Heterogeneous Graph Label Distribution Learning

Author: Yufei Jin, Heng Lian, He Yi, Xingquan Zhu
Florida Atlantic University,
College of William & Mary,
Presenter: Yufei Jin

Supported under grant Nos. IIS-2236578, IIS-2236579, IIS-2302786, IIS-2441449, IOS-2430224, and IOS-2446522.

FLORIDA ATLANTIC UNIVERSITY

WILLIAM & MARY
CHARTERED 1693

NEURAL INFORMATION PROCESSING SYSTEMS

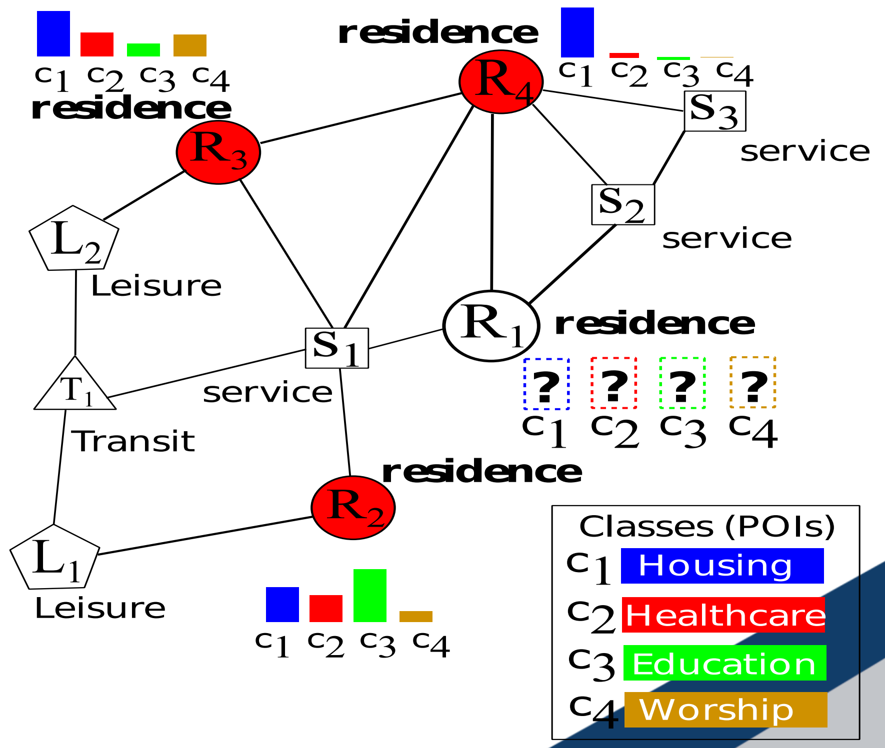
Introduction

- Goal: solve the label distribution learning problem on complex heterogeneous networks.
- Label Distribution Learning (LDL)
- Heterogeneous Graph Learning:
  - Meta-path based approach

Main motivation

Example Study of Urban Planning:

- LDL: extend the prediction from single point of interest (POI) of local regions to distribution of POIs of local regions, providing insights of the functional degrees of all four POIs in this region.
- Heterogeneous Graph: existing LDL methods overlook the information of urban topology.



