

Global and Local Structure Learning for Sparse Tensor Completion

Dawon Ahn

Computer Science and Engineering
University of California, Riverside
Riverside, USA
dahn017@ucr.edu

Evangelos E. Papalexakis

Computer Science and Engineering
University of California, Riverside
Riverside, USA
epapalex@cs.ucr.edu

Abstract—*How can we accurately complete tensors by learning relationships of dimensions along each mode?* Tensor completion, a widely studied problem, is to predict missing entries in incomplete tensors. Tensor decomposition methods, fundamental tensor analysis tools, have been actively developed to solve tensor completion tasks. However, standard tensor decomposition models have not been designed to learn relationships of dimensions along each mode, which limits to accurate tensor completion. Also, previously developed tensor decomposition models have required prior knowledge between relations within dimensions to model the relations, expensive to obtain. This paper proposes TGL (Tensor Decomposition Learning Global and Local Structures) to accurately predict missing entries in tensors. TGL reconstructs a tensor with factor matrices which learn local structures with GNN without prior knowledges. Extensive experiments are conducted to evaluate TGL with baselines and datasets.

Index Terms—Tensor Decomposition, GNNs

I. INTRODUCTION

Given an incomplete tensor, how can we predict missing values learning relations within dimensions?

A tensor is a natural way to represent multi-dimensional data. For example, 3-mode tensor (user, movie, time) represents movie rating data where users give rates to movies at specific times. The goal of tensor completion is to impute missing entries in incomplete tensors which are partially observed. The tensor completion problems have been actively studied across diverse domain of applications such as recommender system [1]–[3], anomaly detection [4], computer vision [5], and air quality analysis [6].

Tensor decomposition methods have played a key role in tensor analysis such that they have significantly been developed to solve the completion tasks. CP (CANDECOMP/PARAFAC) [7], [8] factorization is one of the most widely used tensor factorization models due to its simplicity and interpretability, which factorizes a tensor into a set of factor matrices and a core tensor which is restricted to be diagonal. However, the design of standard CP decomposition model does not consider explicitly relations between dimensions, which degrades the performance in missing entries prediction.

Recently, various types of tensor decomposition methods have been proposed to leverage prior knowledges, such as temporal knowledge, correlations between dimensions, as a

regularization into loss function [6], [9]. However, prior knowledges are expensive to obtain for general datasets.

In this paper, we propose TGL (Tensor Decomposition Learning Global and Local Structures) to learn relationships between dimensions without prior knowledges. We leverage Graph Neural Networks (GNNs) to learn local structures. TGL initially produces representations with a standard CP decomposition and then generates relation matrices with the representations. With representation and relation matrices, TGL updates representations via GNNs for each mode. We show TGL effectively learn relationships and achieve comparable accuracy compared to existing baselines in tensor completion.

II. PROPOSED METHOD

In this study, we utilized both a CP decomposition method and GNNs to learn relations within dimensions of each mode for incomplete tensors without prior knowledges. In each iteration of training, we generate K-nearest neighborhoods (KNN) graphs with the current factor matrices for each mode. Then we update factor matrices using KNN graphs and GNNs. With these newly updated factor matrices, we reconstruct a tensor to jointly train factor matrices and GNNs.

We generate graphs \mathbf{R}_A , \mathbf{R}_B and \mathbf{R}_C with factor matrices \mathbf{A} , \mathbf{B} , and \mathbf{C} by computing pairwise cosine similarity for each mode. A graph G contains a set of nodes $V(|V| = N)$, a set of edges E , and an adjacency matrix $\mathbf{R} \in \mathbb{R}^{N \times N}$ as $G = (V, E)$. Given the number of layers $l (1 \leq l \leq L)$, $l + 1$ th hidden convolutional layer of GNNs takes feature matrix of the l th layer and the adjacency matrix

$$\mathbf{H}^{(l+1)} = f(\mathbf{H}^{(l)}, \mathbf{R}) = \sigma(\hat{\mathbf{R}}\mathbf{H}^{(l)}\mathbf{W}^{(l)})$$

