Semi-supervisedly Co-embedding Attributed Networks

Zaiqiao Meng[1,2], Shangsong Liang[2], Jinyuan Fang[2], Teng Xiao[2]
[1] *University of Glasgow*, [2] *Sun Yat-sun University*

SVAE for Heterogeneous Data

Let $\mathcal{O}_{ij} = (\mathbf{x}_i^g, \mathbf{x}_j^h, r_{ij}, \mathbf{Y}^l)$ be an atomic data point of the heterogeneous data where $\mathbf{Y}^l$ are the labels of entities, $\mathcal{Z}_{ij} = (\mathbf{z}_i^g, \mathbf{z}_j^h, \mathbf{Y}^u)$ be the collection of latent variables of the two entities where $\mathbf{Y}^u$ are the predicted labels for entities without a label, and $\mathbf{Y} = (\mathbf{Y}^u, \mathbf{Y}^l)$. $\mathcal{F}_{ij} = (\phi(\mathbf{x}_i^g), \phi(\mathbf{x}_j^h))$ is the conditional variables of variational posterior where ϕ is a function taking entities' feature as input for filtering posteriors. The ELBO of $\mathcal{O}_{ij}$ is:

$$\log p(\mathcal{O}_{ij}) \geq \mathbb{E}_{q_{\boldsymbol{\phi}}(\mathcal{Z}_{ij}|\mathcal{F}_{ij})}[\log p_{\boldsymbol{\theta}}(r_{ij} \mid \mathcal{Z}_{ij}, \mathbf{Y})] - KL(q_{\boldsymbol{\phi}}(\mathcal{Z}_{ij} \mid \mathcal{F}_{ij}) \mid p(\mathbf{z}_i^g)p(\mathbf{z}_j^h)) \triangleq -\mathcal{L}(\mathcal{O}_{ij}), \quad (1)$$

Incorporating an discriminative component leads to:

$$\mathcal{J}^{\alpha}(\mathcal{O}) = \sum_{r_{ij} \in \mathcal{R}} \mathcal{L}(\mathcal{O}_{ij}) + \alpha \cdot \mathbb{E}_{\widetilde{p}_l(\mathbf{x},\mathbf{y})}\left[-\log q_{\phi}(\mathbf{y}|\mathbf{x})\right]. \quad (2)$$

Problem Formulation

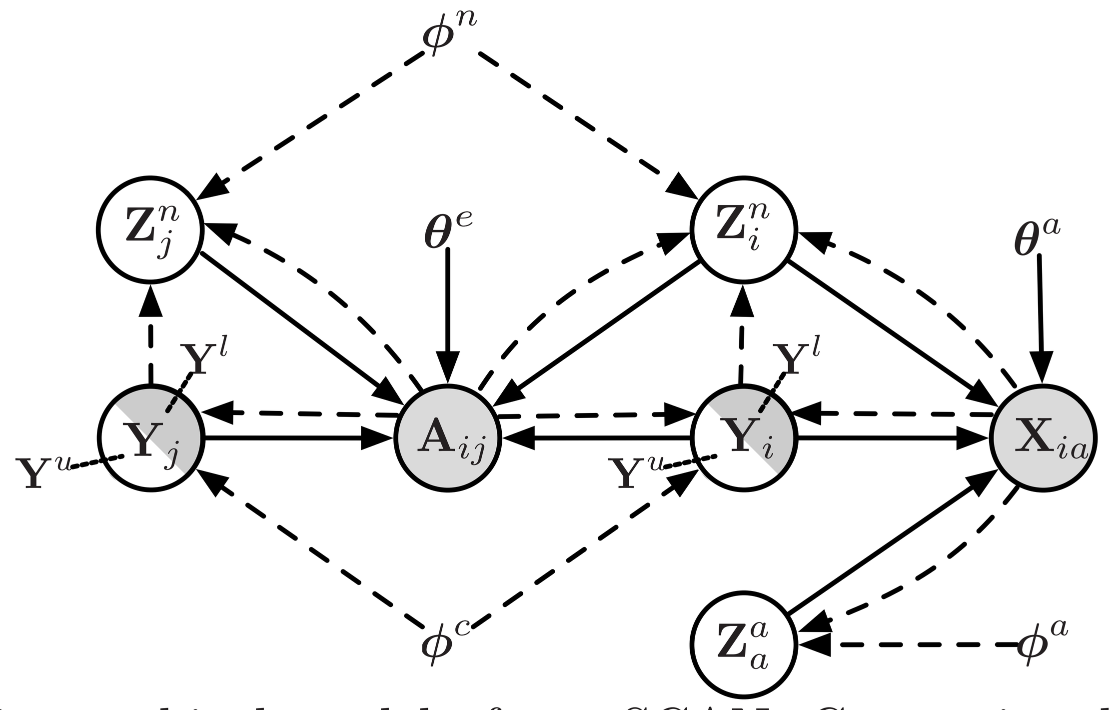
Given a **Partially Labelled Attributed Network** $\mathcal{G} = (\mathcal{V}, \mathcal{A}, \mathbf{A}, \mathbf{X}, \mathbf{Y}^l)$ ($\mathcal{V}$ and $\mathcal{A}$ are the node and attribute sets, $\mathbf{A}$ and $\mathbf{X}$ are the weighted adjacency and node attribute matrices) learn a mapping function Ξ:

