Deep Learning on Graphs

Hongyang Gao

11/13/2019

Reference: Representation Learning on Networks, snap.stanford.edu/proj/embeddings-www, WWW 2018

Questions

- What's convolution?
-
-
-

Questions

- What's convolution?
- What's the key property of convolution layer?
-
-

Questions

- What's convolution?
- What's the key property of convolution layer?
  - weights sharing, sparse connectivity
- What's the input of a 2D convolution layer?
  - H x W x C
-

Questions

- What's convolution?
- What's the key property of convolution layer?
- What's the input of a 2D convolution layer?
  - H x W x C
- Can we apply a 3D convolution layer on this input (H x W x C)? (H x W x C x 1)?

Questions

- What's convolution?
- What's the key property of convolution layer?
- What's the input of a 2D convolution layer?
  - H x W x C
- Can we apply a 3D convolution layer on this input (H x W x C)?
  - NO!!! Pixels have order on H and W dimensions NOT on C dimension.

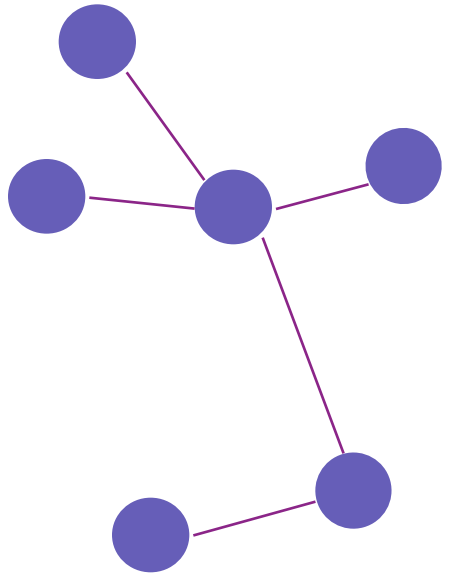
Locality

Sum-up

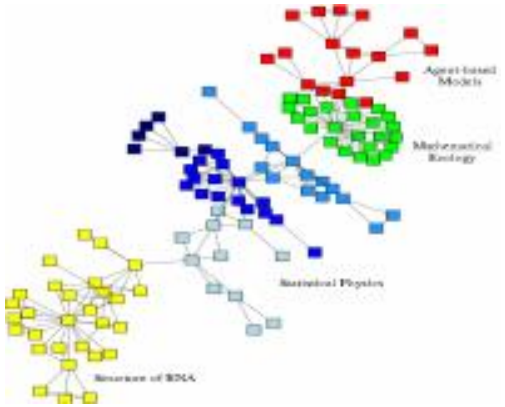
- Locality is important for convolution.
  - Only apply conv on spatial dimensions (H and W)
- How about pooling layer? Is locality important?
- How about data without locality information?

Graph

- Nodes
  - Each node can have features
- Edges
  - Directed or undirected

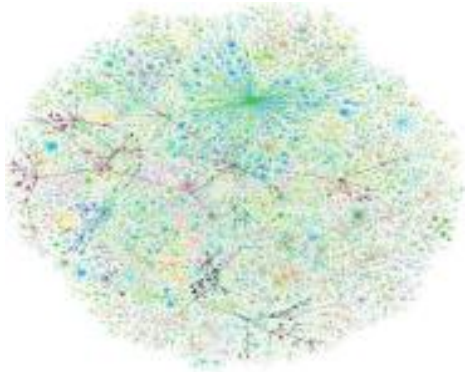


Graph

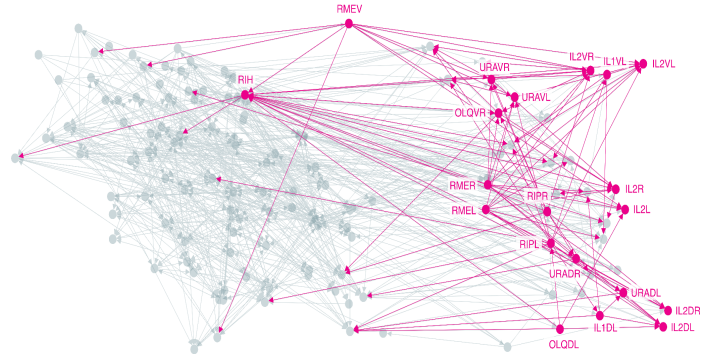


Economic networks

Social networks



Internet



Networks of neurons

Machine Learning with Graphs

Traditional ML tasks in networks:

- Node classification
  - Predict a type of a given node
- Link prediction
  - Predict whether two nodes are linked
- Graph classification
  - Predict a type of a given graph

Migrate the success of deep learning to graph

- Can we?

- Easy or hard?

Image VS. Graph

- Is image a graph?