where $\hat{\mathbf{R}} = \tilde{\mathbf{D}}^{-\frac{1}{2}}\tilde{\mathbf{R}}\tilde{\mathbf{D}}^{\frac{1}{2}}$ is a symmetric normalization of the self-connections added adjacency matrix $\tilde{\mathbf{R}} = \mathbf{R} + \mathbf{I}$, $\tilde{\mathbf{D}}$ is the diagonal node degree matrix with $\tilde{d}_{ii} = \sum_j \tilde{r}_{ij}$, $\mathbf{W}^{(l)} \in \mathbb{R}^{d_l \times d_{l+1}}$ is a trainable weight matrix. σ is a non-linear activation function. Note that initial feature matrix $\mathbf{H}^{(0)}$ is the factor matrix obtained.

Given factor matrices \mathbf{A} , \mathbf{B} , and \mathbf{C} , relation matrices \mathbf{R}_A , \mathbf{R}_B and \mathbf{R}_C , and GNN models $\{f_1, f_2, f_3\}$ for each

mode, the new factor matrices are obtained by minimizing the following loss function:

$$\mathcal{L}(\mathbf{A}, \mathbf{B}, \mathbf{C}) = \sum_{\forall \alpha \in \Omega} \left(x_\alpha - \sum_{r=1}^R \tilde{a}_{ir} \tilde{b}_{jr} \tilde{c}_{kr} \right)^2 \quad (1)$$

where \tilde{a}_{ir} , \tilde{b}_{jr} , and \tilde{c}_{kr} is the (i, r) th, (j, r) th, and (k, r) th entry of $\tilde{\mathbf{A}} = f_1(\mathbf{A}, \mathbf{R}_A)$, $\tilde{\mathbf{B}} = f_2(\mathbf{B}, \mathbf{R}_B)$, and $\tilde{\mathbf{C}} = f_3(\mathbf{C}, \mathbf{R}_C)$. I will explain details of network hyper parameters in Section III.

III. EXPERIMENTS

We perform experiments to evaluate the accuracy of TGL compared to baselines in tensor completion.

A. Experiments Setting

1) *Datasets*: We evaluate the performance of TGL and baselines on three real-world datasets. **Yelp** [10] is a business rating tensor with (user, business, time; rating) which includes 70,817 movies, 15,579 users, 108 year-month time duration, and 100,0836 ratings. We sample each 5,000 in user and business such that the statistics of dataset used is (5,000, 5,000, 108) since the proposed model can not handle too large dataset. **BBC-News** is a BBC news article tensor with (word, word, articles; co-occurrences) which includes 100 words and 400 articles. We split each data into training, validation, and test sets with the ratio of 8:1:1 considering only observable entries and use validation set for early stopping.

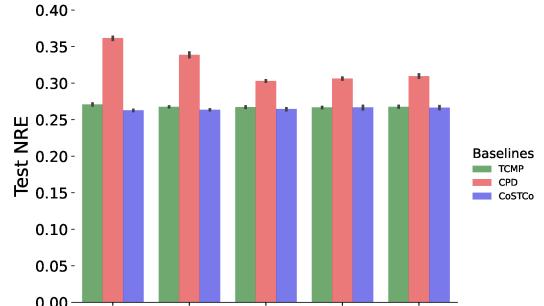
2) *Baselines*: There are two baselines: CPD [11] is a standard CP decomposition model considering only nonzeros with gradient descent optimization. CoSTCo [3] is a state-of-the-art tensor completion method exploiting Convolutional Neural Networks.

3) *Metrics*: We use NRE (Normalized Reconstruction Error) to evaluate the accuracy of tensor completion, which is defined as follows.

