AAAI

DeepETA: A Spatial-temporal Sequential Neural Network Model for Estimating Time of Arrival in Package Delivery System

Fan Wu, Lixia Wu

Artificial Intelligence Department, Zhejiang Cainiao Supply Chain Management Co., Ltd., Hangzhou, China

Background

- Over 100 million packages are delivered each day in China
- Importance of estimating time of arrival (ETA) of packages
  - reducing the anxiety of customers
  - improving the customer experience
  - measuring the service ability and quality of couriers

Delivering time

=

Arrival time - Depature time

Depature

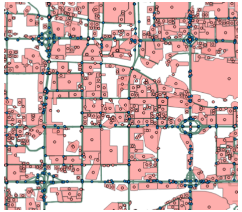
Arrival

Challenges

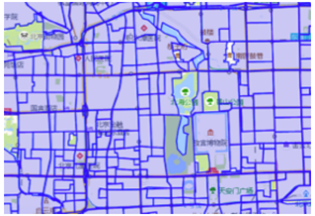
- Multiple destinations
  - All undelivered packages should be predicted at any time
  - The delivery time of different locations may vary due to the delivery sequence and the locations of the packages
- Time-variant delivery status
  - the sequence of the latest route, the regularity of the de-livery pattern and the sequence of packages to be delivered is hard to model
- Time-invariant delivery features
  - The geographical locations of the packages
  - the inherent properties of packages such as the weight or size

Problem Description

- Locations in delivery route
  - Packages are clustered into areas like communities or blocks
  - Utilizing an optimized partition method based on road network and areas of interest (AOI)



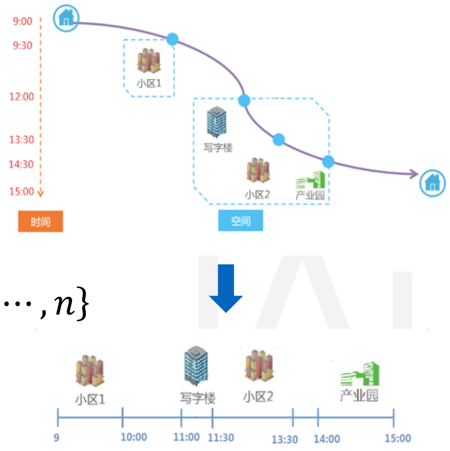
Road and AOI



Partition result

- Delivery route
  - Predicting the delivery time of a set of locations, namely $Set = \{a_i \mid i = 0, \cdots, n\}$
  - the delivery sequence at time t_i in day d

$$Route^{d}_{a_c,t_i} = \{node_{a_s,t_0} \rightarrow node_{a_1,t_1} \rightarrow \cdots \rightarrow node_{a_c,t_i}\}$$

