Graph Wavelet Neural Network

SeungU Yoo

Seoul National University

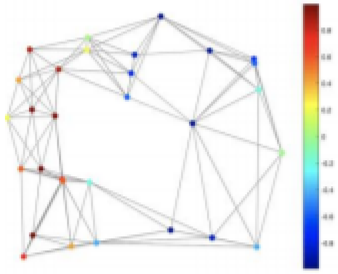
Neural network from Spectral and Graph Fourier Transform

- Simplified ChebNet for GCN

$$\boldsymbol{F}_O = (\widetilde{\boldsymbol{D}}^{-1/2}\widetilde{\boldsymbol{A}}\widetilde{\boldsymbol{D}}^{-1/2})\boldsymbol{F}_I\boldsymbol{\Theta}, \text{ with } \boldsymbol{\Theta} \in \mathbb{R}^{d_1 \times d_2} \text{ and } \boldsymbol{\Theta}[m,n] = \theta^{(mn)}$$

- Spatial view

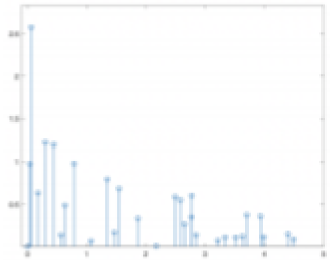
$$\boldsymbol{C} = \widetilde{\boldsymbol{D}}^{-1/2}\widetilde{\boldsymbol{A}}\widetilde{\boldsymbol{D}}^{-1/2} \quad \boldsymbol{F}_O = \boldsymbol{C}\boldsymbol{F}_I\boldsymbol{\Theta}, \text{ with } \boldsymbol{\Theta} \in \mathbb{R}^{d_1 \times d_2}$$



$\hat{\boldsymbol{f}} = \boldsymbol{U}^T\boldsymbol{f}$

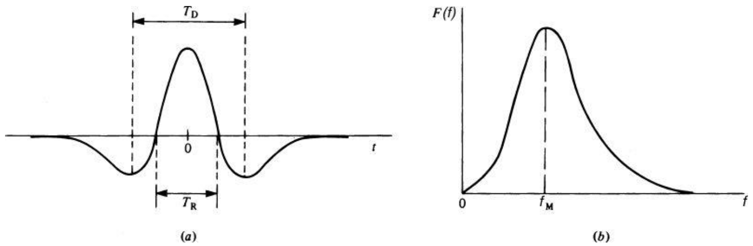
Decompose signal f

Reconstruct signal $\boldsymbol{f}$



What is Wavelet

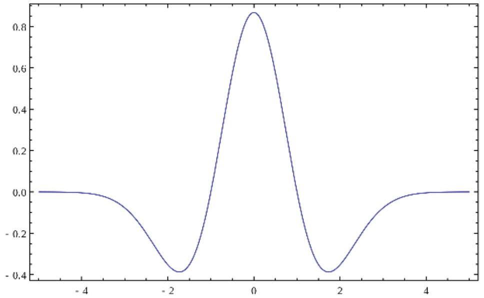
- Wave form



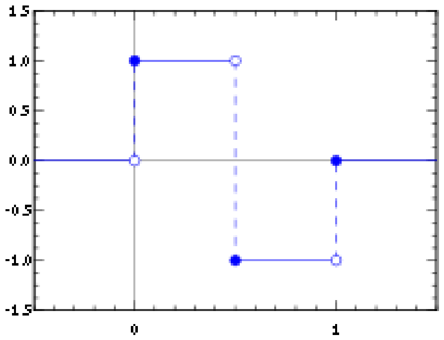
- Orhonormal

$$\int_{\mathbf{R}} \psi_{n_1,k_1}(t)\psi_{n_2,k_2}(t)\,dt = \delta_{n_1,n_2}\delta_{k_1,k_2}$$