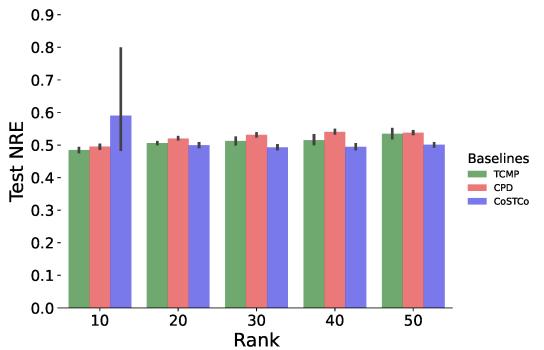
$$NRE = \frac{\sqrt{\sum_{\forall \alpha \in \Omega} (x_\alpha - \hat{x}_\alpha)^2}}{\sqrt{\sum_{\forall \alpha \in \Omega} x_\alpha^2}}.$$

B. Tensor Completion Accuracy

Figure 1 show the performance of TGL and baselines, with regard to test NRE. TGL shows the second best method while CPD using the standard CP decomposition method shows large errors. CoSTCo shows the lowest NRE among all methods. Even if TGL shows better performance than the standard model, one potential reason that TGL shows limited performance is that when training GNNs, TGL gives relation matrices and factor matrices which can have redundant information since relation matrices are created based on factor matrices.



(a) Yelp



(b) BBC-News

Fig. 1. Comparison of test NRE of baselines on Yelp and BBC-News dataset with different size of ranks.

IV. DISCUSSION & CONCLUSIONS

In this paper, we proposed TGL, a tensor decomposition assisted with GNNs to learn global and local structures of a sparse tensor. Experiment results show that TGL outperforms CPD and shows comparable performance as CoSTCo. As mentioned in Experiments section, a graph generated from TGL might give redundant information to GNNs since relation matrices are made from factor matrices. Therefore, we try to make distinct graph structures from node features for GNNs for the future works.

ACKNOWLEDGMENTS

Research was supported by the National Science Foundation under CAREER grant no. IIS 2046086, by the Agriculture and Food Research Initiative Competitive Grant no. 2020-69012-31914 from the USDA National Institute of Food and Agriculture, and by the University Transportation Center for Railway Safety (UTCRS) at UTRGV through the USDOT UTC Program under Grant No. 69A3552348340.

REFERENCES

- [1] A. Karatzoglou, X. Amatriain, L. Baltrunas, and N. Oliver, “Multiverse recommendation: n-dimensional tensor factorization for context-aware collaborative filtering,” in *Proceedings of the fourth ACM conference on Recommender systems*, 2010, pp. 79–86.
- [2] Q. Song, H. Ge, J. Caverlee, and X. Hu, “Tensor completion algorithms in big data analytics,” *ACM Transactions on Knowledge Discovery from Data (TKDD)*, vol. 13, no. 1, pp. 1–48, 2019.

- [3] H. Liu, Y. Li, M. Tsang, and Y. Liu, “Costco: A neural tensor completion model for sparse tensors,” in *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, 2019, pp. 324–334.
- [4] T. Kwon, I. Park, D. Lee, and K. Shin, “Slicenstitch: Continuous cp decomposition of sparse tensor streams,” in *2021 IEEE 37th International Conference on Data Engineering (ICDE)*. IEEE, 2021, pp. 816–827.
- [5] M. A. O. Vasilescu and D. Terzopoulos, “Multilinear analysis of image ensembles: Tensorfaces,” in *European conference on computer vision*. Springer, 2002, pp. 447–460.
- [6] D. Ahn, J.-G. Jang, and U. Kang, “Time-aware tensor decomposition for sparse tensors,” *Machine Learning*, vol. 111, no. 4, pp. 1409–1430, 2022.
- [7] R. A. Harshman *et al.*, “Foundations of the parafac procedure: Models and conditions for an “explanatory” multimodal factor analysis,” 1970.
- [8] H. A. Kiers, “Towards a standardized notation and terminology in multi-way analysis,” *Journal of Chemometrics: A Journal of the Chemometrics Society*, vol. 14, no. 3, pp. 105–122, 2000.
- [9] H.-F. Yu, N. Rao, and I. S. Dhillon, “Temporal regularized matrix factorization for high-dimensional time series prediction,” *Advances in neural information processing systems*, vol. 29, 2016.
- [10] N. Asghar, “Yelp dataset challenge: Review rating prediction,” *arXiv preprint arXiv:1605.05362*, 2016.
- [11] K. Shin, L. Sael, and U. Kang, “Fully scalable methods for distributed tensor factorization,” *IEEE Transactions on Knowledge and Data Engineering*, vol. 29, no. 1, pp. 100–113, 2016.