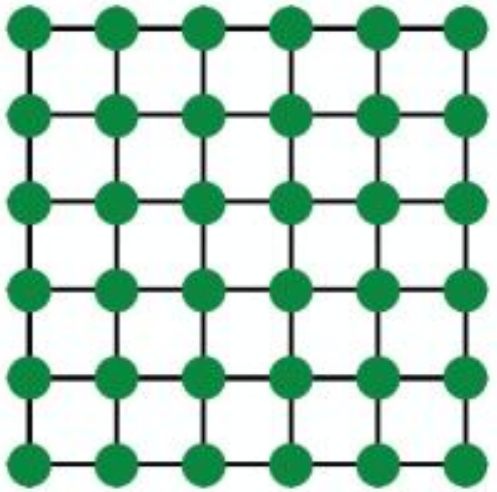
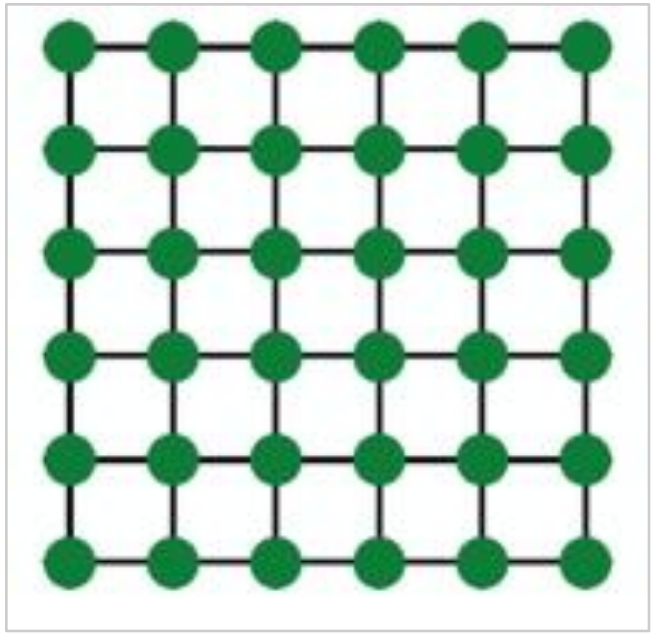
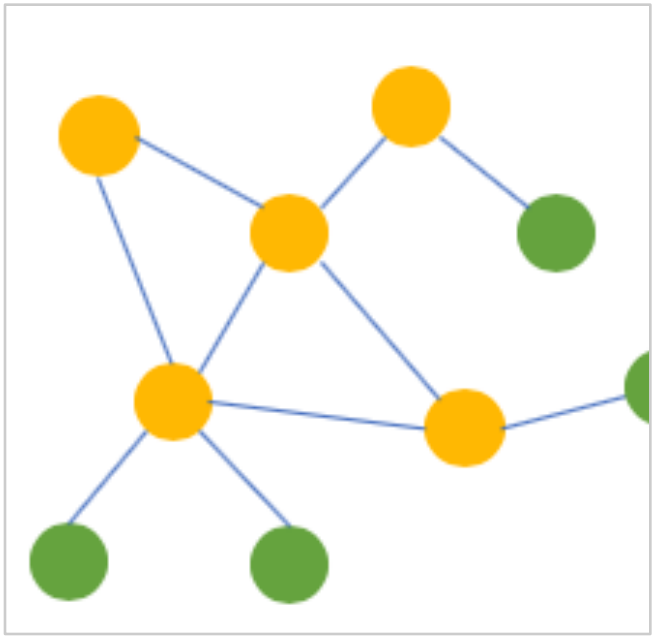


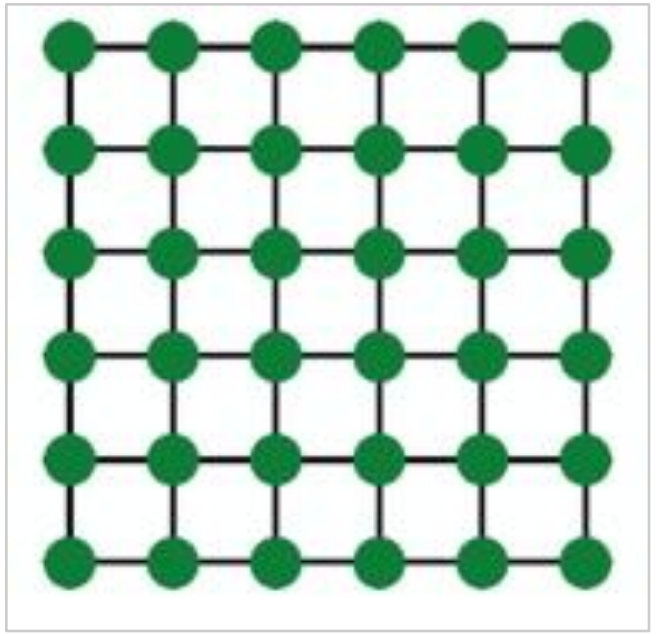
Image VS. Graph

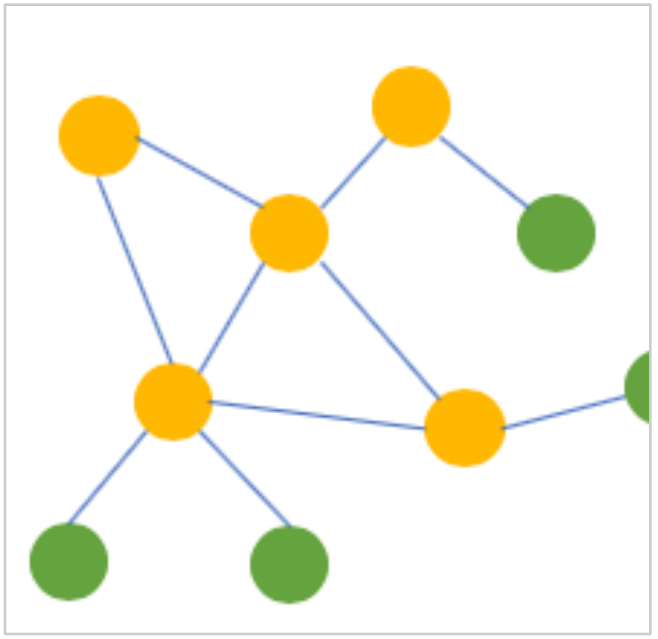




What's the difference between image and graph in terms of structure? (name at least 2)

Image VS. Graph

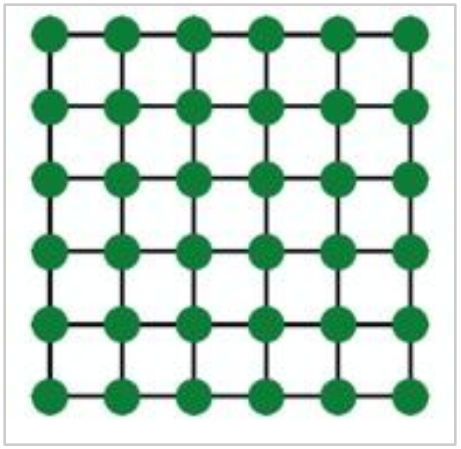


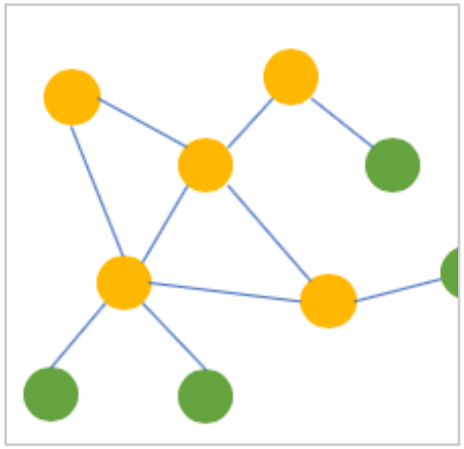


What's the difference between image and graph in terms of structure?