Introduction

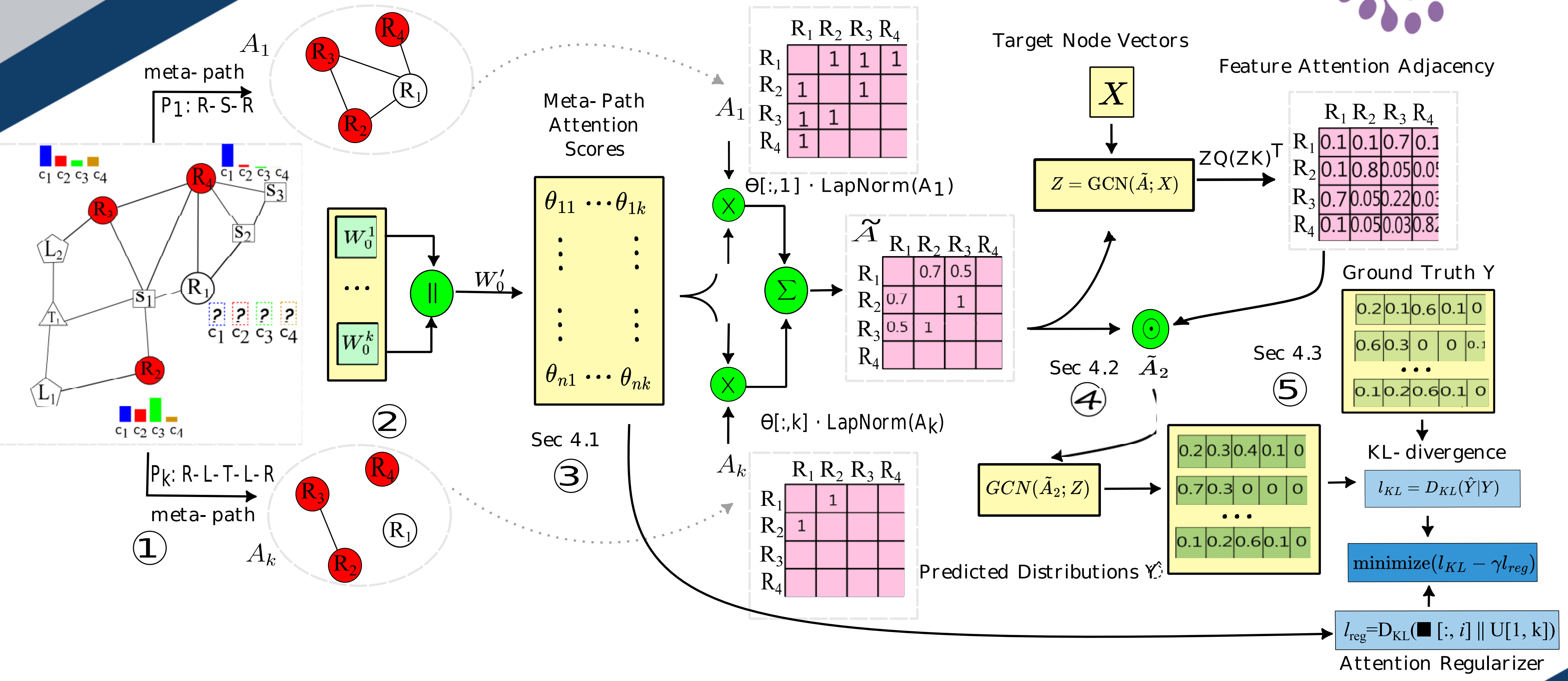
- Contribution:
  - An end-to-end **HGDL** learning approach to jointly learn an optimal meta-path graph topology and align it with nodal features for consistent message-passing.
  - Performance of **HGDL** is theoretically bounded.
  - Empirical study has been carried out over five self-created graph datasets that span domains of bio-medicine, scholarly network, business network, and urban planning.

FLORIDA ATLANTIC UNIVERSITY

WILLIAM & MARY
CHARTERED 1693

NEURAL INFORMATION PROCESSING SYSTEMS

Framework



The proposed HGDL framework. Step 1 - 3: Optimal Graph Topology Homogenization; Step 4 - 5: Local Topology and Global Feature Consistency-Aware Graph Transformer.

Framework

- ***Optimal Graph Topology Homogenization***: Unlike traditional heterogeneous graph learning methods that handle each meta-path separately, we leveraged an attention mechanism to merge all meta-path graph topology information into a single dynamic graph and learn embeddings based on the single learned dynamic graph.
- ***Local Topology and Global Feature Consistency-Aware Graph Transformer***: An attention mechanism is used to induce feature topology. We then integrate it with merged graph topology to learn a sparse graph for message passing embedding learning.
- ***An End-to-End HGDL Objective Function***: a joint loss combining label distribution loss (KL divergence) and attention regularization is used to ensure the diversity of learned different meta-path weights.

FLORIDA ATLANTIC UNIVERSITY

WILLIAM & MARY
CHARTERED 1693

NEURAL INFORMATION PROCESSING SYSTEMS

Benchmarks

- *Single meta-path comparison*
  - GCN_{KL}: A baseline that uses graph constructed from each meta-path to train a vanilla GCN[1], using KL-divergence as loss function, and reports the best meta-path result.
  - GLDL: This is a label distribution learning method proposed specifically for homogenous graph [2].
- *Fused Embeddings from different meta-path*
  - HAN_{KL}: This baseline uses HAN [3] structure (a simple attention mechanism) to integrate embedding from different meta-paths.
  - $SeHGNN_{KL}$: This baseline uses SeHGNN [4], leveraging a transformer based approach to aggregate meta-paths embedding along with KL-divergence loss function.
  - $HINormer_{KL}$: This is a more recent heterogenous graph learning baseline with KL-divergence loss function [5].

FLORIDA ATLANTIC UNIVERSITY

WILLIAM & MARY
CHARTERED 1693

NEURAL INFORMATION PROCESSING SYSTEMS

Dataset

Dataset	Node Types	# features	Edge Types	# Labels
DRUG	D:Drug(894)/S:Disease(454) P:Protein(18877)/G:gene(20561)	191	DP(4397)/DS(2704) PG(18545)/PP(201382) GG(712546)/DD(798316)	28
ACM	A:Author(5810)/P:Paper(12499) S:Subject(73)/C:conference(14) F:Affiliation(1804)	1903	AP(37055)/PC(12499) PS(12499)/PF(12499) AF(30424)	14
DBLP	A:Author(4057)/P:Paper(14328) T:Term(8920)/C:conference(20)	8920	AP(19645)/PT(114273) PC(14328)	4
YELP	U:User(3001)/B:Business(150346) R:Review(6990280)/T:Tip(908915)	19	UR(3001)/RB(6990280) UT(3001)/BT(908915)	9
URBAN	NR:Nature Residence(622) CS:Comprehensive Service(449) GL:Green Leisure(202) TJ:Transit Junction(161)	155	NR-CS (14818) NR-GL (8328) NR-TJ (11736)	10

Table: Five dataset created and used for empirical experiments.
The table shows the details of each dataset.

