Summary Questions of the lecture

- Describe the key aspects of Graph Diffusion-Embedding Networks.
- GDEN formulates three difusion operators $\mathcal{F}_d(A, H)$ from RWR, LapReg, and NLapReg. In particular, LapReg, and NLapReg are formulated from the Laplacian smoothing and regularization for personalized teleport. The graph diffusion operator is used for message passing (or aggregation) in GDEN and the embedding is conducted by the normal convolution on the aggregated graph signals.

Model	Diffusion operator $\mathcal{F}_d(\mathbf{A}, \mathbf{H})$
RWR Eq.(9)	$(1-\lambda)(\mathbf{I}-\lambda\mathbf{A}\mathbf{D}^{-1})^{-1}\mathbf{H}$
LapReg Eq.(11)	$\lambda(\mathbf{D}-\mathbf{A}+\lambda\mathbf{I})^{-1}\mathbf{H}$
NLapReg Eq.(13)	$(1-\gamma)(\mathbf{I}-\gamma\mathbf{D}^{-\frac{1}{2}}\mathbf{A}\mathbf{D}^{-\frac{1}{2}})^{-1}\mathbf{H}$

$$\mathbf{F}^{(l)} = \mathcal{F}_d(\mathbf{A}, \mathbf{H}^{(l)})$$

$$\mathbf{H}^{(l+1)} = \sigma(\mathbf{F}^{(l)}\mathbf{W}^{(l)})$$

$$\mathbf{f}_i^{(t+1)} = \lambda \sum\nolimits_{j, j\neq i} \mathbf{P}_{ij}\mathbf{f}_j^{(t)} + (1-\lambda)\mathbf{h}_i$$

$$\min_{\mathbf{F}} \ \frac{1}{2}\sum_{i,j=1}^{n} \mathbf{A}_{ij}\|\mathbf{f}_i - \mathbf{f}_j\|_2^2 + \lambda\sum_{i=1}^{n}\|\mathbf{f}_i - \mathbf{h}_i\|_2^2$$

$$\min_{\mathbf{F}} \ \frac{1}{2}\sum_{i,j=1}^{n} \mathbf{A}_{ij}\left\|\frac{\mathbf{f}_i}{\sqrt{\mathbf{d}_i}} - \frac{\mathbf{f}_j}{\sqrt{\mathbf{d}_j}}\right\|_2^2 + \lambda\sum_{i=1}^{n}\|\mathbf{f}_i - \mathbf{h}_i\|_2^2$$

Summary Questions of the lecture

- Describe the key aspects of Graph Diffusion Convolution(GDC).
- GDC genralizes graph diffusion by formulating it as an infinite weighted sum of the powers of the transition probability matrix. The output of this formulation, S, is sparsified by extracting salient connections from each column (e.g. top$-k$, thresholding) and normalizing. The resulting matrix is column stochastic and can be used with various graph convolution methods.

Generalized graph diffusion

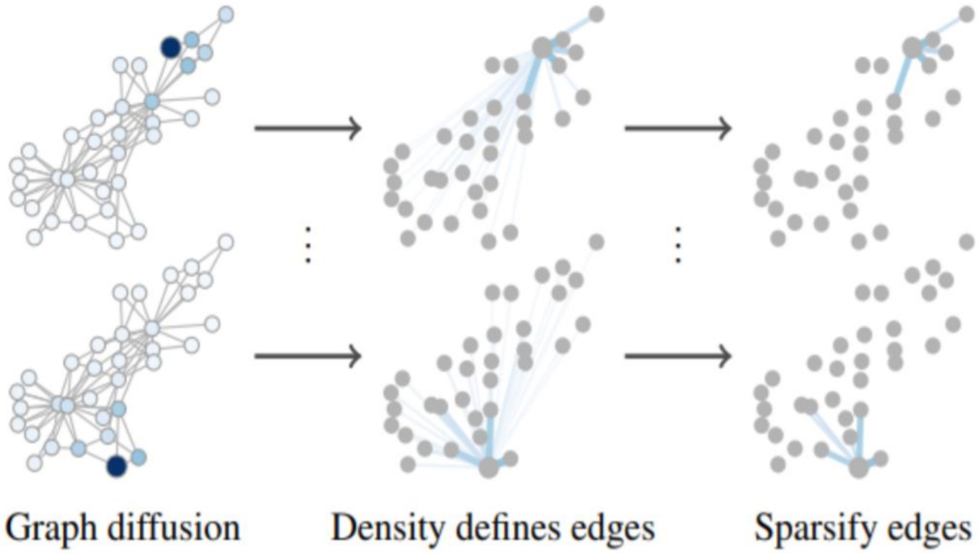
$$S = \sum_{k=0}^{\infty} \theta_k P^k \,,$$

Sparsification is done additionally for $\tilde{S}$

Normalize (symmetric)

$$T_{sym}^{\tilde{S}} = D_{\tilde{S}}^{-1/2} \tilde{S} D_{\tilde{S}}^{-1/2}$$

GCN (ShevNet): $\boldsymbol{F}^{l} = T_{sym}^{\tilde{S}} \boldsymbol{H}^{l}$, $\boldsymbol{H}^{l+1} = \boldsymbol{\sigma}(\boldsymbol{F}^{l}\boldsymbol{\Theta})$

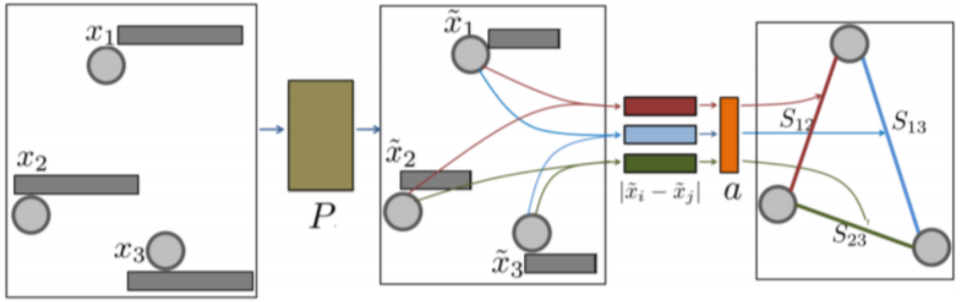


Summary Questions of the lecture

- Describe the key aspects of Graph Learning-Convolutional Networks.
- For SSL, GLCN first predicts the link between nodes and then performs graph convolution. Node connectivity values are assigned similarly as in GAT and trained such that high values are assigned between nodes that have similar features. This property is forced by the graph learning loss proposed in the paper.

$$S_{ij} = g(x_i, x_j) = \frac{A_{ij} \exp(\mathrm{ReLU}(a^T |x_i - x_j|))}{\sum_{j=1}^{n} A_{ij} \exp(\mathrm{ReLU}(a^T |x_i - x_j|))}$$

(when A is not available, $A_{ij} = 1$)



Graph Learning Loss:

$$\mathcal{L}_{\mathrm{GL}} = \sum_{i,j=1}^{n} \|x_i - x_j\|_2^2 S_{ij} + \gamma \|S\|_F^2 + \beta \|S - A\|_F^2$$

$\boxed{a \text{ is trained}}$