Multi-hop Graph Convolutional Networks for Accurate Node Classification

Junghun Kim, U Kang
Seoul National University
bandalg97@snu.ac.kr, ukang@snu.ac.kr

정확한 노드 분류를 위한 다중 홉 그래프 합성곱 신경망

김정헌, 강 유
서울대학교

Abstract

How can we leverage information from multi-hop neighbors in graph convolutional networks (GCNs) without relying on multiple stacking, which often causes over-smoothing? We propose Mh-GCN, a simple yet effective architecture that captures K-hop neighbor information in parallel by learning a weighted combination of hop-specific adjacency matrices. This single-layer design enables efficient and expressive representation learning across multiple scales without the drawbacks of stacked GCNs. Experiments on benchmark citation networks show that Mh-GCN consistently achieves powerful performance with low model complexity.

Ⅰ. Introduction

How can we learn accurate and expressive node representations using only a single-layer GCN? GCN-based methods have been extensively studied for solving various tasks on graph-structured data [1, 2, 3] They rely on stacking multiple layers to capture multi-hop neighborhood information, but they often suffer from over-smoothing and increased training complexity. To overcome these issues, we propose Mh-GCN, a framework that captures multi-hop context using a single layer, by learning a weighted combination of hop-specific adjacency matrices.

Ⅱ. Method

Before introducing our method, consider a naive GCN with K layers. The prediction probability $H \in R^{N\times C}$ is computed as follows:

$$H = \sigma\left(\tilde{A} \ldots \sigma\left(\tilde{A}\left(\sigma\left(\tilde{A}XW_1\right)W_2\right) \ldots\right)W_K\right)$$

where N and C are the numbers of nodes and classes, respectively, W_k for $k = 1, \ldots, K$ are learnable parameters, $X \in R^{N\times F}$ is the input node features, and σ is a nonlinear activation function. This stacking aggregates multi-hop neighbor information but is computationally expensive and prone to over-smoothing. To address this, Mh-GCN expresses multi-hop aggregation in a single layer.

Mh-GCN consists of two core components: (1) a hop-weighted aggregation of adjacency matrices, and (2) a single GCN layer using the aggregated structure for message passing and classification.

Let $\{A^{(1)}, A^{(2)}, \ldots, A^{(K)}\}$ be adjacency matrices representing 1-hop to K-hop connections: $A^{(k)} = \Pi_{i=1}^{k} A^{(1)}$ where $A^{(1)}$ is the adjacency matrix of the given graph. We propose a learnable scalar weight w_k

for each $A^{(k)} \in \{A^{(1)}, A^{(2)}, \dots, A^{(K)}\}$. Then the aggregated adjacency matrix is defined as follows:

$$A^{multi} = \sum_{k=1}^{K} w_k A^{(k)}$$

where weights w_k are normalized by the softmax function along the K layers to ensure stability.

Using the aggregated A^{multi}, the node representations are computed by a single GCN layer:

$$H = \sigma(A^{multi} X W)$$

where W is the weight matrix. Finally, a classifier (e.g., a linear layer followed by a softmax function) is applied on H for node classification.

To prevent A^{multi} from deviating too much from the original $A^{(1)}$, we additionally optimize $L_{reg} = |A^{multi} - A|_2$ along with the classification loss.

Ⅲ. Experiments

We test our Multi-hop GCN on citation networks (Cora, Citeseer, Pubmed) for semi-supervised node classification. Using ReLU, Adam optimizer, and 16 hidden units, we report mean accuracy (±std) over 10 runs with different random seeds.

We compare Mh-GCN with standard GCNs composed of 1, 2, and 3 layers. The results are summarized in Table 1. Mh-GCN consistently outperforms standard GCNs with the same number of parameters, demonstrating its superior ability to capture multi-hop information even with a single layer. Furthermore, Mh-GCN achieves performance comparable to deeper GCNs while maintaining more efficient architecture.

Ⅳ. Conclusion

We proposed Mh-GCN, a single-layer model that efficiently captures multi-hop dependencies via learnable adjacency fusion. Mh-GCN achieves competitive or superior accuracy compared to deeper GCNs, with fewer parameters and lower complexity.

Table 1. Accuracy of Mh-GCN.

Data	Model	layers	Acc.	Std.	Params
Cora	GCN	1	0.746	0.001	10,038
		2	**0.795**	0.011	**23,063**
		3	0.782	0.012	23,335
	Mh-GCN (proposed)	1	0.713	0.005	10,040
		2	**0.787**	0.008	**10,042**
		3	0.770	0.011	10,044
CiteSeer	GCN	1	0.633	0.006	22,224
		2	**0.669**	0.02	**59,366**
		3	0.635	0.015	59,638
	Mh-GCN (proposed)	1	0.559	0.002	22,225
		2	**0.653**	0.007	**22,226**
		3	0.620	0.009	22,227
PubMed	GCN	1	0.735	0.002	1,503
		2	**0.765**	0.002	**8,067**
		3	0.766	0.006	8,339
	Mh-GCN (proposed)	1	0.666	0.003	1,504
		2	0.754	0.006	1,505
		3	**0.756**	0.004	**1,506**

ACKNOWLEDGMENT

This work was supported by IITP grant funded by the MSIT [No.RS-2021-II211343, Artificial Intelligence Graduate School Program (Seoul National University)].

REFERENCES

[1] Kipf, Thomas N., and Max Welling. "Semi-supervised classification with graph convolutional networks." ICLR (2017).

[2] Kim, J., K. H. Park, H. Yoon, and U. Kang. "Accurate Link Prediction for Edge-Incomplete Graphs via PU Learning". AAAI (2025)

[3] Kim, J., K. H. Park, H. Yoon, and U. Kang. "Accurate Graph-based Multi-Positive Unlabeled Learning via Disentangled Multi-view Feature Propagation". KDD (2025)