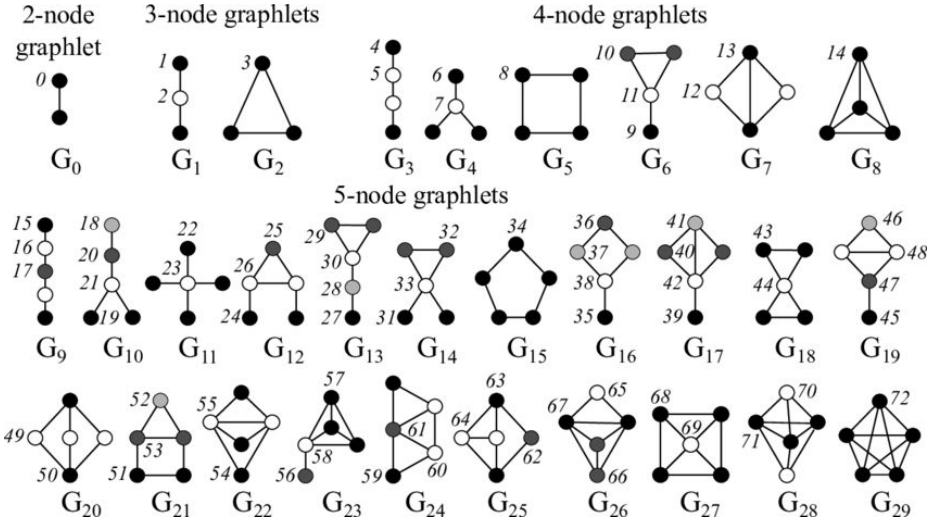
Multimodal Graph Autoencoder

21.04.

Administriva

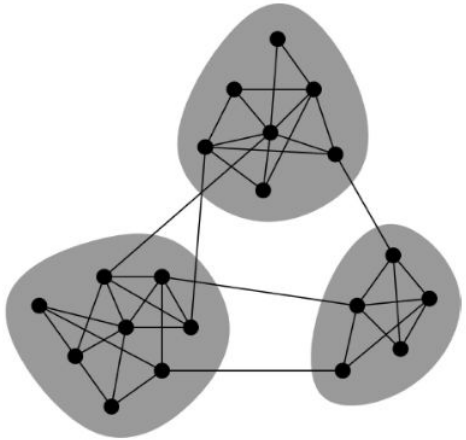
- Slack
  - Upload this morning
- Cluster
  - Received all, do I have to test now?
- Overleaf
  - New updates?

Graphlets



- Graph building blocks
  - Possible granularity reduction?
- Weisfeiler-Lehman kernel
  - Smartly count the amount of the different graphlets
  - O(#edges)
- Often used as similarity measure
    - Also works for changing amount of nodes :)
    - Prefer Graph-Edit-Distance if available and feasible
      - More precise and closer to what happens during generation

Communities/Cliques/Clusters



- Possible level for hierarchy
- Louvain algorithm
  - Each node is it's own cluster
  - Change cluster if it is beneficial
    - Pre-defined function → Less total edges
  - Contract edges inside cluster, sum outgoing edges
  -
  - Probably could be learned with GNNs (if not already done)
- Simple Community-affiliation graph model
  - If two nodes are in same community, make an edge with probability p_c

GraphRNN

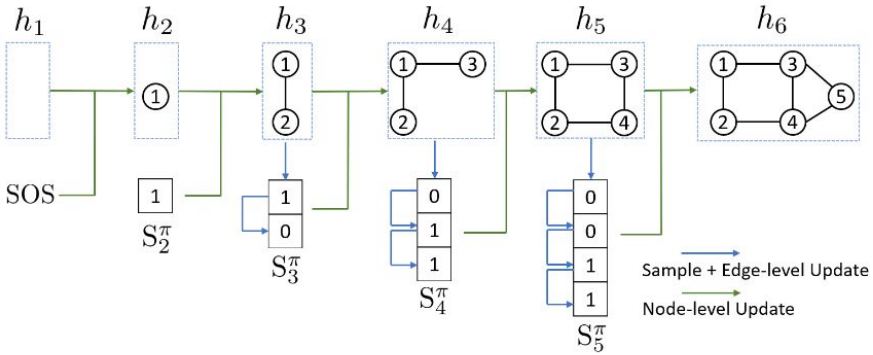
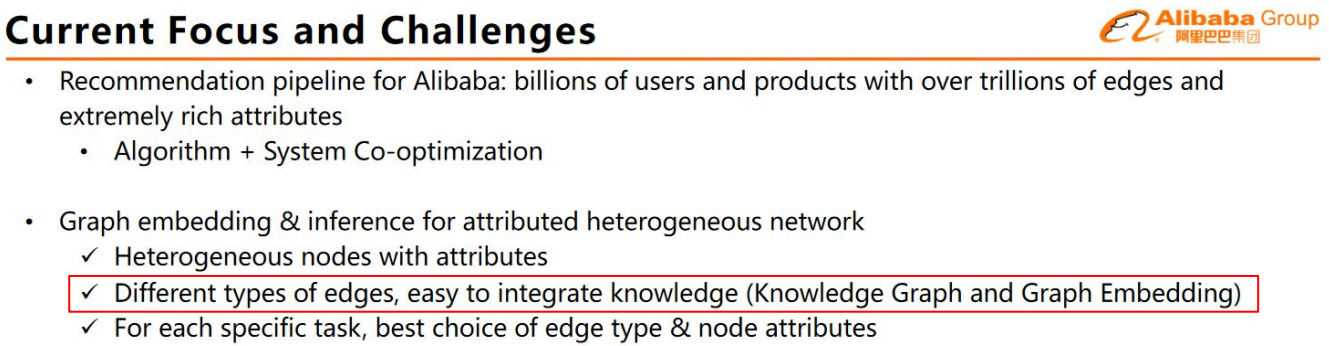


