# When Sparse Graph Representation Learning Falls into Domain Shift: Data Augmentation for Cross-Domain Graph Meta-Learning (Student Abstract)

**Simin Niu[1], Xun Liang[1], Sensen Zhang[1], Shichao Song[1], Xuan Zhang[1,2], Xiaoping Zhou[3]**

[1] Renmin University of China
[2]Peking University
[3]Beijing University of Civil Engineering and Architecture
{niusimin, xliang, sensen0126, songshichao}@ruc.edu.cn, zhangxuan01@jsfund.cn, lukefchou@gmail.com

## Abstract

Cross-domain Graph Meta-learning (CGML) has shown its promise, where *meta-knowledge* is extracted from few-shot graph data in multiple relevant but distinct domains. However, several recent efforts assume target data available, which commonly does not established in practice. In this paper, we devise a novel **C**ross-domain **D**ata **A**ugmentation for **G**raph **M**eta-**L**earning (**CDA-GML**), which incorporates the superiorities of CGML and Data Augmentation, has addressed intractable shortcomings of label sparsity, domain shift, and the absence of target data simultaneously. Specifically, our method simulates instance-level and task-level domain shift to alleviate the cross-domain generalization issue in conventional graph meta-learning. Experiments show that our method outperforms the existing state-of-the-art methods.

## Introduction

Graph Meta-learning (GML) has a capability to alleviate performance degradation in face of a few-shot scenario by utilizing *meta-knowledge* extracted from scarcely annotated data on graphs yet challenging for generalization to out-of-distribution data (OOD). A few methods on how to deal with domain shift have been proposed. The earliest attempt focused on CGML was in (Hassani 2022), where an attention-based graph encoder was proposed to employ contextual and topological views of graphs to obtain task-specific representations for the few-shot graph classification tasks. From domain adaptation perspective, a well-designed cross-domain meta-learning method (Zhang et al. 2022) for the graph encoder called CrossHG-Meta has displayed satisfactory performance when deploying trained GML models in cross-domain few-shot scenarios. However, such a straightforward solution to bypass the OOD data issue relies on a common assumption that several data of the target domain is accessible, which is often challenging in practice because target data is always laborious to obtain or even unknown.

This paper proposes a novel Graph Meta-learning framework combining the advantages of Data Augmentation and Cross-domain Graph Meta-learning strategy which perform well on Domain Generalization to solve the three limitations (i.e., the label sparsity issue, the domain shift problem, and the absence of target data) as shown in Figure 1.



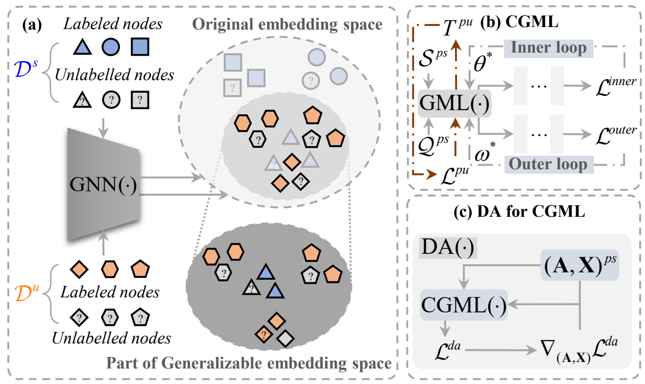


Figure 1: Issues of domain shift problem of existing Graph Meta-learning frameworks and overview of CDA-GML.

Key contributions are as follows: (1) We propose a Cross-domain Graph Meta-learning (CGML) framework to simulate domain-shift between meta-training stage and meta-testing stage to achieve satisfactory performance by conducting meta-optimize. (2) We propose a Data Augmentation strategy for CGML (DA for CGML) based on a transformation, which simulates instance-level domain shift and directly augments data during the meta-training stage.