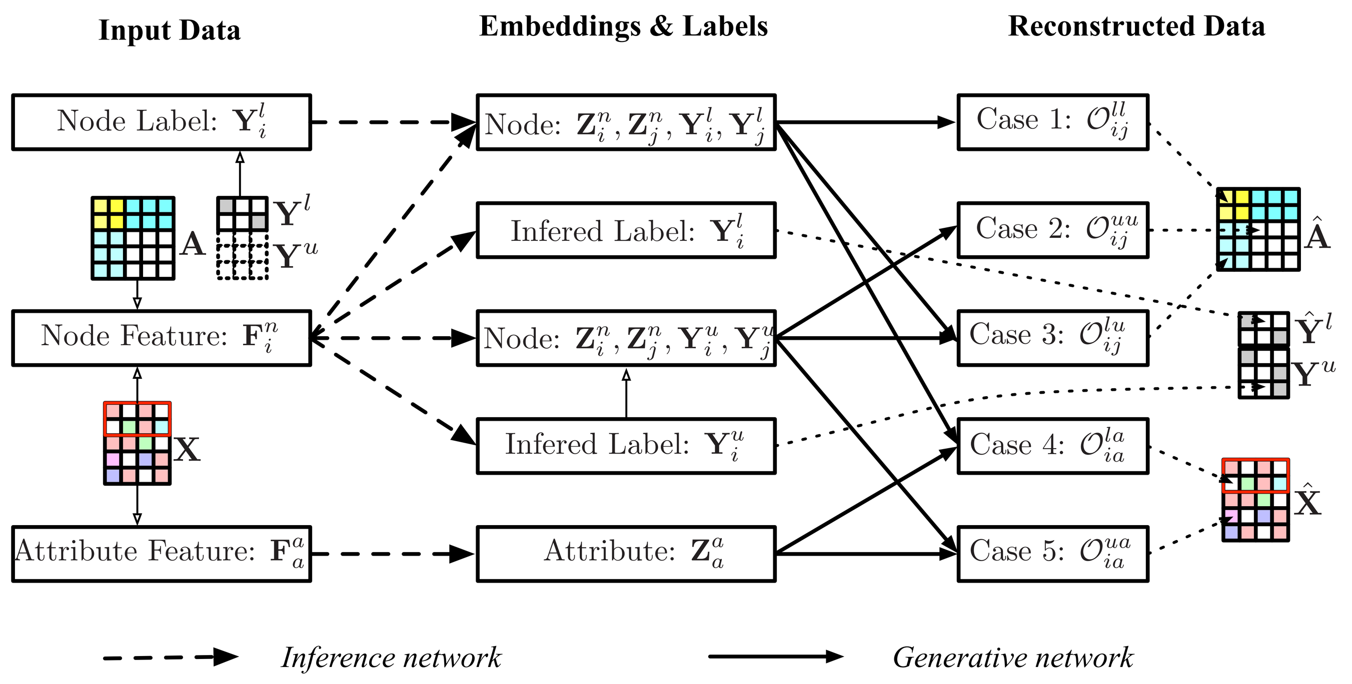
$$\mathcal{G} = (\mathcal{V}, \mathcal{A}, \mathbf{A}, \mathbf{X}, \mathbf{Y}^l) \xrightarrow{\Xi} \mathbf{Z}^n, \mathbf{Z}^a, \mathbf{Y}^u, \quad (3)$$

such that both network structure, node attributes and the partial labels can be preserved by $\mathbf{Z}^n$, $\mathbf{Z}^a$ and $\mathbf{Y}^u$, where $\mathbf{Z}^n \in \mathbb{R}^{N\times D}$ and $\mathbf{Z}^a \in \mathbb{R}^{M\times D}$ represent the latent representation matrices for all the nodes and attributes, respectively, and $\mathbf{Y}^u$ represents the learned labels for all the unlabelled nodes. Here D is the size of the embeddings.

The Semi-supervised Co-embedding Model (SCAN)

(a)





(b)

(a) Probabilistic graphical model of our SCAN. Generative dependencies are solid arrows, while inferences are in dashed arrows. **(b)** The elements in both of matrices (i.e. the edges between nodes and the attribute values of nodes) can be categorized into five cases: (case 1) an edge connecting two labelled nodes; (case 2) an edge connecting two unlabelled nodes; (case 3) an edge connecting a labelled node and an unlabelled node; (case 4) an attribute value associating with a labelled node and an attribute; (case 5) an attribute value associating with an unlabelled node and an attribute.

