Semi-Supervised Hierarchical Recurrent Graph Neural Network for City-Wide Parking Availability Prediction - Appendix

Paper ID 6479

Effect of β

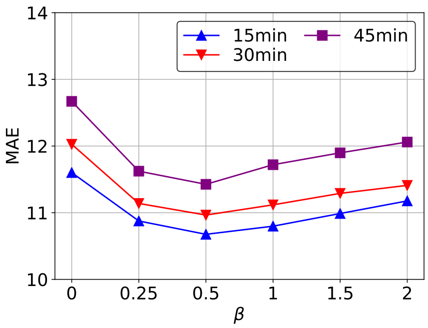


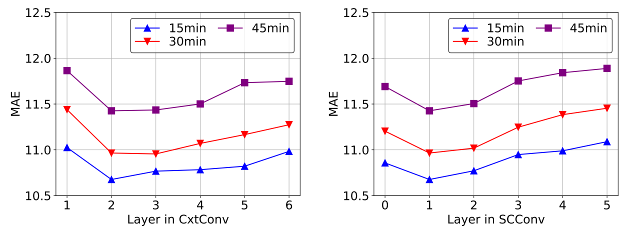
Figure 1: Effect of β on BEIJING

β is the hyper-parameter that controls the importance of two CE losses. Figure 1 reports the effect of β on BEIJING. We can make the following observations. (1) By adding CE loss can remarkably improve the prediction accuracy. (2) SHARE achieves the best accuracy when β is 0.5, and the accuracy degrades when we further decrease or increase β. Based on the above observations, we choose $\beta = 0.5$ in the overall experiment.

Effect of graph parameters

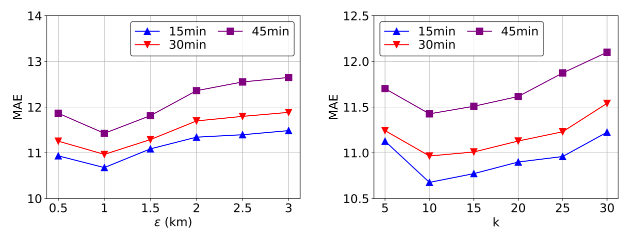
In this section, we report the effect of graph parameters on BEIJING, including the number of GNN layers in CxtConv, the number of GNN layers in SCConv, the distance threshold ϵ, and the top-k nearest parking lots in graph construction. Each time we vary a parameter, we set others to their default values. Note that stacking multiple layers in PropConv will induce label confusion, therefore we only use one GNN layer in PropConv.

First, we vary the number of GNN layers in CxtConv from 1 to 6. The results are reported in Figure 2(a). As can be seen, by setting two GNN layers in CxtConv achieves the best performance. Further decrease or increase GNN layers lead to performance degradation. This is because too few GNN layers can not aggregate sufficient information, whereas too many GNN layers make the model losses discriminative power. Therefore, we use two GNN layers in the overall experiment.



(a) # of of layers in CxtConv (b) # of of layers in SCConv

Figure 2: Effect of graph parameters on BEIJING



(a) Effect of ϵ (b) Effect of k

Figure 3: Effect of graph parameters on BEIJING

Similarly, we vary the number of GNN layers in SCConv from 0 to 5. Note that zero layer denotes not use GNN (i.e., update operation in Equation (8)) in SCConv. The results are reported in Figure 2(b). Different from CxtConv, use only one GNN layer is enough for SCConv to capture distinctive features. Further stacking GNN layers leads to performance degradation. This because the graph in SCConv is more densely connected (according to Equation (9)). Therefore, we use one GNN layers in the overall experiment.

After that, we vary the distance threshold ϵ from 0.5 to 3. The results are reported in Figure 3(a). SHARE achieves the best performance when $\epsilon = 1$Km. This makes sense since too few neighbors limit the information propagation, whereas too many neighbors introduce extra noises in the information propagation process.

Finally, we vary the parameter k when connecting parking lots from 5 to 30. The results are reported in Figure 3(b). Overall, SHARE achieves the best performance when $k = 10$. We observe consistent performance degradation when we decrease or increase k. The reason is similar to the effect of ϵ.