- The number of neighboring nodes are fixed on image but not on graph
- The neighboring nodes on image are ordered by their relative positions but not on graph

How different?



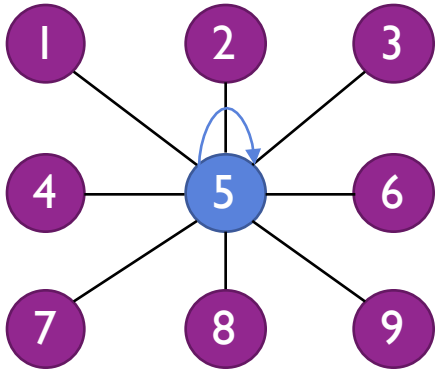


- Graph is not grid-like data
  - There is no fixed-number of neighboring nodes
  - The neighboring nodes are not ordered

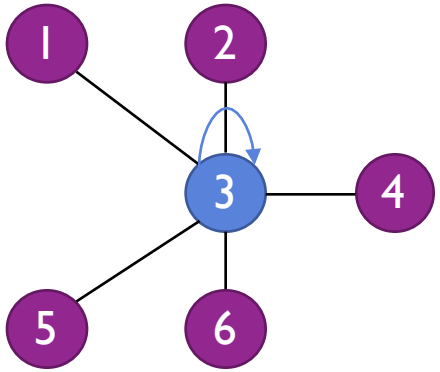
ANY IDEA?

How different?

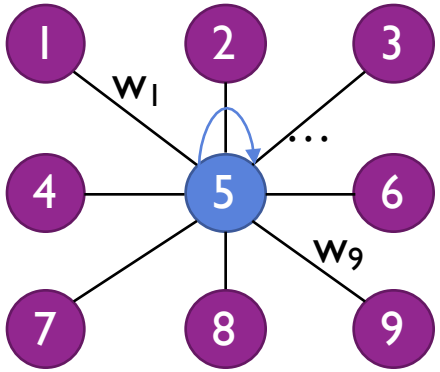
Image



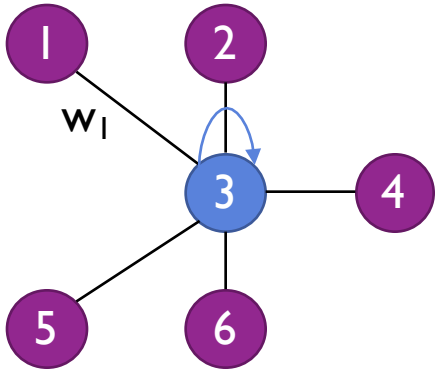
Graph



- $W = [w_1, w_2, w_3, w_4, w_5, w_6, w_7, w_8, w_9]$ # Trained weights
- How to assign these weights for image and graph?



Image



Graph

$$g_{\downarrow}conv(H, A) = AHW$$

$A \in R^{n \times n}$: is adjacency matrix

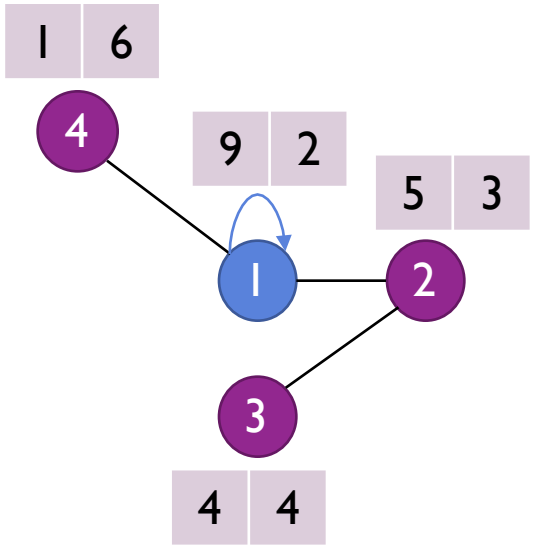
$H \in R^{n \times c}$: is feature matrix

$W \in R^{c \times d}$: is a weights matrix

What does n mean?
What does c mean?
What does d mean?

Graph Convolutional Networks

$$g_{\downarrow}conv(H, A) = AHW$$



	1	2	3	4
1	1	1	0	1
2	1	0	1	0
3	0	1	0	0
4	1	0	0	0

A

1	9	2
2	5	3
3	4	4
4	1	6

H

Calculate for node 1:

1	1	1	0	1

9	2	1
5	3	2
4	4	3
1	6	4

= 1*9 + 1*5 + 0*4 + 1*1 = 15
= 1*2 + 1*3 + 0*4 + 1*6 = 11

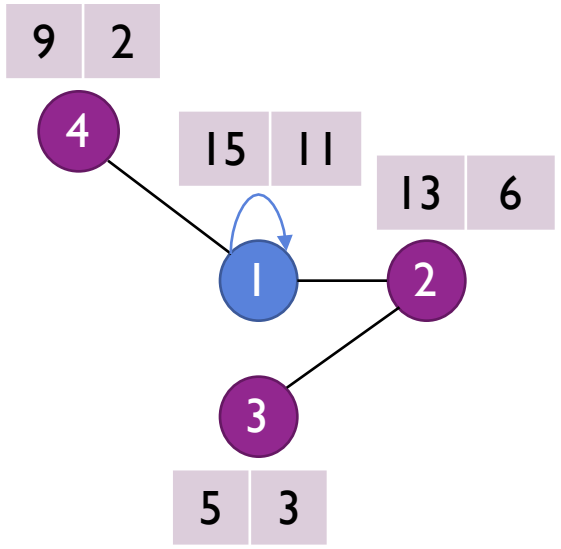
15	11	1
13	6	2
5	3	3
9	2	4

AH

Graph Convolutional Networks

AH: Summation from neighboring nodes

$$g_{\downarrow}conv(H, A) = AHW$$



AH

1	15	11
2	13	6
3	5	3
4	9	2

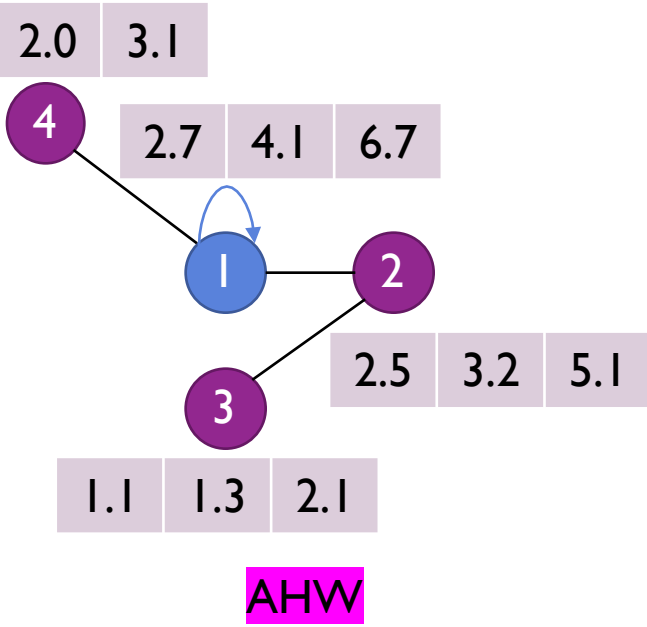
H'

0.1	0.2	0.3
0.2	0.1	0.2

W

=

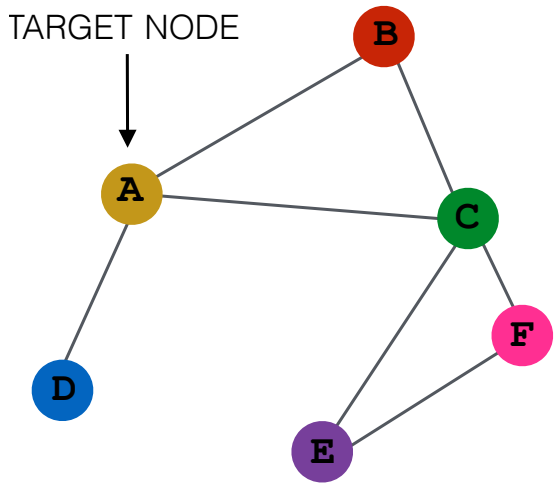
2.7	4.1	6.7	1
2.5	3.2	5.1	2
1.1	1.3	2.1	3
1.3	2.0	3.1	4