The total loss can be obtained by combining the ELBOs for these five types of atomic observations with a discriminative loss:

$$\mathcal{J}(\mathbf{A}, \mathbf{X}, \mathbf{Y}^l) = \sum_{i,j \in \mathcal{V}^l} \mathcal{L}(\mathcal{O}_{ij}^{ll}) + \sum_{i \in \mathcal{V}^u, j \in \mathcal{V}^u} \mathcal{U}(\mathcal{O}_{ij}^{uu}) + \sum_{i \in \mathcal{V}^l, j \in \mathcal{V}^u} \mathcal{M}(\mathcal{O}_{ij}^{lu}) + \sum_{i \in \mathcal{V}^l, a \in \mathcal{A}} \mathcal{B}(\mathcal{O}_{ia}^{la}) + \sum_{i \in \mathcal{V}^u, a \in \mathcal{A}} \mathcal{C}(\mathcal{O}_{ia}^{ua}) + \alpha \cdot \mathbb{E}_{v \sim \mathcal{V}^l}[-\log q_{\boldsymbol{\phi}^c}(\mathbf{Y}_v \mid \phi(\mathbf{F}_v^n))] \quad (4)$$

Semi-supervised node classification

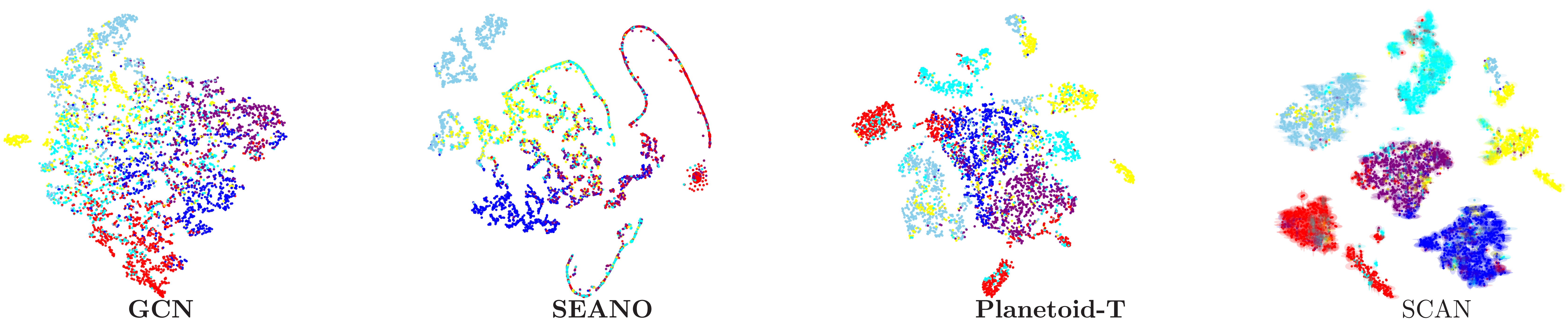
Method	Pubmed			Flickr			BlogCatalog		
	Ma_F1	Mi_F1	ACC	Ma_F1	Mi_F1	ACC	Ma_F1	Mi_F1	ACC
SEANO	.841	.845	.850	.738	.748	.741	.627	.635	.648
Planetoid-T	.815	.825	.823	.721	.743	.733	.803	.811	.817
GCN	.838	.847	.850	.286	.291	.309	.509	.527	.538
GraphSGAN	.839	.842	.841	.697	.715	.702	.698	.703	.719
SCAN_SVM	.847†	.852†	.851	.747†	.750	.747†	.820†	.829†	.835†
SCAN_DIS	**.850**†	**.858**†	**.862**†	**.749**†	**.753**†	**.750**†	**.829**†	**.832**†	**.839**†

SCAN_SVM: SVM trained on embeddings. **SCAN_DIS:** Discriminative network.

Attribute Inference

Method	Pubmed		Flickr		BlogCatalog	
	AUC	AP	AUC	AP	AUC	AP
EdgeExp	.586	.576	.678	.685	.684	.744
SAN	.579	.572	.653	.660	.694	.710
BLA	.622	.602	.730	.769	.787	.792
CAN	.670	.652	.867	.868	.867	.865
SCAN	**.713**†	**.682**†	**.874**†	**.871**	**.893**†	**.895**†

Network Visualization



Contributions

1. We showed how the Semi-supervised Variational Auto-encoders can be generalized to the heterogeneous data.
2. We proposed a semi-supervised co-embedding model (called SCAN) to collaboratively learn representations of nodes and attributes in the same space, meanwhile, a discriminator is trained to generalize from small labelled data to large unlabelled ones.

Acknowledgments

This research was partially supported by the National Natural Science Foundation of China (Grant No. 61906219).