Conference Paper Title*

*Note: Sub-titles are not captured in Xplore and should not be used

1st Given Name Surname
dept. name of organization (of Aff.)
name of organization (of Aff.)
City, Country
email address or ORCID

2nd Given Name Surname
dept. name of organization (of Aff.)
name of organization (of Aff.)
City, Country
email address or ORCID

I. Introduction

A. Current work

Currently typical message passing mechanisms of graph neural networks (GNN) are highly related to the graph structure. For example, graph attention networks (GAT) passes information via attention sharing on global nodes, graph convolution network (GCN) performs approximations on spectral convolution. However, there still exists information which are not position-aware, which indicates they are unable to be shared during the position aware message passing. For example, in relation networks like social networks and citation networks, there may exists orphan nodes which contains features not similar to those of connected nodes. Typical solution is random sampling on the whole node set of the graph, regardless of the edge connections. Thus successful extraction of these structure-independent features may significantly boosts the representation performance of the node. Typical solution is random sampling on the whole node set of the graph, regardless of the edge connections. [1] randomly samples several subsets containing arbitrary count of nodes on the original graph, and stacks as a feature matrix. All subsets further aggregate into a shared vector by aggregation (subset-wise mean), and fused to each node feature vector. While [2] divides the graph via random walk and find shared information from different samples on the unique structure topic. However, random sampling cannot perfectly extract the global information, which performance may influenced by the randomly sampled result. Mechanics with direct operation on the whole graph is required.

Non-negative matrix factorization (NMF) is a matrix factorization mechanism which is able to learn parts-based representations by 2 low rank matrices, as the linear combination of the data-based basis and weights [3], like other typical factorization methods like principle component analysis (PCA) and vector quantization (VQ). Among which, eigen-based methods PCA and VQ mainly extracts holistic features, where basis components manifesting building blocks of the original matrix may missing [4]. NMF makes the computation more efficient and has achieved wide application in engineering, such as data mining and image processing, to name a few. In image processing, [3] and [4] uses NMF to calculate the basis representation of human face data in facial recognition and classification tasks. In data mining, NMF is able to calculate the semantic feature representations in articles [3]. NMF has also become a basic baseline algorithm in recommendation system as representing the individual-item relation by refined individual-topic and topic-item relations. Besides, [2] also uses NMF to factorize the key structure topic of sampled results from different random walks, and lead to the following principal component calculation. However typically, NMF is applied in end to end approaches as feature extractor, where potential distribution shift may exists between the learned representation by NMF and the original node features. Thus separated implement of NMF may still cause performance decreasing in contrast with the approach without NMF.

B. Our contribution

We purpose a novel approach enhancing the message passing of GNNs via non-negative matrix factorization (NMF), for leading the structure-invariant components into the node features. Inspired by above researches, we directly perform NMF on the feature matrix and gained a global basis matrix and the corresponding weight matrix. The operation on the whole matrix imposes the global feature distribution into the factorized basis, which is further embedded into a feature vector and send into message passing together with original node features. Furthermore, to better integrate NMF into message passing framework, the basis matrix is randomly initialized and update via backward propagation with the whole graph rather than explicitly perform NMF on the original feature matrix, to achieve distribution consistency between the feature of basis representation and node features.

Identify applicable funding agency here. If none, delete this.

II. Methods

Non-negative matrix factorization (NMF) is a matrix approximation mechanism of representing the big original matrix A by the product of 2 low rank matrices as $\mathbf{A} \approx \mathbf{WH}$. Among which, the factorization can be represented as expressing the original matrix by a set of more related basis, where $\mathbf{W}$ is the basis matrix and $\mathbf{H}$ is the projection to those basis.

III. Results

We've implemented a naive run on Cora dataset [6], which consists 2708 scientific publication instances classified into 7 classes. The dataset provides a publication-keywords-classification characteristic table and a publication-publication citation relationship table. The implementation uses mean method to aggregate basis matrix into global feature vector, and uses GAT as the message passing backbone. The experiment result is shown in table III

Implements	Inference Accuracy	Time consultant on each epoch (s)
GAT	0.875	0.
Ours	0.842	

References

[1] You, J., Ying, R., & Leskovec, J. (2019). Position-aware graph neural networks. arXiv preprint arXiv:1906.04817.

[2] Long, Q., Jin, Y., Song, G., Li, Y., & Lin, W. (2020, August). Graph Structural-topic Neural Network. In Proceedings of the 26th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining (pp. 1065-1073).

[3] Lee, D. D., & Seung, H. S. (1999). Learning the parts of objects by non-negative matrix factorization. Nature, 401(6755), 788-791.

[4] Feng, T., Li, S. Z., Shum, H. Y., & Zhang, H. (2002, June). Local non-negative matrix factorization as a visual representation. In Proceedings 2nd International Conference on Development and Learning. ICDL 2002 (pp. 178-183). IEEE.

[5] Huang, Z., Zhou, A., & Zhang, G. (2012, October). Non-negative matrix factorization: A short survey on methods and applications. In International Symposium on Intelligence Computation and Applications (pp. 331-340). Springer, Berlin, Heidelberg.

[6] Sen, P., Namata, G., Bilgic, M., Getoor, L., Galligher, B., & Eliassi-Rad, T. (2008). Collective classification in network data. AI magazine, 29(3), 93-93.