

This is a note for the paper: Gated Graph Sequence Neural Networks (ICLR 2016)

对原始GNN的分析

原始GNN在图上进行局部迭代运算时，由于要求局部迭代运算为压缩映射，因此一个节点的特征在传播时随着距离的增加会出现严重的传递衰减，这就导致远处的节点无法接收到当前节点的信息。即，压缩映射虽然保证了收敛但是导致信息传递的损失。

加门控的GNN

为了避免这种信息损失，不再进行迭代运算直到收敛，而是迭代固定次数。节点的新特征由两部分构成：当前节点的特征 $h(t)$ ，周围节点对当前节点的特征 $h_n(t)$ 。周围节点对当前节点的特征计算方法为：线性变换后再相加。其中，这个线性变换对所有节点共享。将 $h_n(t)$ 作为GRU输入， $h(t)$ 作为GRU状态，送入GRU单元，GRU单元全局共享，得到 $h(t+1)$ 。论文中 h_n 被称为activations。如果是多标签的边，那么每种标签建一张图，将所有图的特征相加作为总的特征，其中每张图上的运算如前所述，参数在不同图之间不共享。

输出由状态通过一个NN得到。可以每一时刻都有输出，就得到了一个sequence。

作者还引入了一个annotation状态，是有状态 h 得到的。之所以引入annotation是因为现实中真的可能有annotation，所以这是一个可选择的输入项。

实验

实验部分没太看懂