

Recommender System Based Trajectory Prediction

推薦システムに基づく軌跡予測

学籍番号 47216801
氏 名 莫 宇 (Mo, Yu)
指導教員 宋 軒 准教授
小林 博樹 教授

Introduction

In the ever-evolving landscape of intelligent transportation systems, trajectory prediction has surfaced as a critical element in optimizing transportation networks and devising urban road planning strategies. Accurate and efficient trajectory prediction is paramount for enhancing safety measures, improving traffic management, and facilitating effective route planning. Trajectory prediction serves as a vital approach to address social problems and elevate modern transportation systems. By leveraging a data-driven methodology to ease traffic congestion, optimize transportation networks, and streamline traffic management, trajectory prediction has solidified its position as an indispensable aspect of Mobility as a Service (MaaS). The concept of MaaS integrates various transportation modes into a unified service, aiming to deliver seamless mobility experiences. Precise trajectory prediction is essential for achieving this objective. Moreover, the applications of trajectory prediction extend beyond MaaS, encompassing other fields such as autonomous driving and sports analysis. By offering accurate and perceptive predictions, trajectory prediction holds the potential to

transform multiple sectors and foster a more efficient and sustainable future.

Despite extensive research has been conducted in the field of trajectory prediction, providing accurate predictions of the next location based on past trajectories remains a challenging problem, influenced by the factors like the balance between research area and predicting accuracy, and the complex external factor including driver's habit and effect of holidays. To address these challenges, this thesis proposes a novel approach for city-wide multi-modes trajectory prediction, termed as (RSTP). This approach deviates from conventional trajectory prediction as it employs a grid-based trajectory instead of the original GPS point sequences. Fig.1 gives an intuitive example of grid trajectory.

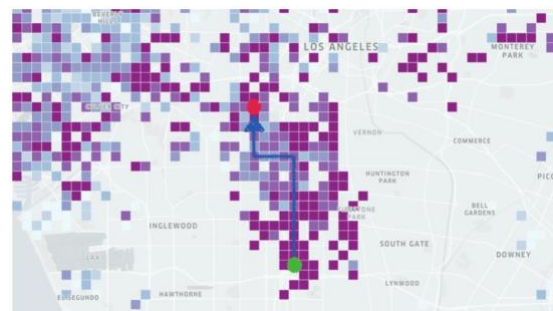


Figure 1. A Demo of Grid Trajectory

Methodology

The original trajectory is represented by a sequence of GPS points (latitude, longitude) with time order. Traditional trajectory prediction method is to directly predict the exactly point. Obviously, directly predict the coordinate is a very challenging task since the solution space is huge. That is why traditional trajectory prediction approach is always conducted in a limited area. To apply the Recommender System approach into the trajectory prediction task, a new representation of trajectory is required. Relace precise GPS point with a grid ID is considerable, where the grid is a set of GPS points. For instance, if we use grids with size of $100\text{m} * 100\text{m}$ to divide map, then each grid will represent any point located in the grid area. In this research, I will use the grid-based map and grid-based trajectory instead of using original trajectory. In this way, the problem can be greatly simplified since our prediction target reduce from infinity to countable (usually 100 thousand grids in a city level).

Based on the model proposed by Covington et al. in their work named “Deep Neural Networks for YouTube Recommendations” (for convenience, using YouTubeDNN to note this model in the later part of this abstract). YouTubeDNN leverages deep learning to generate accurate and personalized recommendations for YouTube users. The model is designed to overcome the challenges associated with the large scale and diverse content available on the platform, as well as accommodate the wide user preferences. Thus, the structure is straightforward and easy to

reproduce. Except for the feature engineering, the core part of YouTubeDNN is to use 3 DNN layers to obtain the vector or embedding of YouTube users according to their behaviors and based on those embeddings to compute the similarity between the users and items (videos).

Borrow this idea, the thesis proposed a model called Recommender System based Trajectory Prediction model, RSTR. The design principle of RSTP is also using Embedding and MLP structure to deal with the problem that a large scale and diverse grid cells to predict. The MLP layer is also know as Deep Neural Network, DNN. Additionally, to capture the temporal information behind the trajectory data, the Long Short-Term Memory (LSTM) layer is added to modeling the temporal feature before the data is feed into the DNN layers. LSTM is a type of Recurrent Neural Network (RNN) architecture that is designed to learn and remember long-range dependencies in sequential data. LSTMs utilize special memory cells and gating mechanisms to control the flow of information, allowing them to effectively capture and retain information over long time periods. As for the spatial feature construction, the deep learning techniques call DeepWalk is applied, which is effective method to capture the information behind the network structures. Moreover, to make sure the full utilization of the spatial information and temporal information at the same time, Attention mechanism is also used in the model modeling the features across the temporal dimension and the spatial dimension, which is important to the final performance.