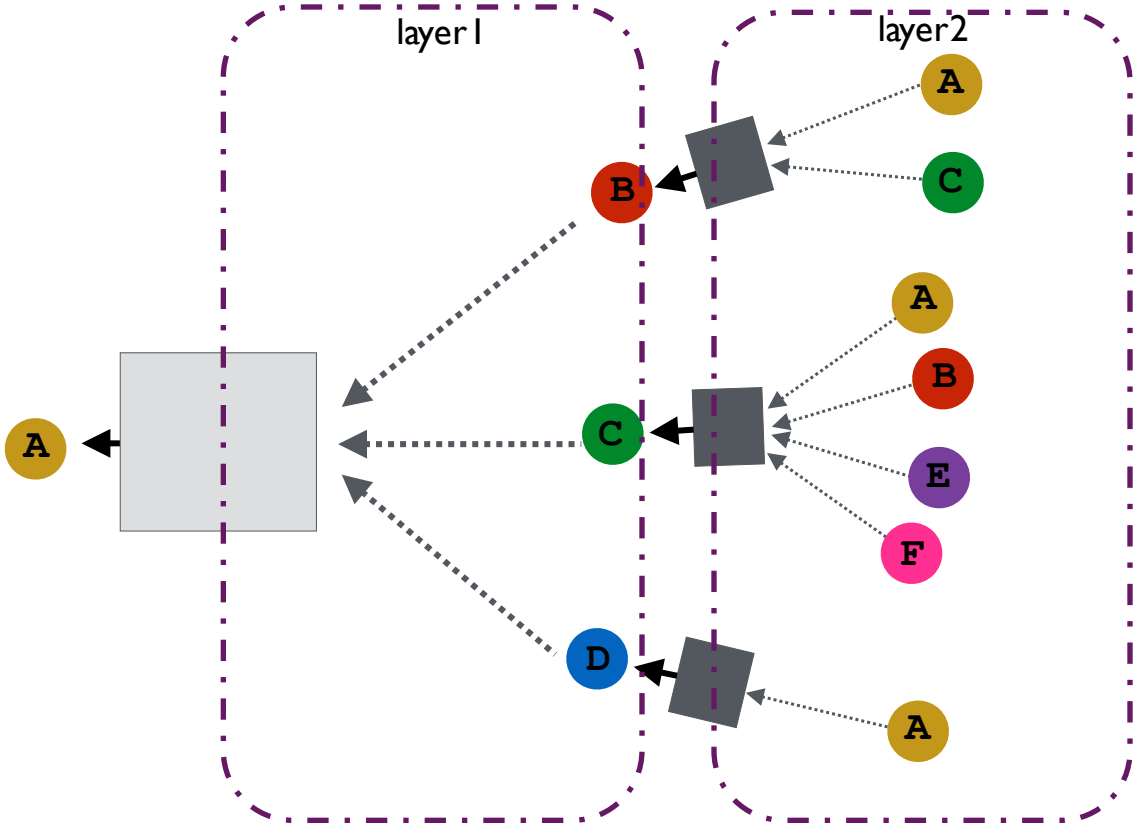
AHW

Graph Convolutional Networks

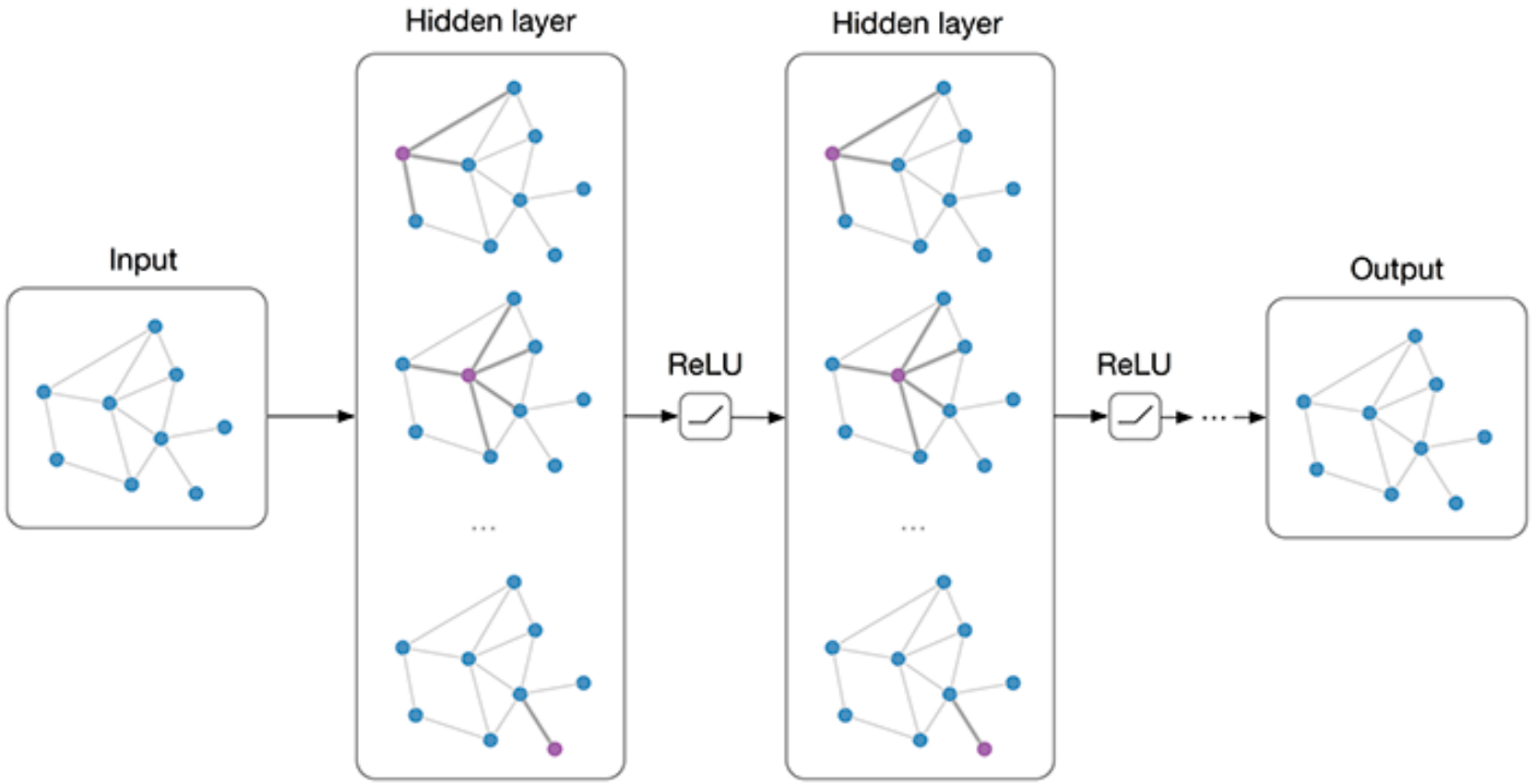
AHW: Summation from neighboring nodes then projection



INPUT GRAPH



Neighborhood Aggregation



Graph Convolutional Networks

Graph Convolutional Networks

- Some questions about GCNs:
  - Where is weights sharing?
  - How is this related to convolution layer or fully-connected layer?
  - What problem does GCN solve?
    - Node classification? Link prediction? Graph classification?
  - How to apply GCN to graph classification tasks?
  - **Any limitations?**

Deep Learning on Graphs (2)

Hongyang Gao

11/15/2019

Questions

- What are challenges of applying deep learning on graphs?
-
-
-

Questions

- What are challenges of applying deep learning on graphs?
- What's formula of GCN?
-
-

Questions

- What are challenges of applying deep learning on graphs?
- What's formula of GCN?

$$g_{\downarrow}conv(H, A) = AHW$$

- What does AH mean?
-

Questions

- What are challenges of applying deep learning on graphs?
- What's formula of GCN?

$$g_{\downarrow}conv(H, A) = AHW$$

- What does AH mean?
- What does AHW mean?

Why pooling

- Why we need pooling?

