A General Model for Learning Node and Graph Representations Jointly

Author

Chaofan Chen, University of Science and Technology of China

Presenter

Chaofan Chen

Motivation

- My work focuses on two fundamental graph recognition tasks: **node classification** and **graph classification**. For the topological graph, local and global structure learning are very **correlated** but usually studied **separately**.
- Benefits of learning the local and global structures jointly:
- (1) The global information content of the entire graph can provide good contexts to learn the **globally relevant** node representation;
- (2) The local features can be utilized to capture the hierarchical information for generating the entire graph representation (**locally relevant**)

Method

- **1. Initial Node Feature**

$$H^{(k)} = mp(A^{(k-1)}, H^{(k-1)}, W^{(k-1)}) \tag{1}$$

- **2. Assignment Matrix Generation**
- 2.1. Joint Community Detection

$$p(z=j|w) = \frac{exp(\Phi_w^T \Psi_j)}{\sum_{k=1}^{K} exp(\Phi_w^T \Psi_k)} \tag{2}$$

$$p(c|z=j) = \frac{exp(\Psi_j^T \Phi_c)}{\sum_{i \in V} exp(\Psi_j^T \Phi_i)} \tag{3}$$

$$q(z=j|w,c) = \frac{exp((\Phi_w \bigodot \Phi_c)^T \psi_j)}{\sum_{i=1}^{K} exp((\Phi_w \bigodot \Phi_c)^T \psi_j)} \tag{4}$$

Method

- 2.2. Dynamic Routing Method

$$u_{j|i} = h_i W_{ij} \quad (5)$$

$$r_i = softmax(r_i) \quad (6)$$

$$h_j = \sum_i r_{ij} u_{j|i} \quad (7)$$

$$r_{ij} = r_{ij} + h_j u_{j|i}^T \quad (8)$$

Method

- 2.3. Locally Relevant Graph Representation Learning

$$Z^{(l+1)} = (CA)^{(l+1)^T} H^{(l)} \tag{9}$$

$$A^{(l+1)} = (CA)^{(l+1)^T} A^{(l)} (CA)^{(l+1)} \tag{10}$$

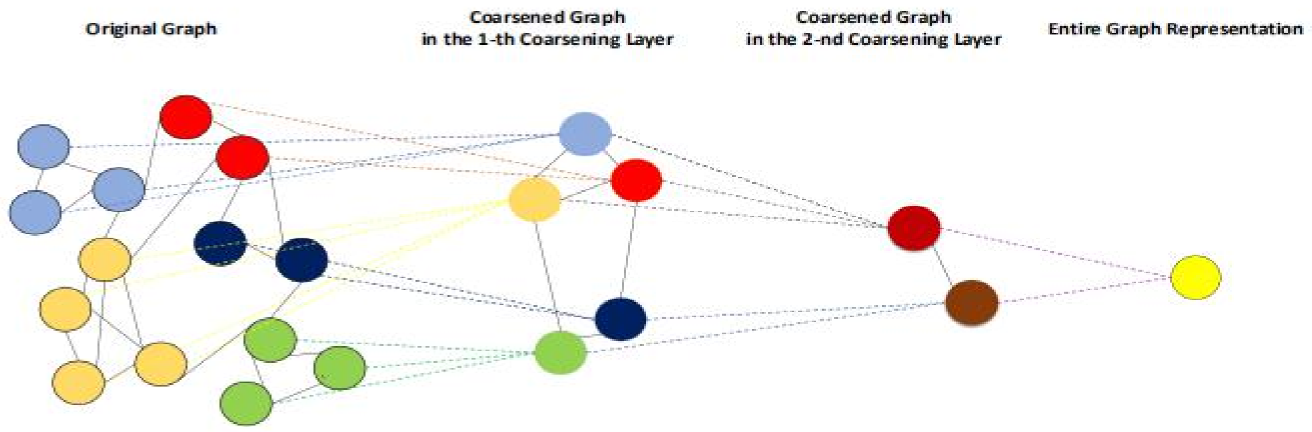

Fig. 1. This figure shows hierarchical clustering over three coarsening layers, where nodes in the latter graph correspond to clusters in the previous graph. Different colors represent different parts/communities of the graph.