Figure 1. GraphRNN at inference time. Green arrows denote the

- Graph generation as a sequence of
- node additions and edge additions
-
- One can use RNNs on sequences
- Two RNNs
  - Node RNN, decides if new node or EOS, gives hidden state for Edge RNN
  - Edge RNN, decide if nodes should be connected, gives hidden state for Node RNN
  - GOTO node RNN
- Tractability and loss tricks needed
- Could be used on each cluster, needs to be tweaked to incorporate intercluster edges
- GCPN uses Reinforcement-Learning and GNNs for targeted generation of proteins

Other from guest lecturers

- SubGNN
  - Subgraph message passing
  - Anchor nodes for position in graph
    - Hardly needed for us
- Aligraph
  - Algorithm warehouse may have some interesting algorithms to check out
  - Graph embeddings

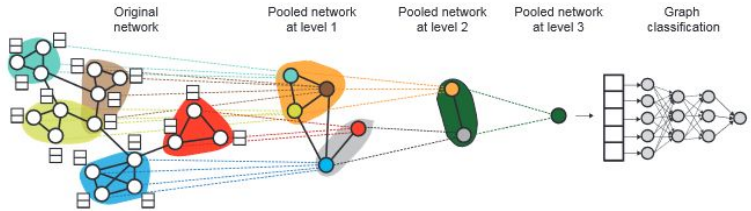


Hierarchical graph embedding in vector space by graph pyramid

- Like in pictures, use “lower resolution” with supernodes/clusters
  - Generate super-nodes by using your favourite clustering algorithm on a task dependent affinity matrix
    - Use k-means to compute the features of the super-node
  - Insert super-edges
    - Check if the amount of between cluster edges is above a threshold
  - Labelling
- Use the concatenated graph representation of many resolutions
  - Experiments done with three resolution layers, each clustering layer are 3 rounds of GNN feature updates
  - Structure and information update on two different layers
- Many algorithms can be augmented with this technique
- Rather formulaic, not much learning done

Hierarchical Graph Representation Learning with Differentiable Pooling

- Predetermined amount of clusters
- Compute soft-assignments to clusters in matrix S
  - S=softmax(GNN_pool(A,X))
  - S “sums up to 1” per row
- Update the features and adjacency matrix accordingly
  - Z=GNN_embed(A,X)
  - X=SZ
  - A=S^T A S
- Use extra losses at each layer to regulate
  - Link prediction (close nodes go together)
  - Entropy loss (node goes to only one cluster)
- Learning and elegant



Multimodal Graphs

- Generally graphs are binary or weighted for one relation
- Each relation can be one modality
- The task is given some relation(s), generate the missing one
- Many (huge) datasets available, rather realistic scenario, already now
- Use the same nodes
  - Same amount
  - Same feature space

Questions/ToDo

- Current stack viable (could be optimised)
  - Multirelational dataset with same nodes
  - Encoding → Clustering
    - Hierarchical Graph Representation Learning with Differentiable Pooling
    - Or Louvier
  - Latent graph of clusters
    - What do if cluster different (your original question)
      - Metalabels? (label1 x label2) per node :(
  - Decoding → Generation inside* clusters
    - Simple Community-affiliation graph model
    - GraphRNN
  - Loss: Graph Edit distance (if possible) or Weisfeiler Lehman
- Now that we have some systems, it is easier to find similar ones for our task
- Find a programming framework that works well
  - Course already gave some pointers
- Find a concrete dataset with many (small) graphs with different relations