TRAFFIC FLOW PREDICTION

Third report

University of Piraeus
Efstratios Karkanis

Table of Contents

1. Introduction

Nowadays, traffic flow prediction is a major problem. Traffic flow prediction can be divided in **short term** (prediction in the next minutes or hours) and **long term** (next day prediction).

It is notable that most of the studies being analyzed use LSTM models (a specific type of RNN), to predict the traffic flow of a car, as this architecture is an efficient method and gives good results.

There are also papers that transform the traffic data into a graph and use whole graphs to predict the future traffic flow. This can be done using Graph Neural Networks (GNNs). By this way, the connectivity of road segments can be modeled as a graph.

2. Related Work Analysis

In this chapter, we analyze papers that contribute to traffic flow prediction domain. From each paper we track information about the problem being solved and the model being used, as well as the input and the output of that model. Finally, we keep information of the dataset that they used in each research.

Paper 1: Deep Crowd Model

Paper's title: DeepCrowd - A Deep Model for Large-Scale Citywide Crowd Density and Flow Prediction [1]

What is the problem being solved?

The problem that this research tries to solve is traffic prediction. In other words, they predict not only how many objects will be in each location at the next timestamp (they call it crowd density prediction), but also how many objects will leave and visit each location at the next timestamp (they call it crowd in-out flow prediction).

How the road network is being modeled?

They split a large urban area to a number of mesh-grids. Then, citywide crowd density in a continuous period can be represented with a four-dimensional tensor R^Timestep×Height×Width×Channel in an analogous manner to video data, where each Timestep can be seen as one video frame, Height and Width is two-dimensional index for mesh-grids, and each Channel stores an aggregated scalar value for each mesh-grid.

What is the model being used?

The model that they use is called DeepCrowd. This model consists of convolutional LSTMs in the form of pyramid. Convolutional LSTMs handles high dimensional data, while the pyramid architecture better utilizes low- and high-resolution feature maps.

What is the input and the output to the model?

The input to the model is multiple historical spatial domains of the road network. The output is the next traffic flow prediction result. Temporally, the next step of citywide crowd density and in-out flow can be correlated with historical data.

What is the dataset being used?

In this research, the dataset being used reflects the real-world citywide data. The data are of high quality (cover a large spatial area, many user samples, use of finer mesh-sizes). Also, this data is collected by GPS using a popular smartphone application. Using this data, researchers convert them into 4D tensors: R^Timestep,Heigh,Width,Channel. Moreover, data are from two big cities: Tokyo and Osaka. (to get this data, an application to the owner is required)

GitHub Link: https://github.com/deepkashiwa20/DeepCrowd

Paper 2: Diffusion Convolutional Recurrent Neural Network (DCRNN)

Paper's title: Diffusion convolutional recurrent neural network: data-driven traffic forecasting [2]

What is the problem being solved?

The problem being solved is traffic forecasting on road networks. Actually, they predict the future traffic speeds of a sensor network, given historical data (previously observed traffic flow) from N correlated sensors. In this study, the data are collected using sensors, which are placed in many places in the road network, and thus creating a sensor network.

How the road network is being modeled?

They represent the sensor network as a weighted directed graph $G = (V, E, W)$, where V is a set of nodes $|V| = N$, E is a set of edges and $W \in R^{N \times N}$ is a weighted adjacency matrix representing the nodes proximity (e.g., a function of their road network distance).

In particular, nodes are sensors and edge weights denote proximity between the sensor pairs measured by the road network distance.

What is the model being used?

The model being used is called Diffusion Convolutional Recurrent Neural Network (DCRNN). The structure of the model is based in the Seq2Seq architecture. Seq2Seq architecture consists of an encoder and a decoder. Both are Recurrent Neural Networks with DCGRUs (Diffusion Convolutional Gated Recurrent Units). Like GRU, DCGRU can be used to build recurrent neural network layers and be trained using backpropagation through time.

During training they feed the historical time series into the encoder and use its final states to initialize the decoder. The decoder generates predictions given previous ground truth observations.

