GARNET: Reduced-Rank Topology Learning for Robust and Scalable Graph Neural Networks

Zepeng Zhang

June 17, 2022

GNNs are vulnerable to adversarial structural attacks. Previous low-rank approximation defense methods eliminate high-rank adversarial components to improve the robustness, while they may also impair the underlying clean graph structure that contributes to GNN training. This paper [1] proposes a spectral method GARNET to boost the robustness of GNNs, which first leverages weighted spectral embedding to construct a base graph and then refines it by pruning additional uncritical edges based on probability graphical model (PGM).

Given the top r smallest eigenvalues $\lambda_1, \dots, \lambda_r$ and their corresponding eigenvectors $v_1, \dots, v_r$ of normalized graph Laplacian matrix, the weighted spectral embedding matrix is defined as

$$V = \left[\sqrt{|1-\lambda_1|}v_1, \dots, \sqrt{|1-\lambda_r|}v_r\right],$$

whose i-th row is the weighted spectral embedding of the corresponding i-th node. Then a base graph is constructed with kNN based on the weighted spectral embedding, which is robust against adversarial attacks.

A further refinement of the base graph is conducted based on attractive Gaussian Markov random fields (GMRFs). The precision matrix is restricted to be a Laplacian-like matrix $\Theta = L + \frac{I}{\sigma^2}$ with σ^2 being a constant prior data variance. Recent methods for estimating attractive GMRFs are to solve the following convex problem:

$$\max_{\Theta} : F = \log \det \Theta - \frac{1}{r}\mathrm{tr}\left(VV^T\Theta\right) - \alpha\|\Theta\|_1,$$

where the first two terms together can be interpreted as log-likelihood under a GMRF and the last term is to promote sparsity. If we use gradient descent to solve this problem, then the importance of edges can be implied based on the magnitude of gradients. By setting $\alpha = 0$, the authors find a way to approximate the gradient of F w.r.t. A_{ij} as

$$\frac{\partial F}{\partial A_{i,j}} \approx \left\|U^T e_{i,j}\right\|_2^2 - \frac{1}{r}\left\|V^T e_{i,j}\right\|_2^2,$$

where $U = \left[\frac{u_1}{\sqrt{\lambda_1 + 1/\sigma^2}}, \dots, \frac{u_r}{\sqrt{\lambda_r + 1/\sigma^2}}\right]$ with u being the corresponding eigenvector of the Laplacian matrix. Then the edges with small spectral embedding distortion $\frac{\left\|U^T e_{i,j}\right\|_2^2}{\|V^T e_{i,j}\|_2^2}$ are pruned.

References

[1] Chenhui Deng, Xiuyu Li, Zhuo Feng, and Zhiru Zhang. Garnet: Reduced-rank topology learning for robust and scalable graph neural networks. *arXiv preprint arXiv:2201.12741*, 2022.