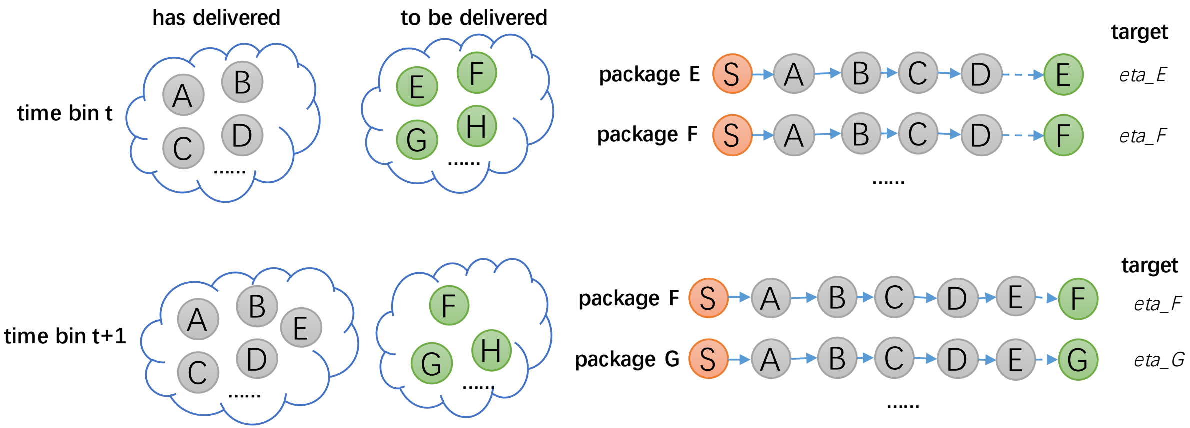


Delivery route

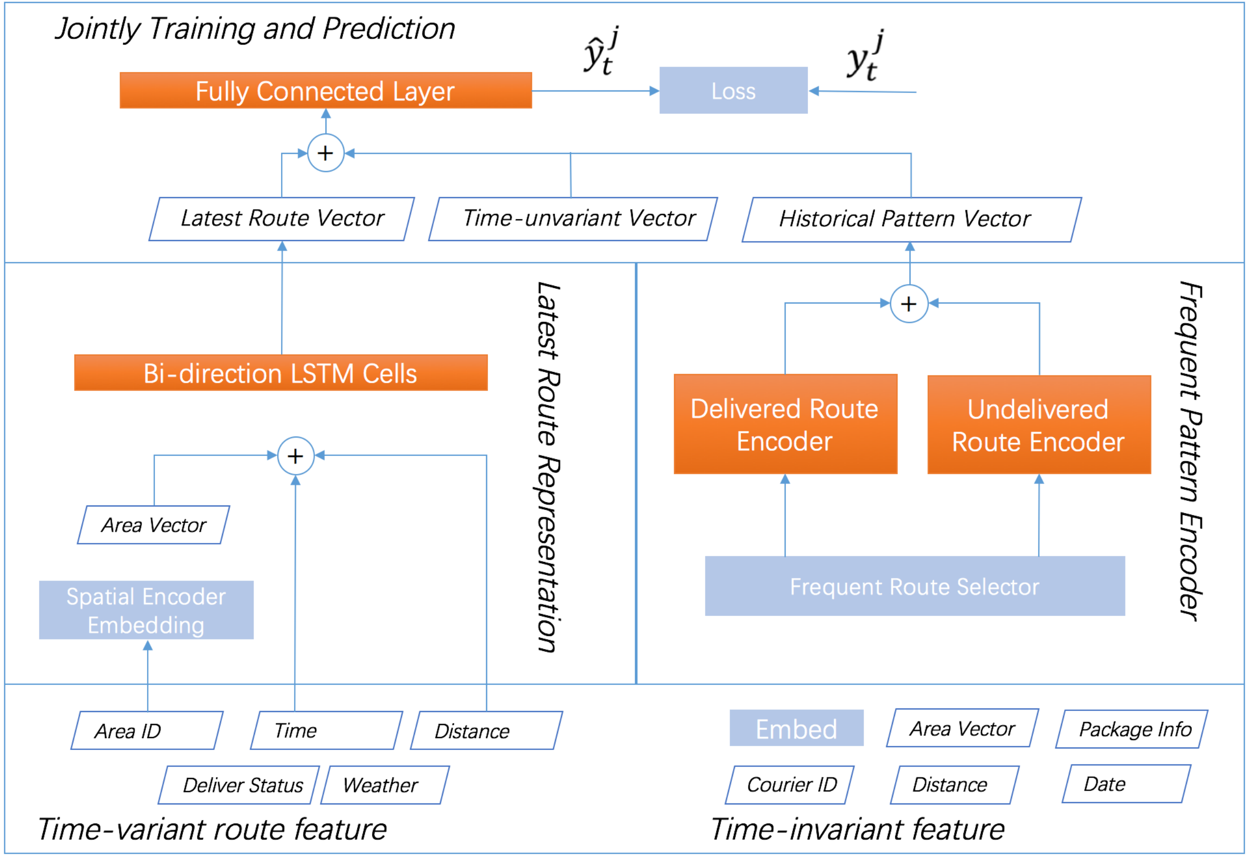
Problem Description

- Problem definition
  - The latest route $Route^{d}_{a_s,a_c}$ and the predicted location a_i
  - Relative routes from history that are similar with the current route symbolized as $\mathcal{H}_{route}$
  - Undelivered location set $Set^{d}_{t_i,}$ and all similar sets from history marked as $\mathcal{F}_{set}$
  - Find the most possible delivery time dt_{a_i} from historical relative routes:

$$dt_{a_i} = f(\mathcal{H}_{route}, \mathcal{F}_{set})$$



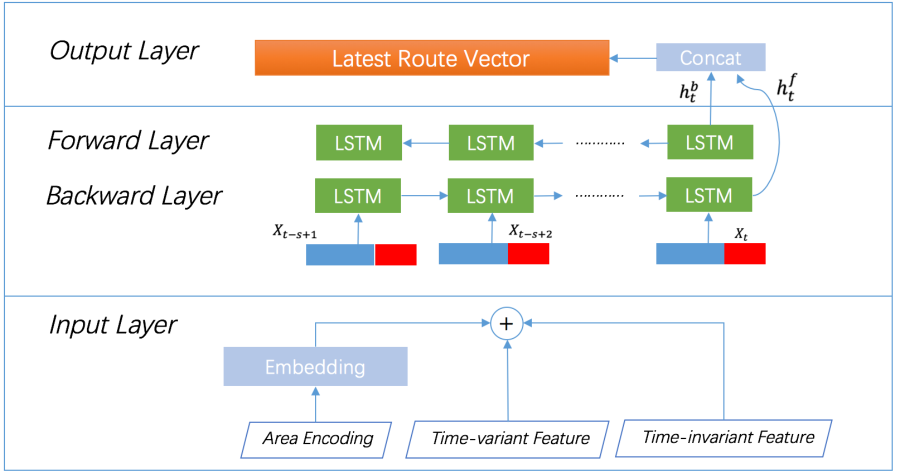
Proposed DeepETA Framework



- End to end structure
- Latest route encoder
- The frequent pattern encoder
- The prediction module

Model Structure

- Latest Route Encoder
  - To capture the complicated sequential information that influences the delivery time
  - Consists of the spatial encoder and the BiLSTM cells



- Spatial encoder
  - Location a_i is transformed into Geohash encoding of 40 digits, namely G_{a_i}

$$V_{a_i} = f\left(W_a G_{a_i} + b_a\right)$$

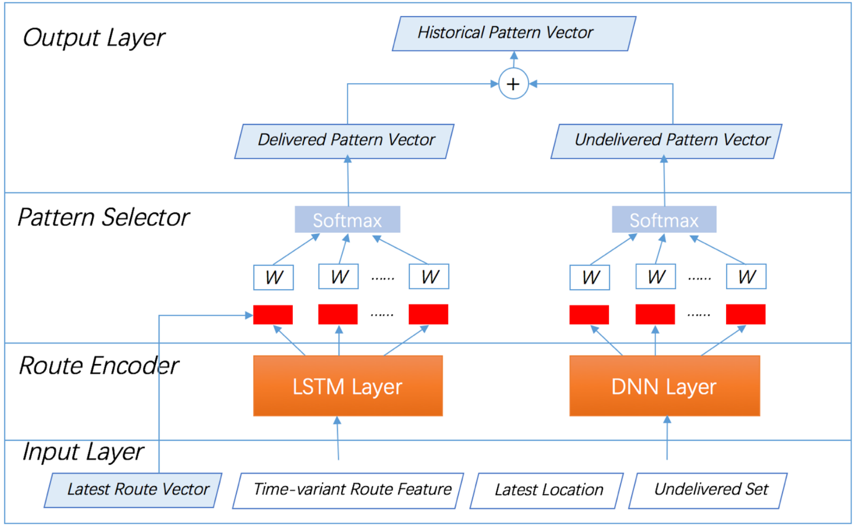
- BiLSTM
  - Spatial encoder V_{a_i}, time-variant features V_{tv} and time-invariant features V_{ti} as $X_t = [V_{a_i}, V_{tv}, V_{ti}]$

$$h_t = LSTM(h_{t-1}, X_t)$$

  - Combine the back and forward vectors $h_t = [h_t^b, h_t^f]$

Model Structure

- Frequent Pattern Encoder
  - To capture the frequent mobility patterns
  - Consists of the route encoder and the pattern selector



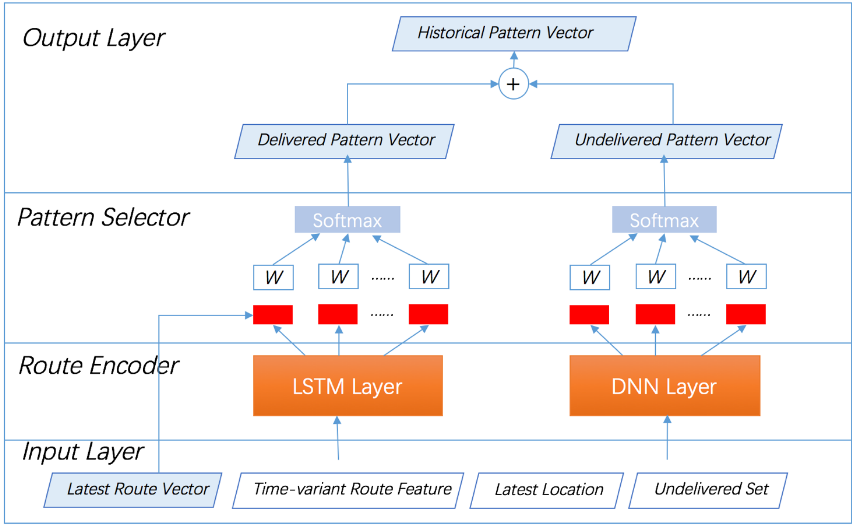
- Route encoder
  - Extracts only the relative frequent patterns
  - Delivered route pattern

