Unveiling Intrinsic Graph Data Properties for Domain Adaptation

Ruiyi Fang

Department of Computer Science
Western University
rfang32@uwo.ca

1 Abstract

Graphs are ubiquitous in real-world applications, capturing rich relational structures and node-level attributes. However, labeled graph data are often scarce, and the distribution shift between source and target graphs poses a critical challenge for **label transfer**. Traditional graph domain adaptation (GDA) methods primarily align structural distributions, but they fail when intrinsic graph properties such as node attributes and homophily patterns differ across domains. To address this, we propose a unified framework of *intrinsic property alignment*, which tackles label transfer from two complementary perspectives: (i) an **attribute-driven alignment** that explicitly models and aligns feature graphs to reduce attribute divergence, and (ii) a **homophily-enhanced alignment** that disentangles and aligns homophilic, heterophilic, and attribute signals via mixed graph filters. Both theoretical analysis and empirical studies on multiple benchmarks demonstrate that aligning these intrinsic properties significantly improves cross-network label transfer, achieving robust and interpretable GDA performance.

2 Introduction

Graph-structured data is powerful but inherently challenging for knowledge transfer. They integrate complex topological connections with high-dimensional node attributes, making it difficult to reuse labeled information when domain shifts occur. In real scenarios, the source and target graphs often differ in structural patterns, feature distributions, and local connectivity (homophily). This causes the learned representations to become domain-specific, hindering **label transfer** to unlabeled graphs. Existing GDA frameworks mainly focus on aligning graph structures or spectral features, which overlook the deeper intrinsic factors governing label propagation.

Our recent works address this gap from two orthogonal but complementary perspectives. First, the **Graph Attribute Alignment (GAA)** [1] framework reveals that attribute divergence dominates topology shift in many datasets and proposes cross-view similarity refinement between topology and attribute graphs to achieve stable label transfer. Second, the **Homophily-Enhanced GDA (HGDA)** [2] framework highlights that local homophily mismatch impairs cross-domain consistency and introduces mixed filters to separately align homophilic, heterophilic, and attribute signals. Theoretical analysis based on PAC-Bayesian bounds supports that these intrinsic property shifts directly affect the generalization risk of GDA, providing a principled explanation for label transfer failure across graphs.

3 Future Work

Future research in GDA can move toward more **scalable and structural generalization** scenarios. One promising direction is the **from small to large graph** paradigm, where models trained on a small or local graph, such as a district-level transportation network, are transferred to a larger or more complex graph, like an entire metropolitan or inter-city network. This transition raises new challenges in hierarchical representation learning, multi-scale alignment, and computational scalability. Another important avenue is the **from dense to sparse** setting, which studies how models adapted on dense, well-connected graphs can generalize to sparse and incomplete networks. Such scenarios frequently arise in emerging domains such as sensor networks, recommendation graphs, and the newly RAG knowledge graph, where limited connectivity intensifies the difficulty of label transfer.

References

[1] Ruiyi Fang, Bingheng Li, Zhao Kang, Qiuhao Zeng, Nima Hosseini Dashtbayaz, Ruizhi Pu, Boyu Wang, and Charles Ling. On the benefits of attribute-driven graph domain adaptation. In *ICLR*, 2025.

[2] Ruiyi Fang, Bingheng Li, Jingyu Zhao, Ruizhi Pu, Qiuhao Zeng, Gezheng Xu, Charles Ling, and Boyu Wang. Homophily enhanced graph domain adaptation. *ICML*, 2025.