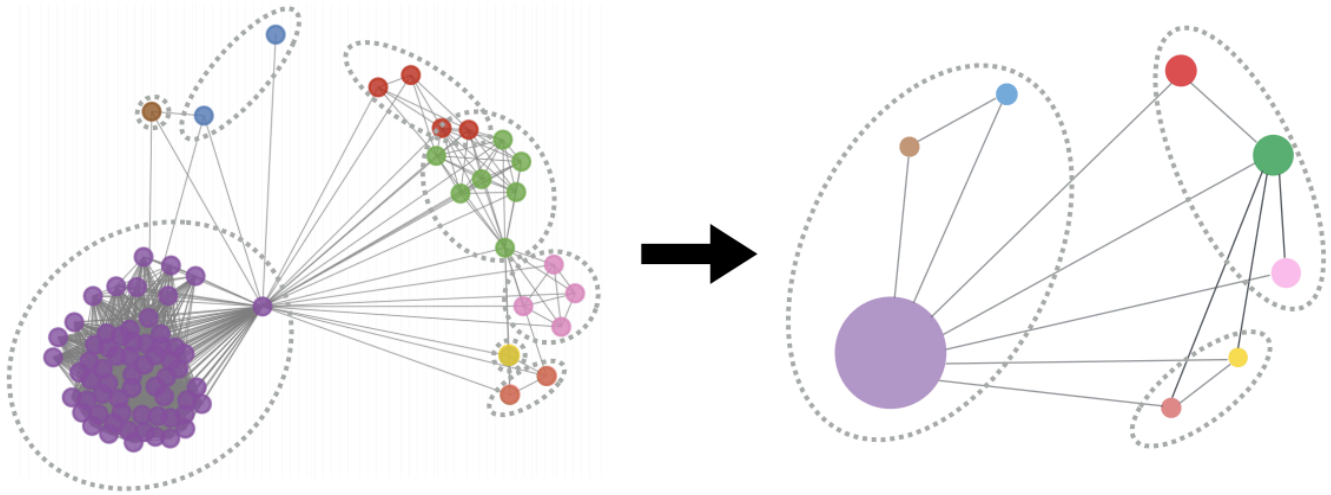
Why pooling

- Why we need pooling?
  - Reduce size
  - Enlarge receptive field
  - Non-linearity (max pooling)

Pooling on graph

- Reduce node numbers
- Hierarchical Graph Representation Learning with Differentiable Pooling

Graph Pooling Using Clustering

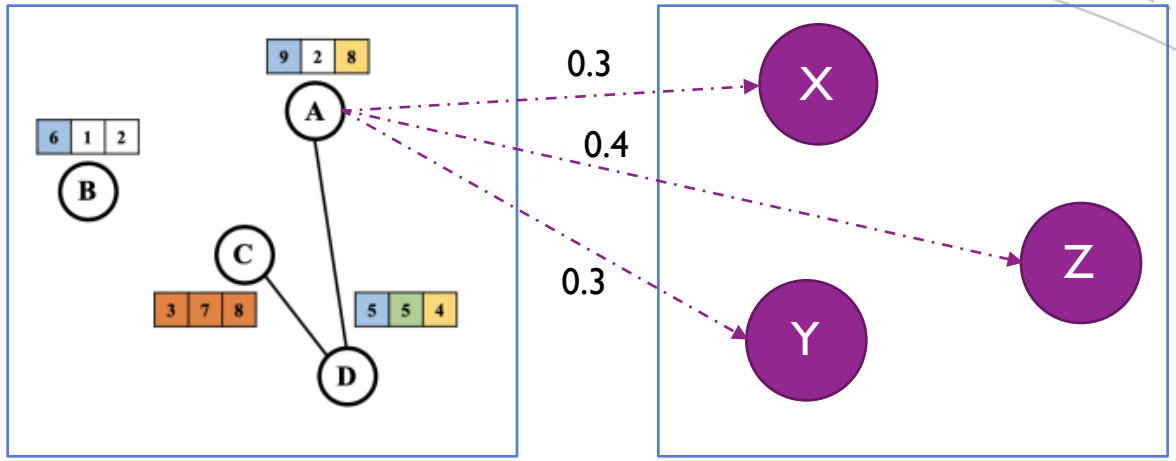


old graph (A, H)

new graph (A', H')

Assign each node to clusters with specific probabilities

Assignment matrix



S:

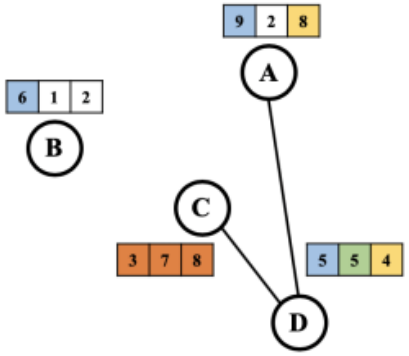
A	0.3	0.4	0.3
B	0.2	0.7	0.1
C	0.1	0.2	0.7
D	0.2	0.3	0.5
	X	Y	Z

- Suppose we have an assignment matrix $S \in R^{4\times3}$.

Soft assignment or hard assignment?

Compute New Graph (Feature Matrix)

- $H' = S^T H$



	A	B	C	D
X	0.3	0.2	0.1	0.2
Y	0.4	0.7	0.2	0.3
Z	0.3	0.1	0.7	0.5

S^T

A	9	2	8
B	6	1	2
C	3	7	8
D	5	5	4

H

- Takes the node embeddings
- Aggregates them according to assignment matrix
- Generates embeddings for new clusters

Compute New Graph (Adjacency Matrix)

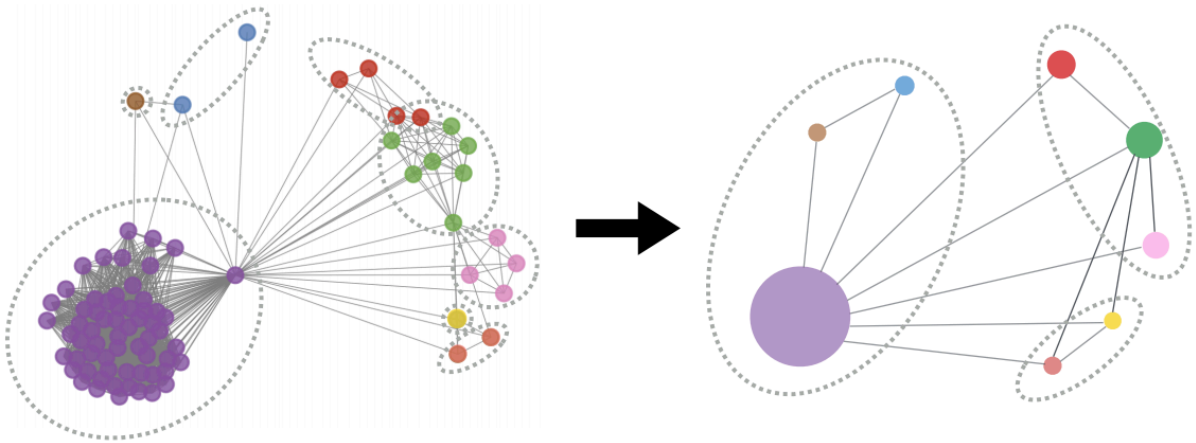
- $A' = S^T A S$

A':

