RiskOracle: A Minute-level Citywide Traffic Accident Forecasting Framework Paper ID: 1664 Supplementary Materials

We present several supporting details for our paper from the perspectives of main contributions and experimental details.

Main Contributions

The main contributions of our framework RiskOracle can be summarized as follows.

- We improve the temporal granularity of real-time accident forecasting from hour levels to minute levels.
- We propose the Multi-task DTGN in our framework to tackle the challenges of extremely short-term forecasting. To the best of our knowledge, it is the first time to address the accident forecasting in the graph convolution network.

  (a) Regions are geographically distant but have the potential correlations due to the similar road structures or tidal flows can be directly connected in a dynamic way.

  (b) Propagations and interactions of abnormal events can be captured with the differential component incorporated and further the relationship between accidents and abnormal variations are learned in this model.

  (c) The multi-task scheme is designed to address the sporadic and heterogenous accident distribution in forecasting. The multi-scale accident distributions can be learned by two tasks and further utilized for highlighting most-likely accident subregions.

  All these strategies are proposed to capture the minute-level variation of dynamic traffic for accident predictions and can be extended to other short-term spatiotemporal forecasting.
- We propose two strategies to deal with the challenging dataset:

  (a) We propose a data enhancement strategy to deal with the zero-inflated issue in extremely rare nonzero items in labels in training dataset.

  (b) We propose a co-sensing strategy (ST-DFM) to deal with sparse sensing data in urban computing by training the model on the intersections of the real-time datasets with intensive spatial attributes involved.

  With the two strategies, our framework can be extended to other sporadic and sparse spatiotemporal forecasting tasks with incomplete data such as drug abuse forecasting, disease distribution inference as well as crime prediction.

Hyper-parameter Study Details

To illustrate the experimental process on hyper-parameter studies, here we show the performances of our model with different parameter setings on 30-minute level. We adjust the number of layers and filters in each layer to make itself reach their best performance at 9 layers with 384 filters. For multi-task learning, we fix the weight of main task as 1, and adjust λ_1, λ_2 as Table 3 shows. Also, we adjust the weight of the dynamic element in overall affinity and reach the best performance when γ equals 0.5 among $\{0, 0.5, 1.0, 1.2, 1.5\}$. And q equals 18 among $\{9, 18, 33\}$ when the MSE of the second auxiliary task reaches the lowest at 0.806.

Number of layers and filters

We show the performance varying with the number of DTGN layers and filters in each layer in Figure 1.

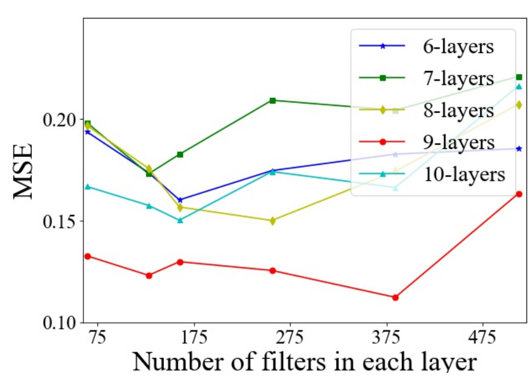

Figure 1: Performance on different number of layers and filters

Table 1: Overall affinity

Proportion of dynamic affinity	Acc@K	Mse of accident risks
0	32.88	0.3253
0.5	**51.94**	**0.2417**
1	46.88	0.2849
1.5	43.31	0.3124

Table 2: The performance on different q

q	Acc@K	Mse of the 2nd Auxiliary task
9	42.88	0.942
18	**51.32**	**0.806**
33	34.19	1.120

The proportion of dynamic affinity accounts for overall affinity

We show the proportion of dynamic affinity accounts for time-varying overall affinity in Table 1. As can be observed in this table, dynamic affinity indeed contributes to the improvement of our main task and the dynamic affinity accounts for 0.5 when it arrives the highest accuracy.

Numbers of different medium-sized rectangles

Here, we visualize the performance of our framework on different q in Table 2.

As can be observed, the the MSE of the second auxiliary task is fairly low to meet the requirements for accident selection and conform the framework to the changes of time and weather.

Effects of weights on different auxiliary tasks

For multi-task learning, we fix the weight of main task as 1, and adjust λ_1, λ_2 as Table 3 shows. As can be observed, the second auxiliary task (λ_2) which predicts coarse-grained accident distributions, plays a pivotal role in fine-grained accident forecasting. The hierarchical accident forecasting consists of the output of the main task and auxiliary task. And due to the high correlations of traffic volumes and accident risks, the λ_1 is adjusted to 0.8 to fit the main task.

Table 3: Effects of λ

λ_1	λ_2	MSE	Acc@K
0.5	1	0.2105	43.64%
0.8	**1**	**0.2275**	**53.82%**
1	1	0.2120	46.59%
1	1.2	0.2335	39.21%
1	1.5	0.2332	36.34%