- example



Mexican hat wavelet

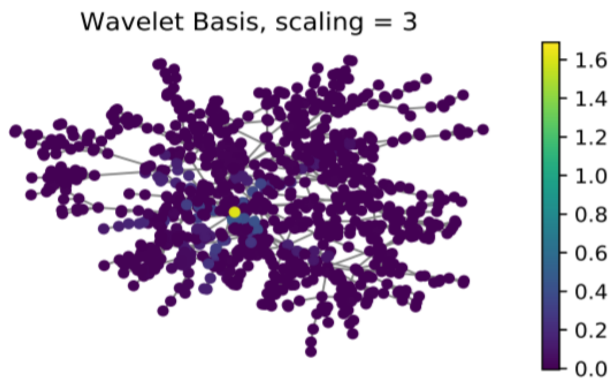


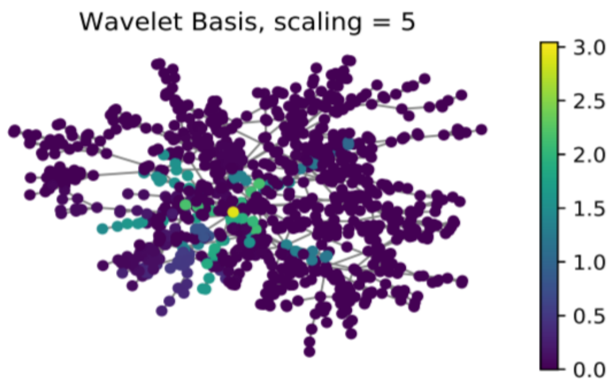
Haar wavelet

Graph Wavelet Transform

- Advantage of Wavelet Transform
  - High efficiency-no eigen-decomposition=>
  - High sparseness => computationally efficient
  - Localized convolution=>reduce parameter=>applicable to large graph
  - Flexible neighborhood(varing scale parameter)

"Wavelets on graphs via spectral ´ graph theory. Applied and Computational Harmonic Analysis"





Graph Wavelet Transform

- Layer and formula

$$\psi_s^{-1}\boldsymbol{f} = \frac{1}{2}c_0\boldsymbol{f} + \sum_{k=1}^{m} c_k T_k'(\boldsymbol{L})\boldsymbol{f}$$

$$T_k'(\boldsymbol{L})\boldsymbol{f} = \frac{2}{a}(\boldsymbol{L}-\boldsymbol{I})(T_{k-1}'(\boldsymbol{L})\boldsymbol{f}) - T_{k-2}'(\boldsymbol{L})\boldsymbol{f} \quad \Rightarrow \quad T_k(y) = cos(k\ arccos(y))$$

$$\text{first layer}: \ \boldsymbol{X}^2_{[:,j]} = \mathrm{ReLU}(\psi_s \sum_{i=1}^{p} \boldsymbol{F}^1_{i,j}\psi_s^{-1}\boldsymbol{X}^1_{[:,i]}) \qquad j = 1, \cdots, q,$$

$$\text{second layer}: \ \boldsymbol{Z}_j = \mathrm{softmax}(\psi_s \sum_{i=1}^{q} \boldsymbol{F}^2_{i,j}\psi_s^{-1}\boldsymbol{X}^2_{[:,,i]}) \qquad j = 1, \cdots, c,$$

- Area complexity

Table 4: Statistics of wavelet transform and Fourier transform on Cora

	Statistical Property	**wavelet transform**	**Fourier transform**
Transform Matrix	Density	2.8%	99.1%
	Number of Non-zero Elements	205,774	7,274,383
Projected Signal	Density	10.9%	100%
	Number of Non-zero Elements	297	2,708

- Performance

Table 3: Results of Node Classification

Method	**Cora**	**Citeseer**	**Pubmed**
MLP	55.1%	46.5%	71.4%
ManiReg	59.5%	60.1%	70.7%
SemiEmb	59.0%	59.6%	71.7%
LP	68.0%	45.3%	63.0%
DeepWalk	67.2%	43.2%	65.3%
ICA	75.1%	69.1%	73.9%
Planetoid	75.7%	64.7%	77.2%
Spectral CNN	73.3%	58.9%	73.9%
ChebyNet	81.2%	69.8%	74.4%
GCN	81.5%	70.3%	79.0%
MoNet	81.7±0.5%	—	78.8±0.3%
GWNN	**82.8%**	**71.7%**	**79.1%**

THANK YOU FOR LISTENING