Deep neural architecture for geospatial trajectory completion

Asif Nawaz
Department of Computer Science and Technology
Nanjing University of Aeronautics & Astronautics
Nanjing, China
asifnawaz@nuaa.edu.cn

Zhiqiu Huang
Department of Computer Science and Technology
Nanjing University of Aeronautics & Astronautics
Nanjing, China
zqhuang@nuaa.edu.cn

Senzhang Wang
Department of Computer Science & Technology
Nanjing University of Aeronautics & Astronautics
Nanjing, China
szwang@nuaa.edu.cn

Amara Naseer
Department of Computer Science & Technology
Nanjing University of Aeronautics & Astronautics
Nanjing, China
mari.choudry@yahoo.com

***Abstract*—GPS data is widely used in many real world applications. The quality of GPS data is critically important to produce high quality results. In real world applications, certain GPS trajectories are sparse and incomplete; which causes challenges to GPS trajectory based applications. Few of existing studies have tried to address this problem using complicated algorithms that are based on conventional heuristics; this requires extensive domain knowledge of underlying applications. Deep learning in recent era has achieved the great success to solve many sequence to sequence prediction problems. In this paper, a deep learning based bidirectional convolutional recurrent encoder decoder architecture using attention mechanism is proposed that predicts the missing data points; resulting in a complete GPS trajectory. The proposed method shows the significant improvement over state-of-the-art benchmark methods.**

***Index Terms*—GPS, Deep Learning, Trajectories**

I. Introduction

GPS data collected from the personnel's travel surveys around the world has grown exponentially over the past few decades, using the most advanced system GNSS (Global Navigation Satellite System), that are equipped with Galileo and GLONASS etc [1]. The process of GPS data collection suffers from some major drawbacks of missing data that usually occurs because of inaccurate signals, device errors and data privacy issues; that in result leads to missing GPS points. Higher quality densed GPS data is important to produce the results with higher accuracy. Common techniques used to deal with missing data is to replace the missing data with some values or ignore them, but it does not provide reliable estimation of the ground truth. Therefore, it becomes very important to innovate the techniques to reproduce missing GPS data.

The process of imputing missing parts of the trajectories involves to find the patterns in similar trajectories using multiple features. The trajectory completion task is quite challenging as it requires to learn complex traffic patterns in big cities environment [2]. Many studies like [3]–[5] try to derive the complex algorithms for trajectory completion with certain limitations.

- Traditional shallow methods like bayesian networks, markov models are not able to effectively capture non-linear relationships in the data.
- It uses sophisticated behavior modeling, that requires the system designers to do in-depth model optimization, and can only model the limited set of patterns.
- It requires large amount of time to hand-engineer the most representative features based on their domain knowledge.
- Few studies take support of GIS data, that is usually not available for all the locations.

Our approach is data-driven and easier to understand in that it automatically learns complex behaviors from the massive amount of GPS trajectories using deep neural network models. Based on the learning of these diverse set of patterns we are able to predict the missing parts of trajectories. GPS trajectory completion is kind of sequence to sequence prediction problem [6], [13]. Based on the recent success of deep learning techniques to solve many challenging problems like flow prediction [12], [14], we use state-of-the-art deep net based encoder decoder architecture [7] to solve our problem. Our study only uses the GPS data without the support of data from any other sensors, that make the problem more challenging.

Following are the core contributions of our work.

- To solve the trajectory completion problem, we for the first time propose deep learning based convolution recurrent encoder decoder architecture, to predict the complete sequence of GPS points from partial input of GPS trajectory points on occupancy grid map.
- Due to the spatiotemporal nature of GPS data, we adopt bidirectional ConvLSTM architecture as encoder to learn both spatial and temporal relationship in the data, with the ConvLSTM decoder using attention mechanism.
- Our model can effectively integrate global features and

auxiliary information that support the encoder-decoder model for trajectory completion task.

