Trajectory Prediction from External Sensor Data using Recurrent Neural Networks

Lívia Almada Cruz
Campus Quixadá
Federal University of Ceará
Quixadá, Brazil
livia.almada@ufc.br

Karine Zeitouni
DAVID Laboratory - CNRS
University of Versailles St Quentin
Versailles, France
karine.zeitouni@uvsq.fr

José A. F. de Macedo
Department of Computing Science
Federal University of Ceará
Fortaleza, Brazil
jose.macedo@dc.ufc.br

Abstract—**This paper presents a study of location prediction applied to trajectories obtained from sensors placed on road-networks. We applied a variation of Recurrent Neural Networks using different combination of features, to measure the impact of each feature on the learning task.**

Index Terms—**mobility analysis, location prediction, recurrent neural networks, sensor trajectories**

I. INTRODUCTION

The high availability of tracking data brought opportunities to provide new methods to analyze and understand mobility patterns. Among the analysis, location prediction has gained some attention given its applicability. The mobility prediction (or location prediction) problem focuses on inferring the next relevant location of a moving object based on historical trajectories and its most recent tracking. Many applications, like smart transportation, traffic control, urban planning and recommendation systems, can benefit from location prediction. Three levels of mobility prediction were considered in [1]: 1) object position; 2) path prediction; and 3) next place prediction. In levels 1 and 2, usually the models are learned from raw trajectories obtained by Global Positioning System (GPS) devices and their predictions consider the movement of the object. Level 3 predicts stops rather than movements, usually by learning from sequences of points of interest or events. Differently of previous works, in this work, we analyze movement of objects tracked by sensors placed on the roads sides. Each sensor captures and registers the passage of moving objects. Assuming that each register contains enough information to uniquely identify the associated moving object, it is possible to derive from them the trajectories of the moving objects observed by the sensors. Our work is in between of level 1 (object position) and level 2 (movement and path).

II. PROBLEM STATEMENT

1) Sensor Trajectory Prediction: Given a set of sensors $S = \{s_1, s_2, \ldots, s_n\}$, when a sensor s_i captures the passage of a moving object m at timestamp t, it registers an observation

$o = (s_i, m, t)$. A trajectory $traj = <o_1, o_2, \ldots, o_l>$ of a moving object m is the sequence of observations associated to m, where $t_i \leq t_{i+1}$. The trajectories are defined over a period of the day.

The problem tackled here is to predict the next sensor given the last k observations of a moving object (also called recent trajectory) and the set of historical trajectories.

Different from classification problems with many classes, we have to consider complex transitions patterns and time-dependence. As the sensors have spatial relationships among them, the proximity of the predicted location to the actual value is also important. Furthermore, the trajectories obtained from sensors have a set of particularities that give us new opportunities and challenges.

1) **Huge data:** Sensors continuously capture huge number of observations per day.
2) **Exhaustive types of trajectories:** Moving objects are not restricted to a specific fleet of vehicles. Commuters, fleet of taxi or buses, deliveries, etc. can all of them be tracked by the sensors. Thus, trajectories can have very different patterns.
3) **Sparsity:** The sensors are located on fixed positions, usually only on the main roads of the city. The entire tracking of moving objects is not available and the trajectories are very sparse in space and time.
4) **Incompleteness and uncertainty:** Sensors may fail to capture the passage a vehicle, producing incomplete trajectories. It is not obvious when one observation is not in the data set because the sensor failed or if it is because the object did not passe by the sensor.

Recently, applying Recurrent Neural Networks (RNN) to location prediction has demonstrated the potential of these approaches to capture the complexity of mobility data. However, at the best of our knowledge, none of those works have studied location prediction for trajectories based on external sensor data. We call this problem "Sensor trajectory Prediction". In this work, we evaluate RNN models based in different set of features in order to understand the limitations of the predictability of such trajectories. We compare these approaches with the ones based on sequence patterns and Markov models. Finally, we discuss the results and future
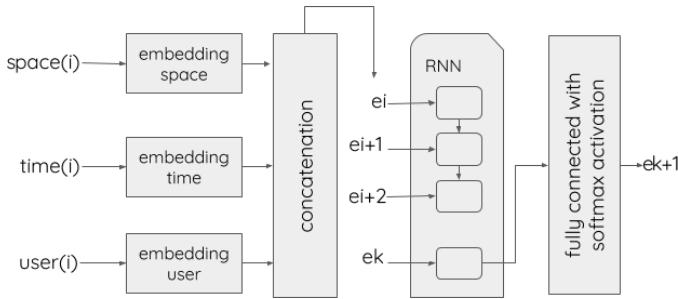
Fig. 1. General architecture of the RNN for location prediction.

directions.

III. RELATED WORKS

TPRED [2] predicts the next stop from GPS trajectories. The work in [3] predicts the next stop based on groups of users who share the same profile (e.g. gender and age). GMove [4] uses spatial-temporal information and geo-tagged text from check-ins to predict the next stop. MyWay [5] predicts the position of moving objects from their GPS trajectories based on the clustering and spatial match. In [6],a Dynamic Bayesian Network predict next location of sparse trajectories from call details records. In [7], a Spatial-Temporal RNN predicts the next location using spatial and temporal continuous values. SERM [8] is a spatial-temporal RNN to predict the next stop. TA-TEM [9] is a recommendation system based on RNN which predicts the next stop. Both works [8] and [9] learned from check-ins. DeepMove [10] uses attention RNN to predict the stop on the next time window.

