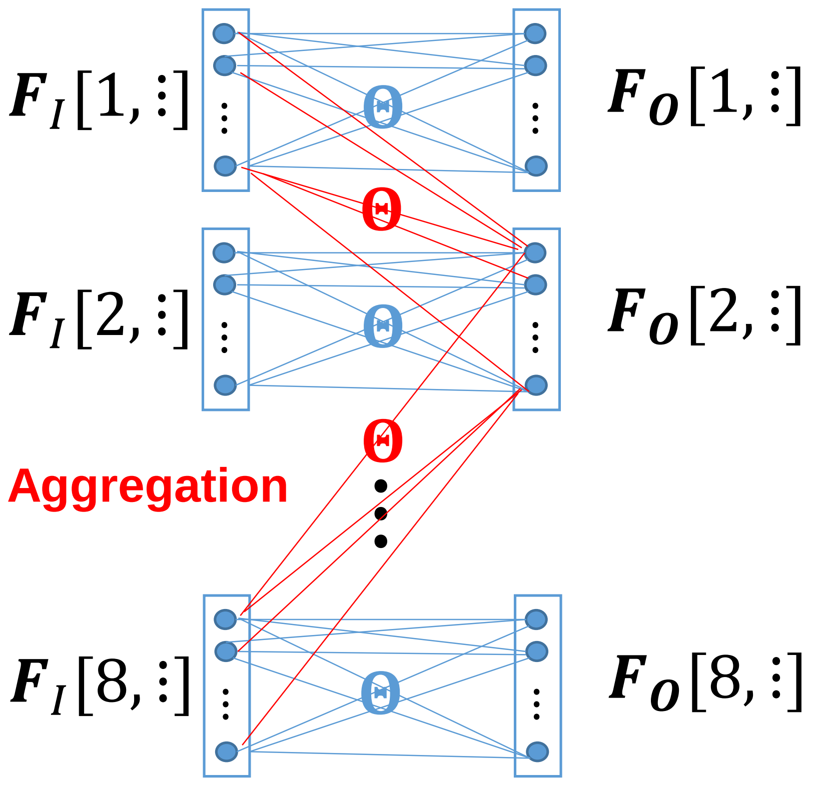
Summary Questions of the lecture

- Present the spatial view of Simplified ShevNet.

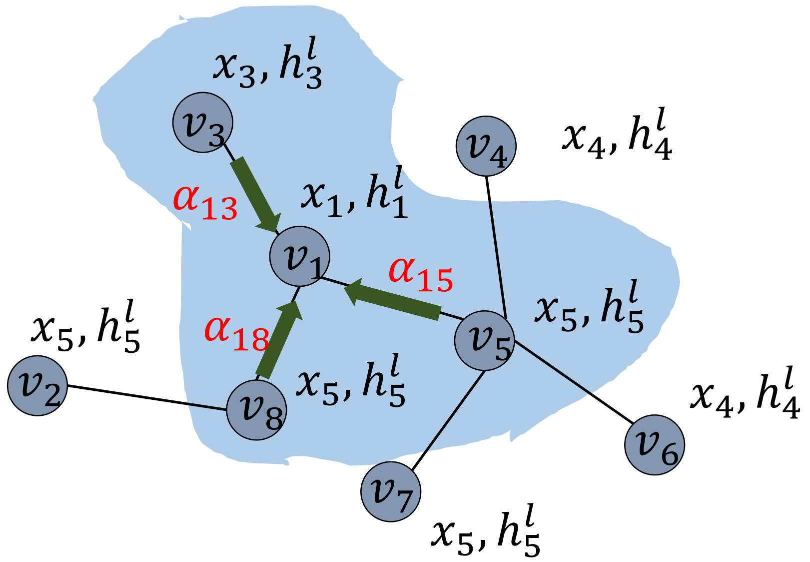
→ From the simplified ChebNet's formulation $F_O = CF_I\Theta$, we observe the output graph signal of the i'th node: $F_O[i,:] = \sum_j C[i,j]F_I[j,:]\Theta$. From the fact that $C[i,j]$ is zero between nodes that are not neighbors, we find that $F_O[i,:]$ is an aggregation from the graph signals of neighbor nodes, weighted by the learnable parameters Θ, which is a spatial smoothing operation and corresponds to a spectral smoothing.



Summary Questions of the lecture

- What is the difference of Simplified ShevNet from a non-graph neural network ?

→ Non-graph neural networks cannot leverage the connectivity information among graph nodes. Thus, the feature of each node (or 'data point' in non-graph networks) is transformed independently. On the other hand, graph neural networks transform each node by aggregating information from connected nodes.



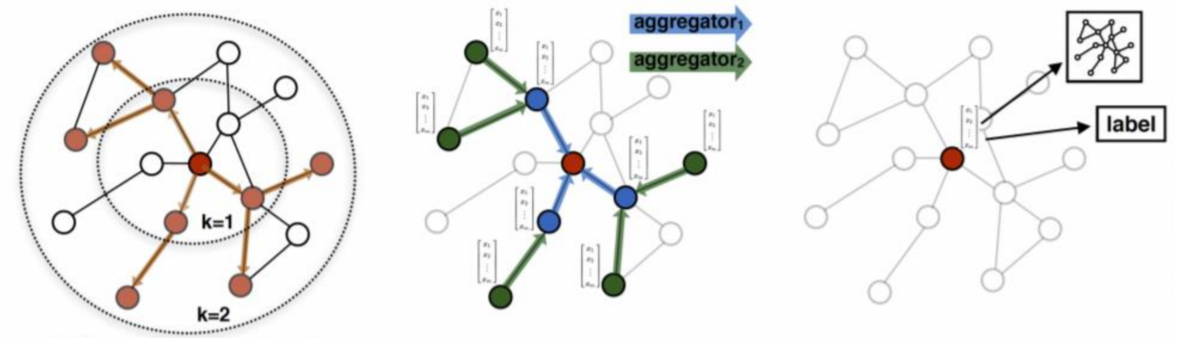
Summary Questions of the lecture

- Explain the key aspects of GraphSAGE (SAmple and aggreGatE).

→ In the context of semi-supervised learning, only aggregating information from nodes of distance 1 has a danger of only encountering unlabeled nodes. Thus, GraphSage tries to sample and aggregate nodes in multi-hop distances. Thus, the resulting message-passing signal is concatenated with the current node's signal and transformed by a learned parameter matrix.

$$\boldsymbol{h}_{N_s(v_i)}^{(\boldsymbol{l}+\mathbf{1})} = \boldsymbol{AGG}\left(\left\{\boldsymbol{h}_{\boldsymbol{j}}^{(\boldsymbol{l})} \mid v_j \in N_s(v_i)\right\}\right)$$

$$\boldsymbol{h}_{i}^{(\boldsymbol{l}+\mathbf{1})} = \boldsymbol{\sigma}\left(\boldsymbol{\Theta} \cdot \left[\boldsymbol{h}_{\boldsymbol{i}}^{(\boldsymbol{l})} \parallel \boldsymbol{h}_{N_s(v_i)}^{(\boldsymbol{l}+\mathbf{1})}\right]\right)$$



1. Sample neighborhood
2. Aggregate feature information from neighbors
3. Predict graph context and label using aggregated information