Method

- 3. RatioCut Loss

$$L_1 = \sum_{j=1}^{n^{(l+1)}} ((CA)^{(l+1)} L (CA)^{(l+1)^T})_{jj} = tr((CA)^{(l+1)} L (CA)^{(l+1)^T}) \qquad (11)$$

- 4. Mutual Information Loss (Globally Relevant Node Representation Learning)

$$L_2 = \frac{1}{N+M} (\sum_{i=1}^{N} E_{X,A}[logD(h_i, s)] + \sum_{j=1}^{M} E_{X',A'}[log(1 - D(h_j^{'}, s))]) \qquad (12)$$

Experiments

- 1. Datasets

	Nodes	Edges	Classes	Features	Label Rate	Training Nodes	Validation Nodes	Testing Nodes
Cora	2708	5429	7	1433	0.052	140	500	1000
Citeseer	3327	4732	6	3703	0.036	120	500	1000
Pubmed	19717	44338	3	500	0.003	60	500	1000

TABLE I

STATISTICS OF THE DATASETS USED IN OUR NODE CLASSIFICATION TASK.

	Graphs	Nodes	Average Nodes	Average Edges	Classes
ENZYMES	600	19580	32.63	62.14	6
D&D	1178	334925	284.32	715.66	2
COLLAB	5000	372450	74.49	2457.78	3
PROTEINS	1113	43471	39.06	72.82	2

TABLE III

STATISTICS OF THE DATASETS USED IN OUR GRAPH CLASSIFICATION TASK.

Experiments

- 2. Results

	Cora	Citeseer	Pubmed
Raw Features	$47.9 \pm 0.4\%$	$49.3 \pm 0.2\%$	$69.1 \pm 0.3\%$
ManiReg	59.5%	60.1%	70.7%
SemiEmb	59.0%	59.6%	71.1%
LP	68.0%	45.3%	63.0%
DeepWalk	67.2%	43.2%	65.3%
DeepWalk+Features	$70.7 \pm 0.6\%$	$51.4 \pm 0.5\%$	$74.3 \pm 0.9\%$
ICA	75.1%	69.1%	73.9%
Planetoid	75.7%	64.7%	77.2%
Chebyshev	81.2%	69.8%	74.4%
MoNet	$81.7 \pm 0.5\%$	-	$78.8 \pm 0.3\%$
GCN	81.5%	70.3%	79.0%
DGI	$82.3 \pm 0.6\%$	$71.8 \pm 0.7\%$	$76.8 \pm 0.6\%$
Ours-JCD	$83.2 \pm 0.4\%$	$73.1 \pm 0.3\%$	$77.6 \pm 0.5\%$
Ours-DR	$83.6 \pm 0.3\%$	$72.8 \pm 0.4\%$	$77.1 \pm 0.6\%$

TABLE II
NODE CLASSIFICATION RESULTS.

	ENZYMES	D&D	COLLAB	PROTEINS
GRAPHLET	41.03%	74.85%	64.66%	72.91%
SHORTEST-PATH	42.32%	78.86%	59.10%	76.43%
WL	53.43%	74.02%	78.61%	73.76%
WL-OA	60.13%	79.04%	80.74%	75.26%
PATCHYSAN	-	76.27%	72.60%	75.00%
DGK	-	-	73.09%	71.68%
GRAPHSAGE	54.25%	75.42%	68.25%	70.48%
ECC	53.50%	74.10%	67.79%	72.65%
CapsGNN	60.34%	79.55%	80.53%	79.91%
SET2SET	60.15%	78.12%	71.75%	74.29%
SUMPOOL	47.33%	78.72%	69.45%	76.26%
SORTPOOL	57.12%	79.37%	73.76%	75.54%
SAGPOOL	64.17%	81.03%	73.28%	78.82%
DIFFPOOL	64.23%	81.15%	75.50%	78.10%
STRUCTPOOL	63.83%	84.19%	74.22%	80.36%
SHGC	64.17%	78.59%	75.54%	75.46%
Ours-JCD	69.50%	85.69%	74.03%	83.12%
Ours-DR	67.28%	85.34%	72.71%	83.39%

TABLE IV
GRAPH CLASSIFICATION RESULTS.

Thanks for your listening!