IV. RNN FOR SENSOR TRAJECTORY PREDICTION

The general architecture is a simplification of the model proposed by [8]. The model (Figure 1) is composed by: *i.* An embedding layer, responsible of reducing the dimensions of input vectors; *ii.* A layer to concatenate the output of embedding layers in order to get a unique input feature vector; *iii.* A recurrent layer to learn the complex patterns from sequences; *iv.* A fully connected layer with softmax function as activation, which converts the result of the recurrent layer into the set of probabilities to be assigned to each class label.

Our experiments were based on different features: the **spatial** feature corresponds to the sensor label; the **temporal** features is the time slot in a day which fits the timestamp of an observation; and user identification captures **user preferences**. We consider models with different combinations of these features: spatial model (**SM**), spatial-temporal model (**STM**), spatial-user model (**SUM**) and spatial-temporal-user model (**STUM**). We use the one-hot representation to transform each feature in a vector. A window of the k last observations is used to learn the next position.

V. EXPERIMENTAL EVALUATION

1) Dataset: Trajectories were collected from 01/09/2017 to 30/09/2017, from 272 sensors in the city of Fortaleza, Brazil. Originally, we received a total of 22,338,916 observations. We filtered the trajectories and keep those with a minimum of 6 observations. We obtained 1,025,040 trajectories from 266,522 distinct vehicles.

2) Results: Models were trained and tested 5 times using holdout 80-20. We also evaluate the accuracy of RNN models, Markov Models of First and 5th orders and TDAG [11] approaches (Table I). The STUM model reached the best accuracy, which confirms the users tend to have similar and time-dependent patterns. Even with only 4.5 trajectories per user (in average), including the user id improves the accuracy. SM suffers from overfitting in all executions, which means that RNN needs complementary information to learn mobility patterns from the trajectory paths only. We measure the error of RNN models according to the closeness given by the road distance from the actual location to the predicted one (Table II). The quality of models in terms of distance to ground truth is also improved by adding additional knowledge. Precisely, with this metric, the RNN models using time and user features showed a better performance.

VI. DISCUSSION AND FUTURE WORKS

In this paper we have shown preliminary results of the application of RNN to sensor trajectory prediction. This type of trajectories may capture very different mobility patterns, since it is not restricted to a fleet or a community of users. They are also sparse, incomplete and uncertain. We have also highlighted the use of the underline road-network to estimate a finer granularity trajectory definition and obtain better models in terms of accuracy and error of distance.

As ongoing work, we are studying how to deal with missing values by means imputation approaches for the sensor trajectories while we take into account the uncertainty. As future works, we want to use road network restrictions to discard undesirable predictions and enrich the models. Finally, we want to study how to improve the accuracy by means others machine learning techniques, like unsupervised learning.

REFERENCES

[1] D. Bucher, "Vision paper: Using volunteered geographic information to improve mobility prediction," in *Proc. of the 1st ACM SIGSPATIAL Workshop on Prediction of Human Mobility*, 2017, p. 2.

[2] C. L. Rocha et al., "Tpred: a spatio-temporal location predictor framework," in *Proc. of the 20th IDEAS*, 2016, pp. 34–42.

[3] E. Naserian, X. Wang, K. Dahal, Z. Wang, and Z. Wang, "Personalized location prediction for group travellers from spatialtemporal trajectories," *Future Generation Computer Systems*, pp. 278–292, 2018.

[4] C. Zhang et al., "Gmove: group-level mobility modeling using geo-tagged social media," in *Proc. of the 22nd ACM SIGKDD*, 2016, pp. 1305–1314.

[5] R. Trasarti, R. Guidotti, A. Monreale, and F. Giannotti, "Myway: Location prediction via mobility profiling," *Information Systems*, 2017.

[6] F. Alhasoun, M. Alhazzani, F. Aleissa, R. Alnasser, and M. González, "City scale next place prediction from sparse data through similar strangers," in *Proc. of ACM KDD Workshop*, 2017.

[7] Q. Liu, S. Wu, L. Wang, and T. Tan, "Predicting the next location: a recurrent model with spatial and temporal contexts," *AAAI*, pp. 194–200, 2016.

[8] D. Yao, C. Zhang, J. Huang, and J. Bi, "Serm: A recurrent model for next location prediction in semantic trajectories," in *Proc. of the 2017 ACM CIKM*, 2017.

[9] W. X. Zhao et al., "A time-aware trajectory embedding model for next-location recommendation," *Knowledge and Information Systems*, 2017.

[10] J. Feng, Y. Li, C. Zhang, F. Sun, F. Meng, A. Guo, and D. Jin, "Deep-move: predicting human mobility with attentional recurrent networks," in *Proc. of the 2018 WWW Conference*, 2018, pp. 1459–1468.

[11] P. Laird and R. Saul, "Discrete sequence prediction and its applications," *Machine learning*, vol. 15, no. 1, pp. 43–68, 1994.