

Abstract—Graph convolutional networks (GCNs) have recently become one of the most powerful tools for graph analytics tasks in numerous applications, ranging from social networks and natural language processing to bioinformatics and chemoinformatics, thanks to their ability to capture the complex relationships between concepts. At present, the vast majority of GCNs use a neighborhood aggregation framework to learn a continuous and compact vector, then performing a pooling operation to generalize graph embedding for the classification task. These approaches have two disadvantages in the graph classification task: （现有Attention两问题）

(1) when only the largest sub-graph structure (k-hop neighbor) is used for neighborhood aggregation, a large amount of early-stage information is lost during the graph convolution step; （早期信息丢失）

(2) simple average/sum pooling or max pooling utilized, which loses the characteristics of each node and the topology between nodes. （简单Pooling丢拓扑）

【新架构】dual attention graph convolutional networks (DAGCN) to address these problems.

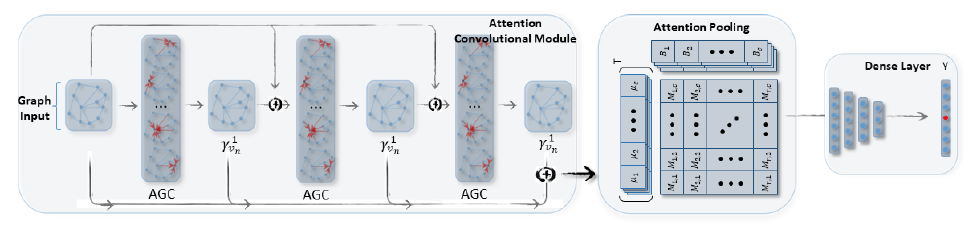


Fig. 1. The architecture of the dual attention graph convolution network (DAGCN).

The model consists of three parts:

(1) The left tier is the attention graph convolution module with three AGC layers (m = 3) which learns the hierarchical local substructure features by aggregating the hops of its neighbors.

(2) The middle part is the attention pooling layer, the matrix B is the attention coefficient matrix.

(3) The final graph embedding matrix M is then sent to a dense layer for final predictions.

【创新1 自动学邻点重要度】DAGCN automatically learns the importance of neighbors at different hops using a novel attention graph convolution layer, and then

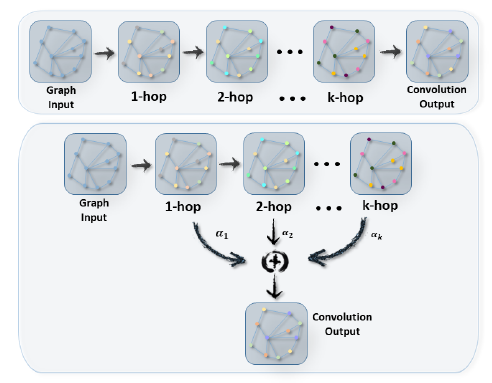
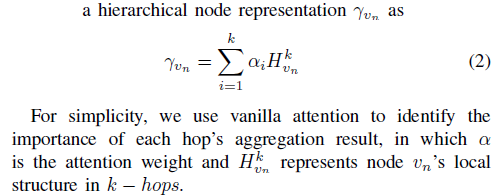
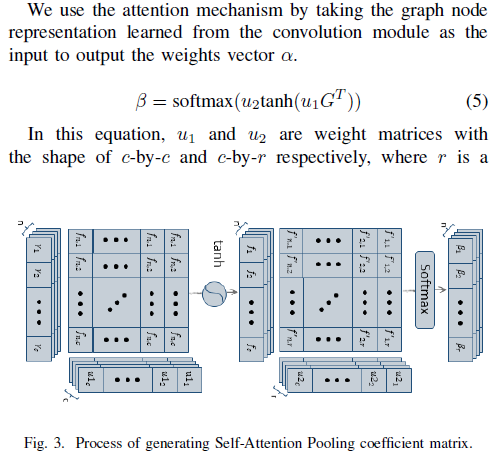


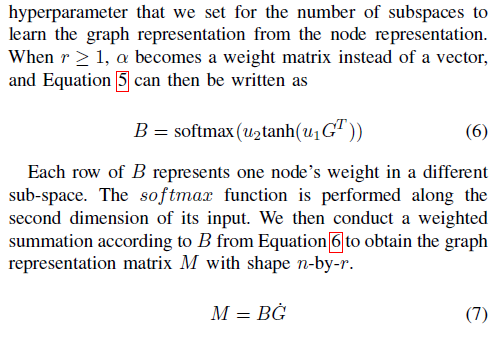
Fig. 2. Traditional Graph Convolution Layer (up): Only the final output which contains the largest sub-structure (k-hop neighbor substructure) is used.

Attention Graph Convolution Layer (down): valuable information is extracted from every convolution step to generate a hierarchical node representation



【创新2 selfattention pooling 通用化图表示，学出不同表示的权重】employs a second attention component, a selfattention pooling layer, to generalize the graph representation from the various aspects of a matrix graph embedding.





【实验】The dual attention network is trained in an end-to-end manner for the graph classification task. We compare our model with state-of-the-art graph kernels and other deep learning methods. The experimental results show that our framework not only outperforms other baselines but also achieves a better rate of convergence.