Challenges

Given historical traffic flow and the road network, predict the future traffic speed.

Diffusion Convolutional Recurrent Neural Network (DCRNN)

Diffusion convolution on the directed graph. Sequence to sequence learning with scheduled sampling for long term forecasting. Code and data: https://github.com/liyaguang/DCRNN

Traffic Forecasting Problem

Model correlation between traffic sensors as a directed graph $G = (V, W)$. $W_{ij} = \exp\left(\frac{\text{dist}(v_i, v_j)}{\alpha}\right)$ if $\text{dist}(v_i, v_j) \leq \alpha$, otherwise 0 $X^{t'} \in \mathbb{R}^{n \times P}$, observed traffic flow on $G$ at time $t$. Problem Statement: Learn a function $g(\cdot)$ that maps $T$ historical observations to future $T$ observations. $X^{t}, \ldots, X^{t'}, Y^{t}, \ldots, Y^{t+T} \rightarrow X^{t'}, \ldots, X^{t+T}$

Spatial Dependency Modeling

Diffusion Convolution

Model the traffic flow as a diffusion process, e.g., diffusion of vehicles. Generalize convolutional operation to graph by defining the filter $f_{\theta}$ as a combination of different diffusion steps.

$X_{p} = \sum_{k=0}^{K} \theta_{1,k} D_{t,k} W_{t,k}^{\top} X_{p}$ for $p \in \{1, \ldots, P\}$

$\star_{\alpha}$: diffusion convolution, $D_{t,k}$: diagonal out-degree matrix, $W_{t,k}$: diagonal in-degree matrix.

Efficient: time complexity $O(K|E|) \ll O(|V|^2)$

Expressive: many convolutional operations defined on graph, e.g., ChebNet [2], can be represented using diffusion convolution.

Filter Visualization

An example of the learned convolutional filter with weights localized around the center, and diffuse alongside the road network.

Temporal Dependency Modeling

Diffusion Convolutional Gated Recurrent Unit (DCGRU)

Sequence to Sequence Framework with Scheduled Sampling

Scheduled Sampling [1]: Curriculum learning method to mitigate error propagation.

Diffusion Convolutional Recurrent Neural Network

Model spatial dependency with bidirectional diffusion convolution on the graph. Model temporal dependency with the sequence to sequence framework.

Experiments

Data statistics

METR-LA: 267 traffic sensors in Los Angeles, 4 months in 2012, 6.5M observations

PEMS-BAY: 345 traffic sensors in Bay Area, 6 months in 2017, 17M observations

Baselines

Historical Average (HA); Auto-regressive Integrated Moving Average with Kalman filter (ARIMA); Vector Auto-Regression (VAR); Support Vector Regression (SVR)

Deep Feed forward Neural Network (FNN); Fully-connected LSTM with sequence to sequence framework (FC-LSTM)

Forecasting performance

DCRNN achieves the best performance for all forecasting horizons for both datasets.

Ablation Study

Removing either spatial or temporal modeling results in significantly worse results.

Performance vs number of filters and number of diffusion steps.

Example Forecasting Results

DCRNN is more likely to accurately predict abrupt changes in the traffic speed.

References


Scheduled sampling for sequence prediction with recurrent neural networks.


Convolutional neural networks on graphs with fast localized spectral filtering.


Sequence to sequence learning with neural networks.