## Proposed Method

**CGML.** In cross-domain graph meta-learning, given a graph $\mathcal{G} = \{\mathcal{V}, \mathcal{E}, \mathbf{A}, \mathbf{X}\}$, where $\mathcal{V}$ denotes the node set ($n = |\mathcal{V}|$); $\mathcal{E}$ is the edge set; $\mathbf{A} \in \mathbb{R}^{n \times n}$ denotes the adjacency matrix; $\mathbf{X} \in \mathbb{R}^{n \times d}$ is the feature matrix, and $d$ is the dimension of node feature vector, we can assume $M$ seen domains $\mathcal{D}^s = \{\mathcal{T}_i^s\}_{i=1}^M$ accessible and denote a domain $\mathcal{D}$ consisting of $m$ meta-tasks as $\mathcal{T} = \{T_i\}_{i=1}^m$, deeply, each task $T$ contains a support set $\mathcal{S} = \{((\mathbf{A}, \mathbf{X})_s, \mathcal{Y}_s)\}$ and a query set $\mathcal{Q} = \{((\mathbf{A}, \mathbf{X})_q, \mathcal{Y}_q)\}$, where $\mathcal{Y}$ is the corresponding node label set. We further train a predictive model $\hat{\mathcal{Y}} = \sigma(\mathrm{CGML}_{\theta,\omega}(\mathbf{A}, \mathbf{X})^{ps})$ parameterized by $\{\theta, \omega\}$ in the meta-training stage as follows:

$$\omega^* = \arg\min_{\omega} \mathcal{L}^{outer}(\mathcal{Q}^{ps}; \theta^*, \omega) \tag{1}$$

$$\text{s.t.} \qquad \theta^* = \arg\min_{\theta} \mathcal{L}^{inner}(\mathcal{S}^{ps}; \theta, \omega), \tag{2}$$

where $\mathcal{L}^{inner}$ and $\mathcal{L}^{outer}$ refer to the loss of support set and query set respectively; $\theta^*$ produced is capable of performing well on query sets after optimizing the outer loss executed on $\omega$ in the meta-training stage. $\omega^*$ indicates the learned *meta-knowledge*. Note that the selection of the output layer activation function $\sigma(\cdot)$ determines whether to carry out classification or regression tasks end-to-end.

Immediately, we simulate domain-shift between meta-training tasks and meta-testing tasks to meta-optimize for satisfactory performance under domain-shift. Specifically, we adopt the learned meta-knowledge $\omega^*$ in the meta-testing stage to perform training on the pseudo-unseen task $T^{pu} = \{(\mathbf{A},\mathbf{X})_s^{pu}, \mathcal{Y}_s^{pu}), ((\mathbf{A},\mathbf{X})_q^{pu}, \mathcal{Y}_q^{pu})\}$ in a pseudo-unseen domain $\mathcal{T}^{pu}$ as $\theta' = \arg\min_{\theta^*} \mathcal{L}(\mathcal{S}^{pu}; \theta^*, \omega^*)$, then we evaluate the generalizability of the updated optimization-based graph meta-learning model on the pseudo-unseen domains by the loss of meta-testing task $\mathcal{L}^{pu}(\mathcal{Q}^{pu}; \theta', \omega^*)$. Eventually, the parameters of the predictive model are optimized by $(\theta,\omega) \leftarrow (\theta,\omega) - \alpha\nabla_{\theta,\omega}\mathcal{L}^{pu}$ to ensure that the final model can work well on the unseen domains after training on the seen domains. See Figure 1(b) for illustrations.

**DA for CGML.** Progress in CGML methods is bided fair to be spurred by well designed Data augmentation strategy. Specifically, we augment the original input by applying a transformation DA$(\cdot)$ parameterized by $\{\theta,\omega\}$ to simulate domain shift on each instance as shown in Figure 1(c). More detailed, we first obtain the adversarial gradient $\nabla_{(\mathbf{A},\mathbf{X})}\mathcal{L}^{da}(\hat{\mathcal{Y}} - \mathcal{Y})$ calculated from CGML classifier for each node in the meta-training stage. Then, we combine the original node embeddings $\mathbf{Z}$ and the interference factor $\mathbf{Z}^*$ generated by the adversarial gradient to enhance and expand original data distributions, thus allowing CGML to learn more generalizable embeddings. Here, we achieve the adjusted node embeddings $\mathbf{Z}'$ by performing the element-wise addition between $\mathbf{Z}$ and $\mathbf{Z}^*$ as follows:

$$\mathrm{DA}_{\theta,\omega}(\mathbf{A},\mathbf{X}) = \mathrm{CGML}_{\theta,\omega}(\mathbf{A},\mathbf{X}) \oplus \nabla_{(\mathbf{A},\mathbf{X})}\mathcal{L}^{da}. \quad (3)$$

By executing the above process, data augmentation can be brought in to benefit CGML by directly augmenting the meta-training data, and thus alleviating the inferior generalizability in meta-learning under the cross-domain setting.