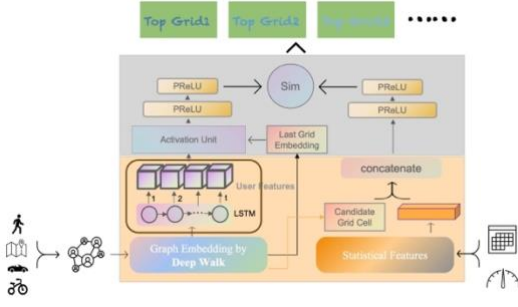


Figure 2. Framework of RSTP

As shown in Fig.2, RSTP is consist of two modules as a typical Recommender System do: the Embedding Construction module (the lower part) and the Candidate Generation module (the upper part). In each module, it contains several functional components to deal with different information. More specifically, Embedding Construction module is comprised of three components: Temporal Component, Spatial Component and Attribute Component. And the Candidate Generation module is comprised of two components: Interaction Component and Output Layers.

Datasets

Two real-word datasets are used in following experiments, which are the Tokyo dataset and the Osaka dataset. Tokyo dataset comprises 612,010 trajectories, encompassing over 120 million GPS records. These data were collected from 2,319 users over a period from August 23rd to August 30th, 2022. Osaka dataset is comprised of 368,216 trajectories, which is collected from 10,53 users over a period from August 23rd to August 30th, 2022.

Experiments

The extensive experiments are conducted to

verify the feasibility of the proposed RSTP model. All the experiment result are summarized into the following 3 tables.

Datasets	Methods	K@1	K@10	K@100
Tokyo	AN	22.98% (9)	25.46% (25)	40.71% (81)
	LSTM	0.11%	-	-
	YouTubeDNN	0.76%	3.92%	9.74%
	Temporal RSTP	23.37%	25.58%	30.68%
	Spatial RSTP	1.23%	8.32%	19.39%
	RSTP (proposed)	23.66%	28.33%	34.76%
Osaka	AN	18.59%(9)	26.66%(25)	48.37%(81)
	LSTM	0.13%	-	-
	YouTubeDNN	0.18%	3.32%	6.84%
	Temporal RSTP	20.18%	28.32%	29.23%
	Spatial RSTP	1.49%	6.32%	20.33%
	RSTP (proposed)	25.77%	30.43%	48.53%

Table 1. Performance Comparison

Datasets	Methods	K@1	K@10	K@100
Tokyo	Temporal RSTP	23.37%	25.58%	30.68%
	Spatial RSTP	1.23%	8.32%	19.39%
	RSTP	23.66%	28.33%	34.76%
	Combined	23.37%	25.90%	31.48%
Osaka	Temporal RSTP	20.18%	28.32%	29.23%
	Spatial RSTP	1.49%	6.32%	20.33%
	RSTP	25.77%	30.43%	48.53%
	Combined	20.33%	29.42%	34.53%

Table 2. Verification of the functionality of Spatiotemporal Component

Datasets	Methods	K@1	K@10	K@100
Tokyo	AN	22.98% (9)	25.46% (25)	40.71%(81)
	RSTP	23.66%	28.33%	34.76%
	Combined	30.71%	36.09%	48.61%
Osaka	AN	18.59%(9)	26.66%(25)	48.37%(81)
	RSTP	25.77%	30.43%	48.53%
	Combined	31.33%	36.83%	51.66%

Table 3. Functionality Analysis of Proposed Model

The metrics is called Hit Rate and hit rate equals the number of Hit sample divided by the number of samples. Method AN means we choose the neighbored grids of last appeared grid as the predict results. And the number in the brackets means how many neighbors we chosen, since it is not possible to only pick on nearest neighbor in this map division. Method **Combined** means the numerical combination of the results given by

two different models. Depending on the experiments, the combined model is different. Though the experiments, 3 findings can be summarized as: Firstly, the spatiotemporal components, the proposed model has better performance than the baseline model in all the cases, but sometimes it weaker than the brute force method AN. Secondly, As the number K increases, the performance of proposed model will increase accordingly, but it is not quick as the AN method in some scenarios. Lastly, the proposed model predicts the next location with a different idea the brute force method, since in table 3, the performance of combined method is better than any base method. If the proposed mode can assort the advantage of the brute force AN method, the further improvement on proposed RSTP model can be expected.

Conclusion

In this thesis, a Recommender System based Trajectory Prediction (RSTP) model is proposed to address the grid-based trajectory prediction problem, striking a balance between prediction accuracy and the scope of the research area. The effectiveness of this modified Recommender System approach, which specializes in modeling spatiotemporal information, has been substantiated through extensive experiments, reinforcing the practicality of applying a Recommender System framework to the trajectory prediction problem. From a comprehensive literature review and the empirical investigations, it is observed that the proposed model exhibits

clear practical utility and shows potential to forge a new approach for tackling trajectory prediction tasks. While the results may not be superior in all scenarios, the case studies suggest that this approach could potentially be a state-of-the-art technique, particularly when larger and more diverse datasets are taken into consideration. In future work, I will incorporate more context information such as the Point-Of-Interest (POI) information and long-term temporal information into the model to further improve the prediction performance.