

Graph Convolutional Networks

Exploiting domain sparsity

$$\mathbf{h} = \mathbf{X} \mathbf{a}$$

Self-attention (I)

$$\{\mathbf{x}_i\}_{i=1}^t = \{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_t\} \rightsquigarrow \mathbf{X} \in \mathbb{R}^{n \times t}, \quad \mathbf{x}_i \in \mathbb{R}^n$$

$$\mathbf{h} = \alpha_1 \mathbf{x}_1 + \alpha_2 \mathbf{x}_2 + \dots + \alpha_t \mathbf{x}_t = \mathbf{X} \mathbf{a} \in \mathbb{R}^n$$

$$\alpha_i > 0$$

$$\mathbf{X} \doteq \begin{bmatrix} | & | & & | \\ \mathbf{x}_1 & \mathbf{x}_2 & \dots & \mathbf{x}_t \\ | & | & & | \end{bmatrix} \in \mathbb{R}^{n \times t}$$

$$\begin{array}{l} \text{soft attention: } \left[\|\mathbf{a}\|_1 = 1 \right. \\ \text{hard attention: } \left[\|\mathbf{a}\|_0 = 1 \right. \end{array}$$

Input layer / samples

$$\mathcal{X} = \{\mathbf{x}^{(i)} \in \mathbb{R}^n \mid \mathbf{x}^{(i)} \text{ is a data sample}\}_{i=1}^m \quad \text{input samples}$$

$$\mathcal{X} = \{ \overset{\text{domain}}{\underset{\text{channels}}{\boldsymbol{x}^{(i)}}} : \Omega \rightarrow \mathbb{R}^c, \boldsymbol{\omega} \mapsto \boldsymbol{x}^{(i)}(\boldsymbol{\omega}) \}_{i=1}^m$$

$$\Omega = \{1, 2, \dots, \overset{\text{total time}}{T/\Delta t}\} \subset \mathbb{N}, \quad c \in \{1, 2, \overset{\text{stereo}}{5+1}, \dots\}$$

sampling interval mono Dolby 5.1

$$\Omega = \{1, \dots, \overset{\text{height}}{h}\} \times \{1, \dots, \overset{\text{width}}{w}\} \subset \mathbb{N}^2, \quad c \in \{1, 3, 20, \dots\}$$

colour
hyperspectral

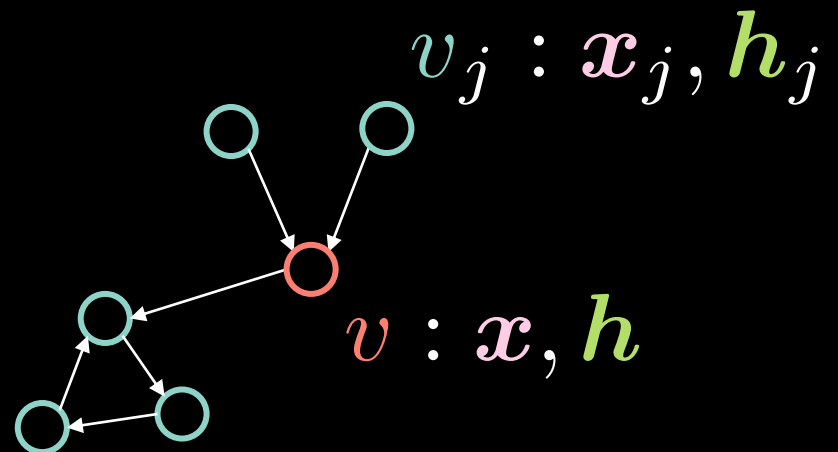
$$\begin{array}{ccc} & \text{four-momentum} & \\ \Omega = \mathbb{R}^4 \times \mathbb{R}^4, & c = 1 & \\ \text{space-time} & \text{Hamiltonian} & \end{array}$$

$$\mathbf{x}(\omega_1, \omega_2) = \begin{pmatrix} r(\omega_1, \omega_2) \\ g(\omega_1, \omega_2) \\ b(\omega_1, \omega_2) \end{pmatrix}$$

GCN

\mathbf{a} : adjacency vector

$$\alpha_j \stackrel{\downarrow}{=} 1 \Leftrightarrow v_j \rightarrow v$$

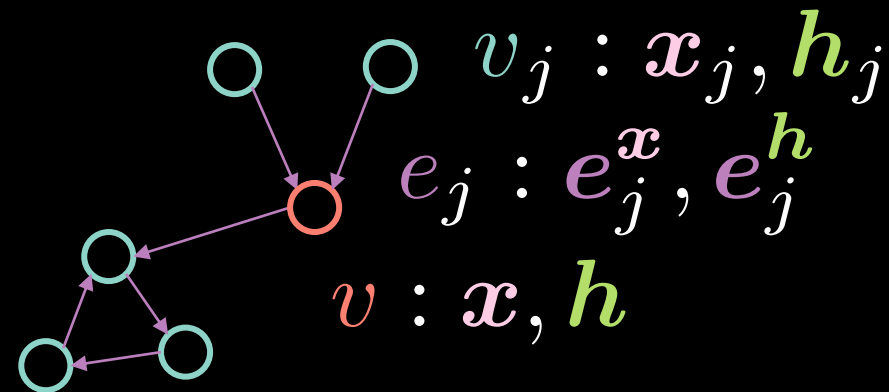


$$d = \|\mathbf{a}\|_1 : \text{degree (\# of incoming edges)}$$

$$\mathbf{h} = f(\mathbf{U}\mathbf{x} + \mathbf{V}\mathbf{X}\mathbf{a}d^{-1}) \quad f(\cdot) : (\cdot)^+, \sigma(\cdot), \tanh(\cdot)$$

$$\{\mathbf{x}_i\}_{i=1}^t \rightsquigarrow \mathbf{H} = f(\mathbf{U}\mathbf{X} + \mathbf{V}\mathbf{X}\mathbf{A}\mathbf{D}^{-1}) \quad \mathbf{D} = \text{diag}(d_i)$$

Residual gated GCN



$$h^{\ell+1} \rightarrow h = x + \left(Ax + \sum_{v_j \rightarrow v} \eta(e_j) \odot Bx_j \right)^+$$

$$\eta(e_j) = \sigma(e_j) \left(\sum_{v_k \rightarrow v} \sigma(e_k) \right)^{-1} h_j^{\ell}$$

$$e_j = Ce_j^x + Dx_j + Ex, \quad e_j^h = e_j^x + (e_j)^+$$