recommendation. However, that is not the only factor highly impacting on the recommendation effectiveness as it is confirmed by the results showing that UPSR outperforms both FMC and PSR.

The main limitation of the FMC approach is that it models the sequential influence by utilizing only the latest visited location in the check-in sequence of a user to derive her visiting probability to new locations. As a result, it does not take full advantage of sequential patterns in location recommendations, since it ignores the impact of the earlier visited locations in the sequence on the new likely visiting locations. Thus, FMC returns the most inaccurate venues in terms of accuracy and misses most venues actually visited by target users in terms of prediction rate, as depicted in Figures 3 and 4.

The improvement brought by UPSR on PSR is also due to the more accurate user preference model implemented. These results outline that the spatial information together with the mobility data of friends with similar travelling attitudes plays a key role in the performance of a location recommendation system.

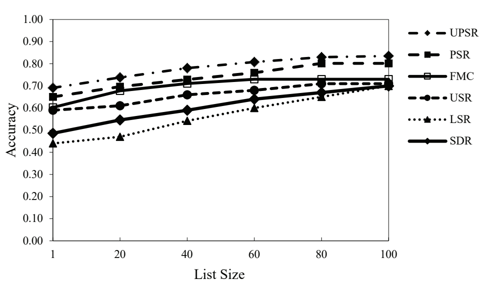

Figure 3. Accuracy w.r.t. List Size, for the different recommendation approaches.

From such results, we can conclude that the proposed method is a promising solution for recommending new locations exploiting human mobility from social media data. It is formulated as a novel hybrid solution combining user similarity mobility patterns and spatio-temporal historical visiting information.

VI. Conclusions

In the paper is presented a novel approach for location recommendation from social network exploiting typical patterns usually travelled by users and the spatio-temporal features characterizing the movements among locations. The recommendation problem is formulated as a similarity problem among the visiting and mobility profiles of users. The candidate locations are then ranked according to the similarity measure. In particular, given a user u, a ranking score is computed for u across all the unvisited locations. A recommendation list for users u is produced by selecting the top K locations with the highest ranks. A dataset of tweets collected in London city from January to June 2013 have been analyzed. The experimental results have shown that the proposed method is effective in recommending new places achieving remarkable accuracy and prediction rates.

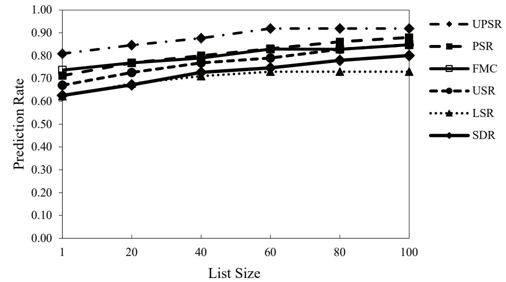

Figure 4. Prediction w.r.t. List Size, for the different recommendation approaches.

References

[1] C. Cheng, H. Yang, M. R. Lyu, and I. King. Where you like to go next: Successive point-of-interest recommendation. In *Proceedings of the Twenty-Third International Joint Conference on Artificial Intelligence*, IJCAI '13, pages 2605–2611.

[2] C. Comito. Exploiting social media for recommending new locations. In *28th International Conference on Computer Communication and Networks, ICCCN 2019, Valencia, Spain, July 29 - August 1, 2019*, pages 1–6. IEEE, 2019.

[3] I. Konstas, V. Stathopoulos, and J. M. Jose. On social networks and collaborative recommendation. In *Proceedings of the 32nd Annual International ACM SIGIR Conference on Research and Development in Information Retrieval, SIGIR 2009, Boston, MA, USA, July 19-23, 2009*, pages 195–202.

[4] J. J. Levandoski, M. Sarwat, A. Eldawy, and M. F. Mokbel. Lars: A location-aware recommender system. In *Proceedings of the 2012 IEEE 28th International Conference on Data Engineering*, ICDE '12, pages 450–461, 2012.

[5] J. Pei, J. Han, B. Mortazavi-Asl, H. Pinto, Q. Chen, U. Dayal, and M.-C. Hsu. Prefixspan: Mining sequential patterns efficiently by prefix-projected pattern growth. In *ICDE*, pages 215–, 2001.

[6] M. C. Pham, Y. Cao, R. Klamma, and M. Jarke. A clustering approach for collaborative filtering recommendation using social network analysis. *J. UCS*, 17(4):583–604, 2011.

[7] T. Stepan, J. M. Morawski, S. Dick, and J. Miller. Incorporating spatial, temporal, and social context in recommendations for location-based social networks. *IEEE Trans. Comput. Social Systems*, 3(4):164–175, 2016.

[8] M. Ye, P. Yin, and W. Lee. Location recommendation for location-based social networks. In *18th ACM SIGSPATIAL International Symposium on Advances in Geographic Information Systems, ACM-GIS 2010*, pages 458–461.

[9] M. Ye, P. Yin, W. Lee, and D. L. Lee. Exploiting geographical influence for collaborative point-of-interest recommendation. In *Proceeding of the 34th International ACM SIGIR Conference on Research and Development in Information Retrieval, 2011*, pages 325–334.

[10] H. Yildirim and M. S. Krishnamoorthy. A random walk method for alleviating the sparsity problem in collaborative filtering. In *Proceedings of the 2008 ACM Conference on Recommender Systems, RecSys 2008, Lausanne, Switzerland, October 23-25, 2008*, pages 131–138, 2008.

[11] Y. Zheng and X. Xie. Learning travel recommendations from user-generated GPS traces. *ACM TIST*, 2(1):2:1–2:29, 2011.