$$\tilde{\mathcal{H}}_t^k = \mathrm{LSTM}(\tilde{\mathcal{H}}_{t-1}^k, \mathcal{H}_{route})$$

  - To-be-delivered pattern

$$\tilde{\mathcal{F}}_t^k = DNN(\tilde{\mathcal{F}}_{t-1}^k, \mathcal{F}_{set})$$

Model Structure



- Pattern selector
  - To find which frequent pattern has the biggest impact on the current situation
  - An attention-based layer
    - the frequent pattern vectors are combined with the latest route vector through a score function

$$f\left(\widetilde{\mathcal{H}}_t^k, h_t\right) = tanh\left(\widetilde{\mathcal{H}}_t^k W_{score} h_t\right)$$

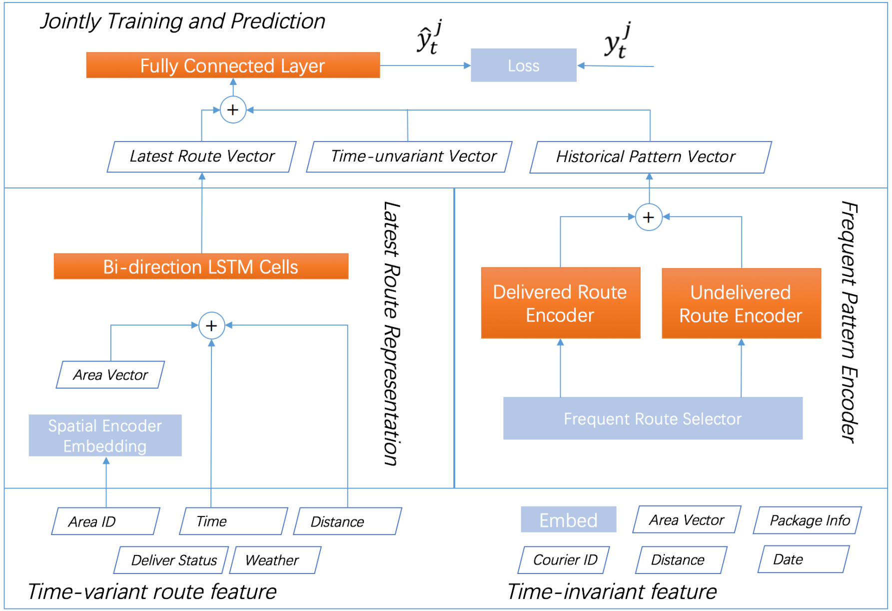
    - All scored vectors are exported to the softmax layer to calculate the weight of each vector

$$\sigma(z)_k = \frac{e^{z_k}}{\sum_{i=1}^{K} e^{z_i}}$$

    - The value of the softmax is multiply with the historical pattern vectors

$$V_{fp} = \sum\left(\sigma(z)_j \widetilde{\mathcal{H}}_t^j\right)$$

Model Structure



- Jointly Training and Prediction
  - The latest route vector h_t, the frequent pattern encoder V_{fp}, and the time-invariant features V_{ti}, as $\tilde{X}_t = [h_t, V_{fp}, V_{ti}]$
  - The fully connected layer

$$\tilde{y}_t = \sigma\left(W_{fc}\tilde{X}_t + b_{fc}\right)$$

- Loss function
  - The mean square error (MSE): *measurement of bias and variance*
  - The mean absolute percentage square error (MAPSE): *not sensitive to outliers*

$$\mathcal{L}(\theta) = \sum_{i=1}^{N}\left((\tilde{y}_t - y_t)^2 + \lambda\left(\frac{\tilde{y}_t - y_t}{y_t}\right)^2\right)$$

- Dataset Description
  - Real-world package delivery dataset collected by Cainiao Ltd.
    - The delivery routes of 331 couriers from Jun. 1, 2018 to Aug. 1, 2018 (60 days), in Beijing, China
    - The previous 50 days are used as training set and the last 10 days as testing set