Evaluation Metric

- **Cosine Distance(COD):**

$$\mathrm{COD}(y, \hat{y}) = 1 - \frac{y \cdot v}{\|y\|\|\hat{y}\|}$$

- **Canberra Distance(CAD):**

$$\mathrm{CAD}(y, \hat{y}) = \sum_i \frac{|y_i - \hat{y}_i|}{|y_i| + |\hat{y}_i|}$$

- **Chebyshev Distance(CHD):**

$$\mathrm{CHD}(y, \hat{y}) = \max_i |y_i - \hat{y}_i|$$

- **Clark Distance (CLD):**

$$\mathrm{CLD}(y, \hat{y}) = \sqrt{\sum_i \frac{(y_i - \hat{y}_i)^2}{(y_i + \hat{y}_i)^2}}$$

- **Intersection Score(IND):**

$$\mathrm{IND}(y, \hat{y}) = \sum_i \min(y_i, \hat{y}_i)$$

- **Kullback-Leibler Divergence(KL):**

$$\mathrm{KL}(y, \hat{y}) = \sum_i y_i \cdot \log \frac{y_i}{\hat{y}_i},$$

Metric used to evaluate model performance, same from normal label distribution learning.

Experiment Results

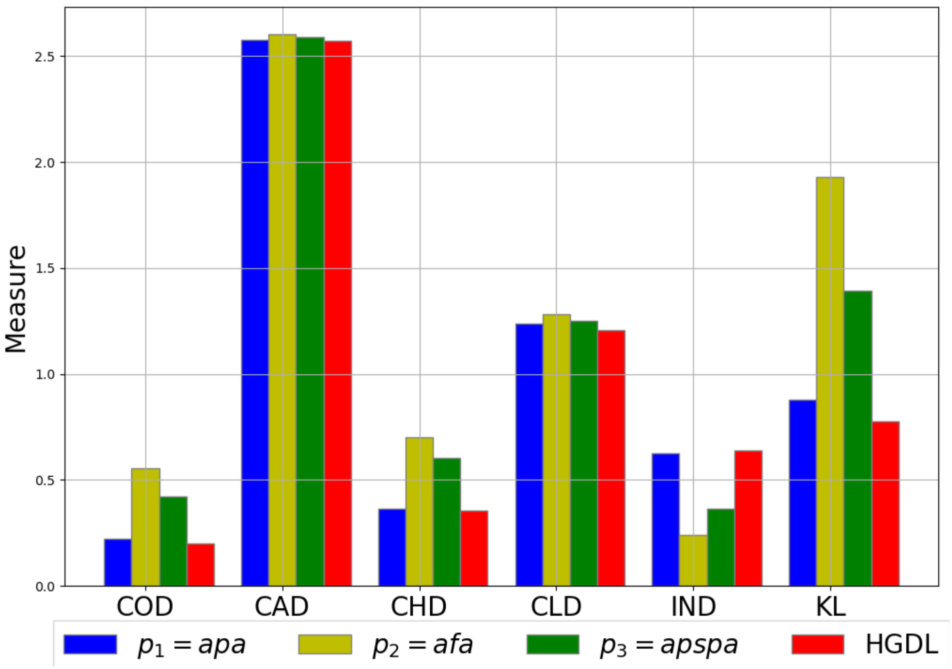


Figure 3: an ablation result of HGDL methods versus its variants on each single meta-path. All metrics except IND is better for lower values.

- The detailed results compared with each benchmark are shown in the paper.
- We provide an ablation result shown on the left to confirm the design of active graph topology homogenization.
  - Our proposed **HGDL** integrates information from different meta-paths and therefore lead to better results compared with each single meta-path for all metrics, proving our graph topology homogenization successfully integrates different meta-path information through attention mechanism.

Key Takeaway

- Success of HGDL shows the importance of learning a proper graph topology before message passing.

FLORIDA ATLANTIC UNIVERSITY

WILLIAM & MARY
CHARTERED 1693

NEURAL INFORMATION PROCESSING SYSTEMS

Reference

- [1] T. N. Kipf and M. Welling, “Semi-supervised classification with graph convolutional networks ,”in ICLR, 2017.
- [2] Y. Jin, R. Gao, Y. He, and X. Zhu, “Gldl: Graph label distribution learning,” in Proceedings of the 38th Annual AAAI Conference on Artificial Intelligence,2024
- [3] W. Xiao, J. Houye, S. Chuan, W. Bai, C. Peng, Y. P., and Y. Yanfang, “Heterogeneous graph attention network,” WWW, 2019.
- [4] X. Yang, M. Yan, S. Pan, X. Ye, and D. Fan, “Simple and efficient heterogeneous graph neural network,” Proceedings of the AAAI Conference on Artificial Intelligence, pp. 10816–10824,2023.
- [5] Qiheng Mao, Zemin Liu, Chenghao Liu, and Jianling Sun. "Hinormer: Representation learning on heterogeneous information networks with graph transformer", in Proceedings of the ACM Web Conference, 2023.

Thanks for listening!