## Experiment and Discussion

**Datasets and Evaluation.** We evaluated CDA-GML on few-shot node classification tasks on the AMiner dataset (Zhang et al. 2022) with four domains. We took the same settings as (Zhang et al. 2022), i.e., selected the best classification accuracy on domain *System* as evaluation metric. Next, we conducted the few-shot node classification tasks under the settings of {2,3}-way (the number of categories) and {1,3}-shot (the number of labeled nodes for each category) on the basis of the optimization-based meta-learning method MAML. Last, we compared our framework with four types of state-of-the-art baseline methods, including GNNs algorithms (Kipf and Welling 2017; Schlichtkrull et al. 2018), Few-shot Learning (FSL) methods (Snell, Swersky, and Zemel 2017; Finn, Abbeel, and Levine 2017), Graph Meta-learning (GML) models (Zhou et al. 2019; Wang et al. 2020), and the advanced CrossHG-Meta (Zhang et al. 2022).

| Category | Method | 2-way | | 3-way | |
|---|---|---|---|---|---|
| | | 1-shot | 3-shot | 1-shot | 3-shot |
| GNNs | GCN | 73.90 | 84.90 | 64.15 | 71.64 |
| | R-GCN | 74.89 | 85.27 | 65.77 | 74.61 |
| FSL | ProtoNet | 72.07 | 80.79 | 62.79 | 71.81 |
| | MAML | 76.45 | 84.94 | 63.73 | 74.33 |
| GML | Meta-SGC | 78.41 | 87.27 | 67.41 | 75.85 |
| | AMM-GNN | 82.10 | 90.18 | 72.24 | 82.59 |
| CGML | CrossHG-Meta | 83.81 | 91.44 | 75.59 | 84.57 |
| | CDA-GML\CD | 84.36 | 92.25 | 77.14 | 85.49 |
| | CDA-GML\DA | 84.19 | 92.16 | 76.70 | 85.22 |
| | **CDA-GML (ours)** | **85.76** | **93.83** | **78.21** | **86.68** |

Table 1: Node classification accuracy (%) for *System* domain. Ablation study: \CD denotes without the CGML module; \DA denotes Data Augmentation module is removed.

**Result Analysis and Ablation Study.** Table 1 illustrates the performance of our model CDA-CGML in comparison with all baselines. CrossHG-Meta model performed the best in baselines and the CDA-CGML method outperformed all the vanilla baselines on the target domain *System*, which idicate that cross-domain graph meta-learning method has the capability to achieve generalizability in unseen domains. Meanwhile, removing each module typically suffers performance drops, which indicates that our proposed components help graph meta-learning get higher generalizability.

## Conclusion

In this paper, we attempt to address limited generalization performance in cross-domain few-shot graph scenarios due to the problem of label sparsity, domain shift, and the absence of target data with CDA-GML. Results show the superiority of CDA-GML over the current advanced methods. Xun Liang is the corresponding author of this paper.

## References

Finn, C.; Abbeel, P.; and Levine, S. 2017. Model-Agnostic Meta-Learning for Fast Adaptation of Deep Networks. In *ICML*.

Hassani, K. 2022. Cross-Domain Few-Shot Graph Classification. In *AAAI*, 6856–6864.

Kipf, T.; and Welling, M. 2017. Semi-Supervised Classification with Graph Convolutional Networks. In *ICLR*.

Schlichtkrull, M.; Kipf, T.; Bloem, P.; Berg, R.; Titov, I.; and Welling, M. 2018. Modeling Relational Data with Graph Convolutional Networks. In *ESWC*, 593–607.

Snell, J.; Swersky, K.; and Zemel, R. 2017. Prototypical Networks for Few-shot Learning. In *NeurIPS*, 4077–4087.

Wang, N.; Luo, M.; Ding, K.; Zhang, L.; Li, J.; and Zheng, Q. 2020. Graph Few-shot Learning with Attribute Matching. In *CIKM*, 1545–1554.

Zhang, Q.; Wu, X.; Yang, Q.; Zhang, C.; and Zhang, X. 2022. Few-shot Heterogeneous Graph Learning via Cross-domain Knowledge Transfer. In *SIGKDD*, 2450–2460.

Zhou, F.; Cao, C.; Zhang, K.; Trajcevski, G.; Zhong, T.; and Geng, J. 2019. Meta-GNN: On Few-shot Node Classification in Graph Meta-learning. In *CIKM*, 2357–2360.