What is the input and the output to the model?

Researchers denote the traffic flow observed on G as a **graph signal $X \in R^{N \times P}$**, where P is the number of features of each node (e.g., velocity, volume). Let $X(t)$ represent the graph signal observed at time t, the traffic forecasting problem aims to learn a function $h(\cdot)$ that maps T' historical graph signals to future T graph signals, given a graph G: $[X(t-T'+1), \cdots, X(t); G] \longrightarrow [X(t+1), \cdots, X(t+T)$.

What is the dataset being used?

The experiments are conducted on two big datasets: **METR-LA** and **PEMS-BAY**. The first dataset (METR-LA) contains traffic information collected from loop detectors in the highway of Los Angeles County. They select 207 sensors and collect 4 months of data ranging from Mar 1st, 2012 to Jun 30th 2012 for the experiment.

The second dataset (PEMS-BAY) is collected by California Transportation Agencies (CalTrans) Performance Measurement System (PeMS). Scientists select 325 sensors in the Bay Area and collect 6 months of data ranging from Jan 1st, 2017 to May 31th 2017 for the experiment.

In both datasets, 70% of data is used for training, 20% are used for testing while the remaining 10% for validation.

GitHub Link: https://github.com/liyaguang/DCRNN

Paper 3: T-LSTM Model

Paper's title: T-LSTM: a long short term memory network enhanced by temporal information for traffic flow prediction [3]

What is the problem being solved?

This paper uses the temporal information (time information) to predict traffic flow on a single road section. In other words, they use features associated with time to predict traffic flow. The prediction is short-term. Researchers believe that traffic flow is not the same in each time period of the day – traffic flow is high in rush hours and low in the early hours of each day.

What is the model being used?

The model being used is called Temporal information enhancing Long Short-Term Memory neural network (T-LSTM). This model combines recurrent time labels with recurrent neural networks, which makes the best use of the temporal features to improve the accuracy of short-term traffic flow prediction.

What is the input and the output to the model?

The input to the model is a vector x = (x1,x2,x3,…,xn), where each xi contains a time label and the traffic flow information. The input to the LSTM model is a 8x2 matrix (8 historical data are used to predict traffic flow of the next moment). The output of the model is a 2D vector y = (time label, traffic flow) of the next moment prediction.

What is the dataset being used?

The traffic detector data from Shibalidian Bridge to Hongyan Bridge of the East Fourth Ring Road in Beijing (March 1 to August 30, 2014) is selected to validate the T-LSTM. This dataset has the following columns: Speed, Flow, Date, Density. Also, there is not a link to this dataset.

For purpose of research and analysis, the data are aggregated into the time intervals of 16 minutes. So, there are 90 (720×2÷16) pieces of data and corresponding time labels for each day. The data of the first five months are used for training and the data of August are used for testing. Finally, the data are normalized to [0, 1] by the Min-Max Scaler normalization method in the Scikit-learn library.

Paper 4: Hybrid LSTM

Paper's title: Hybrid LSTM Neural Network for Short-Term Traffic Flow Prediction [4]

What is the problem being solved?

This paper tries to predict the short-term traffic flow of cars in a real road network. Traffic flow here means the expected number of cars that will flow in the road network.

What is the model being used?

The model is a Hybrid LSTM model, which is based on the LSTM model. Scientists did not use the simple LSTM model here as it does not provide good accuracy results on traffic flow prediction. The Hybrid LSTM model consists of dense, dropout, LSTM, and activation layers. The **Dropout layer** means that some neural units are temporarily discarded from the neural network according to a certain probability during the training process, preventing over-fitting. The **Activation layer** is mainly used to improve the learning ability of the model.

What is the input and the output to the model?

Training data (70%) of the format X = [Current Node, Source Node, Traffic Flow]. Output Y is the traffic flow prediction on each road network of the road.

What is the dataset being used?