	X	Y	Z
X	0.16	0.24	0.4
Y	0.24	0.36	0.6
Z	0.4	0.6	1

- Takes the adjacency matrix
- Generates a coarsened adjacency matrix
- Denotes the connectivity strength between each pair of clusters

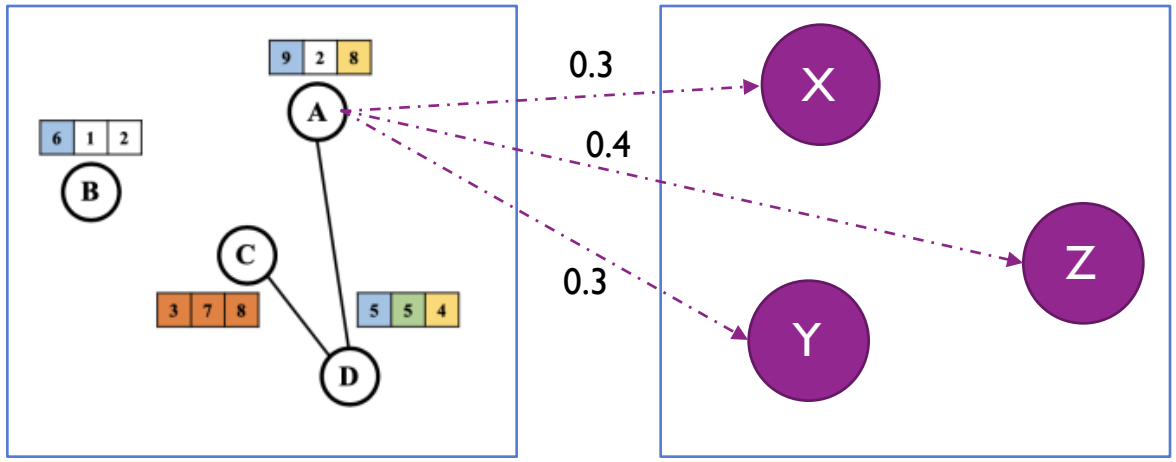
Graph Pooling

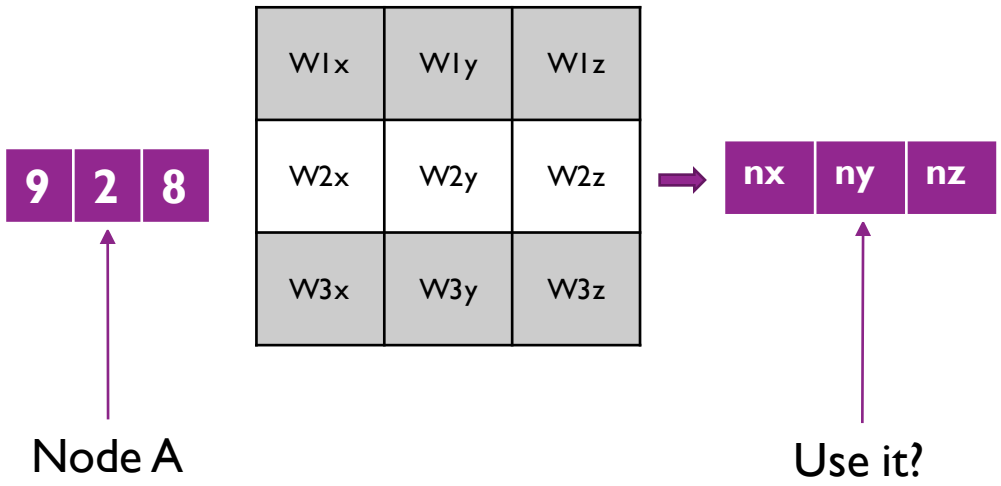


- Assume we have an assignment matrix S
- $A' = S^T A \ S$
- $H' = S^T H$

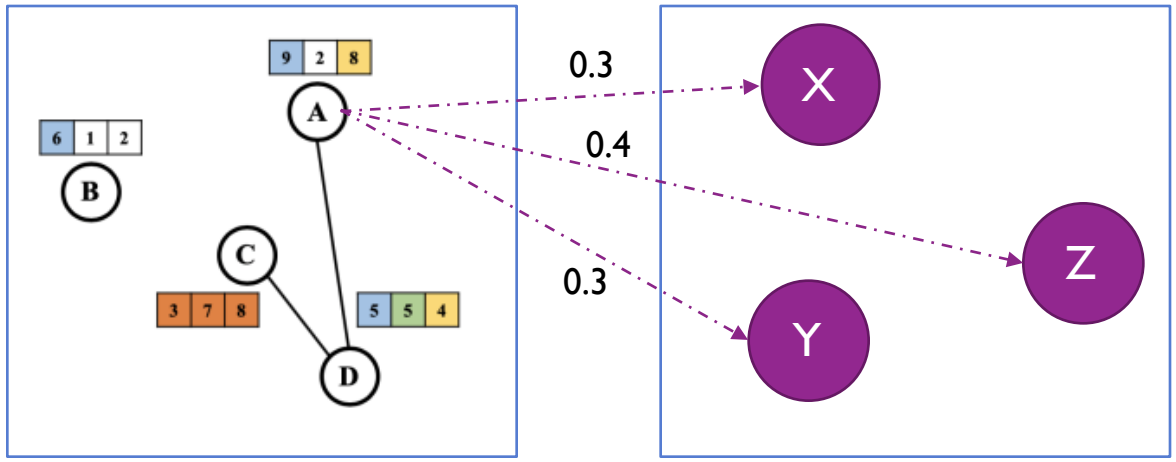
Problem?