II. Methodology

1) Problem Definition: A GPS point is represented by 2D coordinates (x_t, y_t), where x and y are position of an object at time instance t. Given the observed locations of moving object at time $t = 1$ to $t = m-1$ and $t = n+1$ to $t = s$. Our aim is to predict the positions of objects from time $t = m$ to $t = n$. Thus, given the incomplete trajectory $X_{incomp} = (x_1, y_1), (x_2, y_2) \ldots (x_{m-1}, y_{m-1)} \ldots (x_{n+1}, y_{n+1}) \ldots .., (x_s, y_s)$, we aim to predict the missing part $X_{miss} = (x_m, y_m), (x_{m+1}, y_{m+1}) \ldots (x_{n-1}, y_{n-1}), (x_n, y_n)$ of length l in the completed trajectory $X_{comp} = (x_1, y_1), (x_2, y_2) \ldots (x_s, y_s)$, where l is the length of missing sub-trajectory X, and $l = n - m$.

2) Model Description: Our model is based on deep learning based encoder decoder architecture that is proved to be more successful in sequence prediction problems [6]. The encoder transforms the input GPS segment into fixed length context vector using attention mechanism. Context vector is then passed to the decoder, that is responsible for stepping through the output time steps while reading the context vector. Figure 1 elaborates our proposed architecture for trajectory completion task.

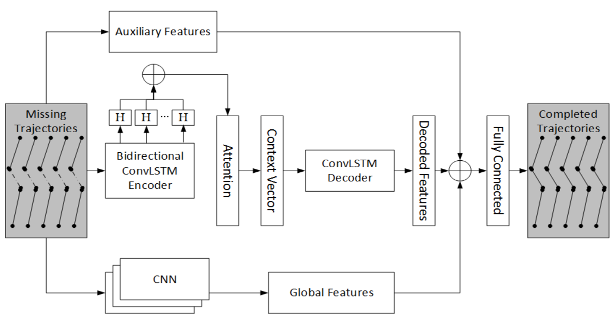

Fig. 1. Proposed approach architecture for trajectory completion.

Input to the architecture is a GPS segment. The geographical space is divided in to two dimensional grid, where all grid cell sizes are of same size [9]. Each grid cell is represented by a row and column number in a 2D grid, and each GPS point is represented by its respective row and column number in 2D grid. The architecture is trained into three parallel steps. In first step, basic input features like velocity, acceleration, jerk are feed into bidirectional encoder decoder architecture, that produces the decoded features. In addition to the decoded features, auxiliary features are extracted directly from input data. These are the features that remains the same throughout all timesteps of input samples. Instead of passing these features as encoder's input, we bypass these features and merge directly with the decoded features to reduce the model complexity. Global features are also extracted in parallel to auxiliary and decoded features. CNN is useful architecture using large filter sizes in order to extract the global features [10]. These features covers the wider geographical context of a trajectory. Global features in our context are the correlations among neighboring grid cells, representing the movement probabilities between different regions based on the densities of GPS points in the grid cells. Global features are obtained by training another convolutional network in parallel to encoder decoder network as shown in Figure 1. Decoded features are merged with auxiliary and global features and passed to fully connected layer, that predict the probability distributions of GPS points for all time steps. We used the widely used beam search algorithm [6] to sample the output sequence of GPS points with highest probability. To validate our model, we used mean squared error loss function J during the training using the equation 1 given below.

$$Loss(y, \hat{y}) = \frac{1}{N} \sum_{i=1}^{N} (y_i - \hat{y}_i)^2 \tag{1}$$

III. Experiments and Results

We used Microsoft geolife trajectory dataset [11] to evaluate the performance of our proposed approach. To conduct experiments, we used Esri ArcMap v.10.2 for visualization and analysis of GPS data, python v.3.7 for preparation of input samples and deep learning library keras v.2.2.4 to implement our proposed architecture. The model is trained on 100 epochs with efficient adam optimizer of stochastic gradient descent using the default configuration. For our experiments, we used Average Displacement Error (ADE) as the evaluation metric [15], which is the measure of euclidean distance between the actual trajectory location and predicted trajectory location for all the trajectories, as shown in equation 2.

$$ADE = \frac{\sum_{j=1}^{N} \frac{\sum_{i=1}^{n} (\hat{x}_i^j - x_i^j)^2 - (\hat{y}_i^j - y_i^j)^2}{n}}{N} \tag{2}$$

Where $(\hat{x}, \hat{y})$ are coordinates of predicted trajectories and (x, y) are coordinates of ground truth trajectories. n is the total number of points in a segment and N is the total number of segments. 80 percent of total data set is used for training purpose, out of which 10 percent is dedicated for validation purpose, Whereas the rest 20 percent used as the test set.