The experimental data set is the local road network traffic data from September 1st to September 8th in Yunyan District, Guiyang City, Guizhou Province. Inside the road network there are traffic lights. In this research, people tried to measure the flow between traffic lights every 30 seconds. The data collected are of the following format (have the following columns):

Current Node (TrafficLightID, Source Node (FromID), Traffic Flow (traffic_flow)

TrafficLightID denotes the traffic light identifier of the vehicle currently, while FromID denotes the traffic light identifier of the vehicle at the last moment. Because the experimental data of each day is collected at a 30s interval from 6:00 a.m. to 8:00 p.m., the traffic_flow field is a matrix of dimension 1680, which is used to represent the traffic flows of a road section at a 30s interval. For example, traffic_flow[0] represents the traffic flow from road intersection FromID to road intersection TrafficLightID at t0.

In this paper, they try to transform time series into supervised learning problem and evaluate the accuracy of model prediction. So, they divided the traffic flow data set into **training set** and **test set**, where the first 70% of the traffic flow data set is the training set and the last 30% of the traffic flow data set is the test set. After training a model based on the training set, the use of RMSE error metric will be used to evaluate the prediction accuracy.

3. Strict Path Queries (SQPs)

3.1 About the research

Strict Path Queries (SPQs) try to detect every vehicle trajectory that follows exactly a path π within a specific time-period, without making detours. In order to understand the idea of this paper, the following terms are defined:

Edge (e): the line that connects two intersections.

Segment (s): a piece of the edge between any two consecutive points that make it up, without necessarily being intersections

Route (r): number of different edges of the traffic network.

Trajectory (t): the set of edges that a car (moving object) crossed in the road network.

Set (T): a number of vehicle trajectories that are stored in a database.

Path (π): a set of edges

tstart: time when the moving object enters a specific edge of the road network.

tend: time when the moving object leaves a specific edge of the road network.

Consequently, the result of an SPQ includes every trajectory t from the set T that has path π as its SUB-PATH, enters π at time tstart or later and exits π at time tend or earlier.

3.2 Technical details about data organization

The traffic network is essentially a directed graph G = (V,E). Vehicle movement data is collected by a GPS system and includes points with a clear position in three-dimensional space (x, y, t), where x is longitude, y is latitude and t is time. A moving object (vehicle) reports its position via a GPS device at times that usually follow a fixed periodicity. Each position reference is a tuple of parameters of the type **loc=(moid, ts, pos)**, where moid is the identifier of the moving object, ts is a specific time and pos is the position of the object at the time ts in the form of a spatial specification (coordinates).

With appropriate processing the initial positioning points **are mapped to a sequence of network edges,** thereby defining a trajectory within the network. Each record belonging to this sequence is a tuple of parameters **locmm=(tid, eid, tsenter, tsleave).** Therefore, the initial data collected from GPS is converted into a locmm format. A trajectory t is a set of such records: t =[locmm1, locmm2, …, locmmn].

4 Bibliography

[1] R. Jiang *et al.*, “DeepCrowd: A Deep Model for Large-Scale Citywide Crowd Density and Flow Prediction,” *IEEE Trans. Knowl. Data Eng.*, vol. 4347, no. c, pp. 1–15, 2021, doi: 10.1109/TKDE.2021.3077056.

[2] Y. Li, R. Yu, C. Shahabi, and Y. Liu, “Diffusion convolutional recurrent neural network: Data-driven traffic forecasting,” *6th Int. Conf. Learn. Represent. ICLR 2018 - Conf. Track Proc.*, pp. 1–16, 2018.

[3] L. Mou, P. Zhao, H. Xie, and Y. Chen, “T-LSTM: A long short-term memory neural network enhanced by temporal information for traffic flow prediction,” *IEEE Access*, vol. 7, pp. 98053–98060, 2019, doi: 10.1109/ACCESS.2019.2929692.

[4] Y. Xiao and Y. Yin, “Hybrid LSTM neural network for short-term traffic flow prediction,” *Inf.*, vol. 10, no. 3, 2019, doi: 10.3390/info10030105.