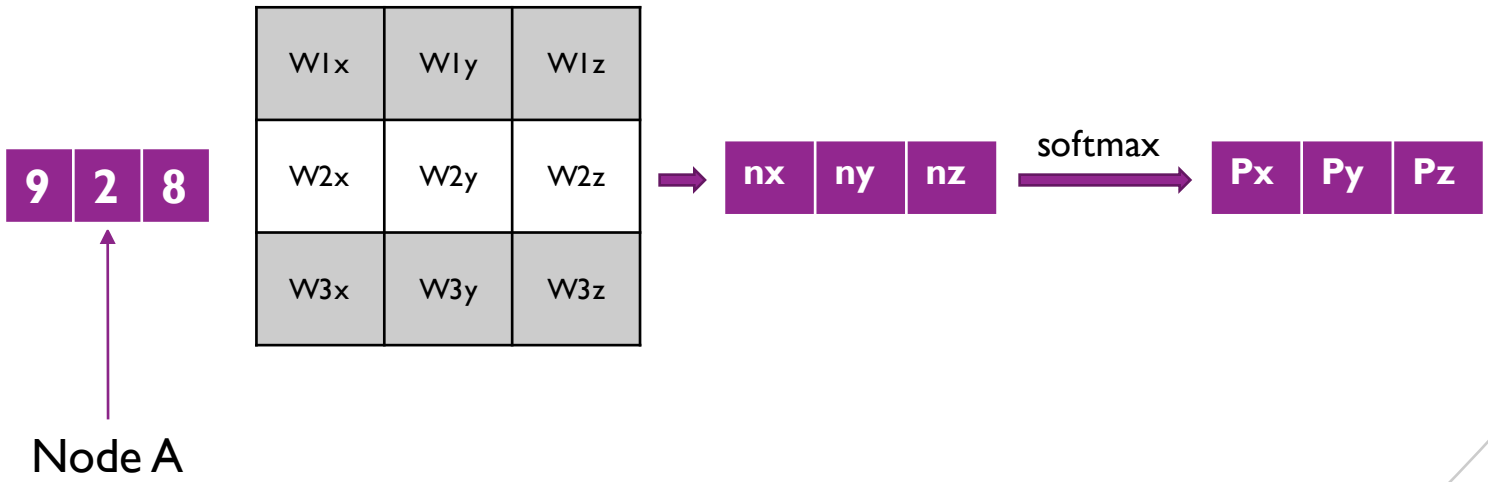
How to learn the assignment matrix

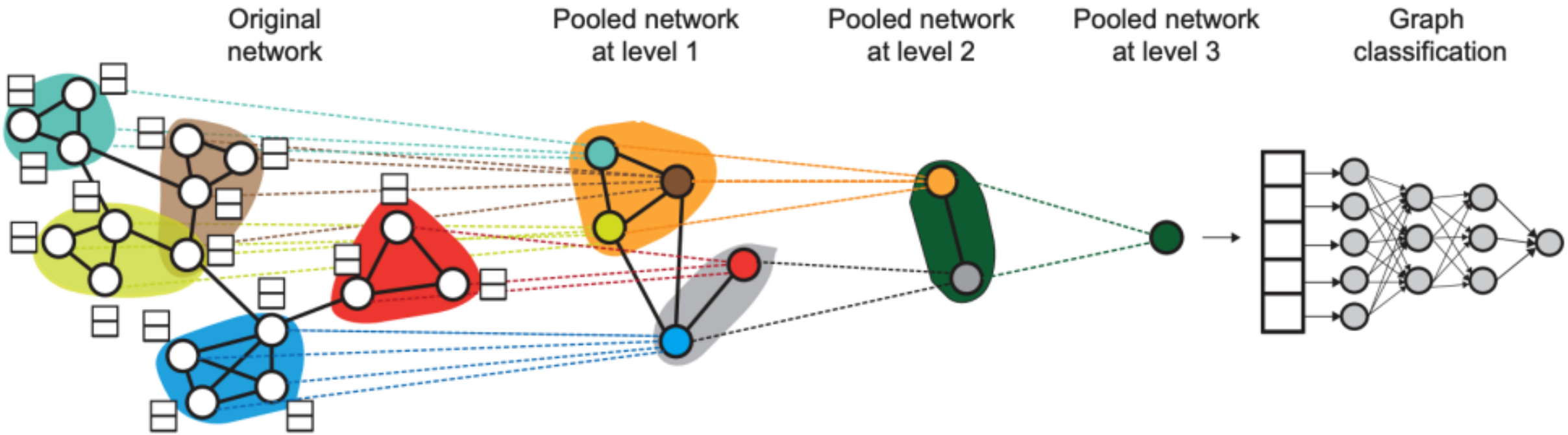




How to learn the assignment matrix







High-level illustration of DIFFPOOL.

Ying, Z., You, J., Morris, C., Ren, X., Hamilton, W., & Leskovec, J. (2018). Hierarchical graph representation learning with differentiable pooling. In *Advances in Neural Information Processing Systems* (pp. 4800-4810).

Graph Pooling Networks

- Some questions:
  - What problem does it solve?
    - Node classification? Link prediction? Graph classification?
  - How is this related to pooling layer?
  - **Any limitations?**
    - **(number of parameters, graph connectivity)**


Спасибо
Gracias
شكرا
Obrigado
Спасибо
Dank U
Grazie
Ευχαριστώ
Danke
Ευχαριστώ
Merci
Thank You
Ngiyabonga
Dank U
Diolch
Thank You
Obrigado
Ngiyabonga
Tack
Merci
Ευχαριστώ
Dzięki
Danke
Grazie
תודה
Dank U
Terima Kasih
Diolch
Grazie
Dzięki
Merci
Tack
Tack
Diolch
Gracias
תודה