A. Comparison with state-of-the-art methods

To evaluate the effectiveness of our approach, this is also necessary to make comparison with different state-of-the-art approaches for sequence prediction. We compare our model with eight different benchmark models shown in Table I. It can be seen from Table I that our encoder decoder architecture performs better in terms of average displacement error as compared to all benchmark models for all level of grid resolutions from 2m to 10m.

The results in Table I shows that our method provides the stable results for different resolution of grids from 10m to 2m. It can be seen from Table I that LSTM Encoder/Decoder (E/D) gives better performance in terms of average displacement error among all other benchmark models. It can also be seen that performance of ConvLSTM encoder decoder has significantly improved as compared to LSTM encoder decoder architecture.

TABLE I
ADE LOSS COMPARISON OF METHODS AT DIFFERENT LEVEL OF GRID RESOLUTION (10 PTS)

Method	Grid Resolution/ADE			
	10m	5m	3m	2m
Linear Regression	9.1	12.18	16.47	22.79
Regression Tree	8.52	11.45	15.52	21.47
MLP	6.51	9.27	12.63	17.75
Simple RNN	4.17	7.39	9.96	14.36
CNN	3.31	4.41	5.88	7.72
GRU	2.86	3.38	5.46	7.11
Vanilla LSTM	2.73	3.57	4.88	7.07
LSTM (E/D)	2.38	3.04	3.43	5.28
ConvLSTM (E/D)	1.57	1.94	2.41	3.13
ConvLSTM (E/D) + Att	1.25	1.42	1.69	1.91
U-ConvLSTM (E/D)+Att+AuxGlobal	1.15	1.28	1.51	1.72
B-ConvLSTM (E/D)+Att+AuxGlobal	1.18	1.27	1.38	1.51

We further analyzed the impact of attention mechanism, it can be observed that ADE loss has further decreased using the attention mechanism in ConvLSTM encoder decoder architecture. We also analyzed the impact of auxiliary and global features (AuxGlobal) in our proposed architecture, the ADE loss is further decreased when AuxGlobal is incorporated with attention model. Bidirectional ConvLSTM encoder decoder using attention and AuxGlobal features is the final configuration of our architecture, the overall ADE loss has further decreased using the bidirectional encoder decoder architecture.

B. Evaluation of proposed Methods

We further investigate the performance of our final configuration of encoder decoder model to interpret different missing lengths of sub segments, and it can be seen that our approach has proved the acceptable performance. We evaluate the performance of our method for different lengths of missing points; i.e. from 10 GPS missing points to 30 missing GPS points. Table II shows the comparison of ADE loss performance of our proposed convolutional recurrent encoder decoder model to identify different lengths of missing parts.

TABLE II
ADE LOSS COMPARISON OF METHODS AT DIFFERENT LEVEL OF GRID RESOLUTION (10 PTS)

Missing Length (No. of points)	Grid Resolution			
	10m	5m	3m	2m
10 Pts	1.18	1.27	1.38	1.51
15 Pts	1.21	1.32	1.51	2.03
20 Pts	1.24	1.42	1.78	2.88
25 Pts	1.28	1.52	2.17	3.73
30 Pts	1.41	1.71	2.93	5.26

The results in Table II shows that the loss slightly increases with the increasing length of missing sub segments (top to bottom). It can also be seen that the value of ADE loss increases with the increased grid resolution (left to right), especially in high resolution grids like (2m grid). The overall loss can be reduced further by training the architecture on larger dataset, that is the key major factor in neural nets to optimize the parameters.