Pkgs	Route	Nodes	Time	AvgDt
80	2	20	8h	3.5h

- Evaluation Metric
  - Mean Average Percentage Error (MAPE)

$$MAPE = \frac{1}{N}\sum_{i=1}^{N}\frac{\left|\hat{y}_t^j - y_t^j\right|}{y_t^j}$$

  - Rooted Mean Square Error (RMSE)

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\left(\hat{y}_t^j - y_t^j\right)^2}$$

Experiment

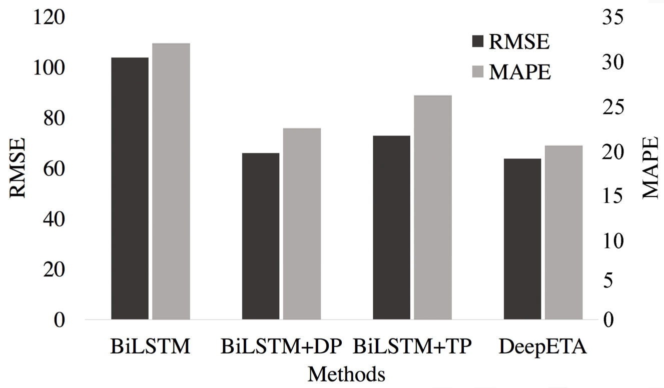
- Methods for Comparison
  - Linear regression (LR)
  - XGBoost
  - Deep neural network (DNN)
  - LSTM
  - DeepTTE
  - DeepMove
  - Latest route representation (BiLSTM)
  - Frequent pattern encoder (BiLSTM+DP and BiLSTM +TP)

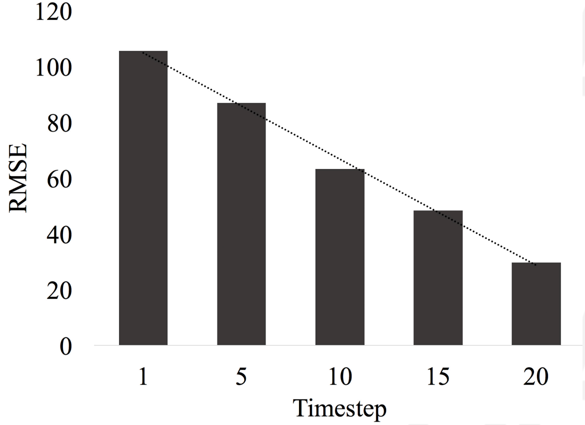
Methods	RMSE (min)	MAPE (%)
LR	144.18	43.5
DNN	127.58	37.4
XGBoost	123.66	38.2
LSTM	110.37	34.9
DeepTTE	97.85	29.7
DeepMove	72.43	24.3
DeepETA	63.58	20.6

- Model Comparison
  - DeepETA achieves the lowest RMSE (63.58 minutes) and the lowest MAPE which improves the best performance of the baseline methods by 13.8% (RMSE) and 16.5% (MAPE)
    - Attention mechanism to enhance the ability of learning long sequences
    - LSTM-based layer to extract the sequential features
    - DNN-based layer to model undelivered sets

Experiment

- Effectiveness of model components
  - An improving of 6% purely rely on the BiLSTM layer
  - The delivered pattern encoder improves 36.5%
  - The to-be-delivered pattern encoder improves 29.8%
  - The combination of all modules improves 3%
- Performance of prediction at different time
  - Re-predicting all the undelivered packages at any time
  - RMSE decrease significantly as the increasing of time step
  - More current route information can be inferred and the scope of undelivered sets can be reduced





Summary

- Propose a deep spatial-temporal sequential model for estimating the package delivery time (DeepETA)
  - The latest route encoder embeds the location of packages that remain geographical relations and BiLSTM is used to model the sequential features
  - The frequent pattern encoder selects the frequent routes from historical data and uses LSTM and DNN to represent the routes
  - An attention-based layer is developed to calculated the most similar patterns with the current route
  - Jointly training is utilized to minimize the loss function
- Experiments on real logistics dataset show that the proposed method overwhelms the start-of-art methods and the effectiveness of three modules are illustrated

CAINIAO 菜鸟 AAAI

THANKS

Fan Wu
wf118503@cainiao.com

In the near future, codes will be released in: https://github.com/wufandm/deepeta