Conclusion: In this paper, we presented the study of missing trajectory completion on GPS trajectory dataset. We come across the unusual shortcomings in existing works and proposed the solution to solve the trajectory completion problem using state-of-the-art deep learning methodology. Our proposed methodology bidirectional convolutional recurrent decoder architecture shown significant improvement over the benchmark models in terms of average displacement error loss function. The results are further improved using attention mechanism and AuxGlobal features with our basic ConvLSTM based encoder decoder model. The results can be improved further, if trained on a bigger dataset in order to learn more complex and distinct patterns.

REFERENCES

[1] Alkan, R. M., Himmet Karaman, and M. Sahin. "GPS, GALILEO and GLONASS satellite navigation systems & GPS modernization." In Proceedings of 2nd International Conference on Recent Advances in Space Technologies, 2005. RAST 2005., pp. 390-394. IEEE, 2005.

[2] Liao, Binbing, et al. "Deep sequence learning with auxiliary information for traffic prediction." Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining. ACM, 2018.

[3] Zheng, Kai, et al. "Reducing uncertainty of low-sampling-rate trajectories." 2012 IEEE 28th International Conference on Data Engineering. IEEE, 2012.

[4] Li, Yang, et al. "Knowledge-based trajectory completion from sparse GPS samples." Proceedings of the 24th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems. ACM, 2016.

[5] Shen, Guicheng, et al. "An Area-Based Method for Missing Trajectory Completion: A STZ Algorithm." 2015 8th International Symposium on Computational Intelligence and Design (ISCID). Vol. 2. IEEE, 2015.

[6] Wu, Yonghui, et al. "Google's neural machine translation system: Bridging the gap between human and machine translation." arXiv preprint arXiv:1609.08144 (2016).

[7] Asadi, Ahmad, and Reza Safabakhsh. "The Encoder-Decoder Framework and Its Applications." Deep Learning: Concepts and Architectures. Springer, Cham, 2020. 133-167.

[8] Xingjian, S. H. I., et al. "Convolutional LSTM network: A machine learning approach for precipitation nowcasting." Advances in neural information processing systems. 2015.

[9] Mariescu-Istodor, Radu, and Pasi Fränti. "Grid-based method for GPS route analysis for retrieval." ACM Transactions on Spatial Algorithms and Systems (TSAS) 3.3 (2017): 8.

[10] Han, Xiaoguang, et al. "High-resolution shape completion using deep neural networks for global structure and local geometry inference." Proceedings of the IEEE International Conference on Computer Vision. 2017.

[11] Zheng, Yu, Xing Xie, and Wei-Ying Ma. "Geolife: A collaborative social networking service among user, location and trajectory." IEEE Data Eng. Bull. 33.2 (2010): 32-39.

[12] Bowen Du, Hao Peng, Senzhang Wang, Md. Zakirul Alam Bhuiyan, Lihong Wang, Qiran Gong, Lin Liu, Jing Li: "Deep Irregular Convolutional Residual LSTM for Urban Traffic Passenger Flows Prediction." IEEE Trans. Intell. Transp. Syst. 21(3): 972-985 (2020).

[13] Senzhang Wang, Jiannong Cao, Hao Chen, Hao Peng. "SeqST-GAN: Seq2Seq Generative Adversarial Nets for Multi-Step Urban Crowd Flow Prediction." To appear in ACM Transactions on Spatial Algorithms and Systems, 2020.

[14] Yuxuan Zhang, Senzhang Wang, Bing Chen, Jiannong Cao, and Zhiqiu Huang. "TrafficGAN: Network-Scale Deep Traffic Prediction with Generative Adversarial Nets." IEEE Trans. on Intelligent Transportation Systems, 2020.

[15] A. Sadeghian, Amir and Kosaraju, Vineet and Gupta, Agrim and Savarese, Silvio and Alahi, "TrajNet: Towards a Benchmark for Human Trajectory Prediction," arXiv Prepr., 2018.