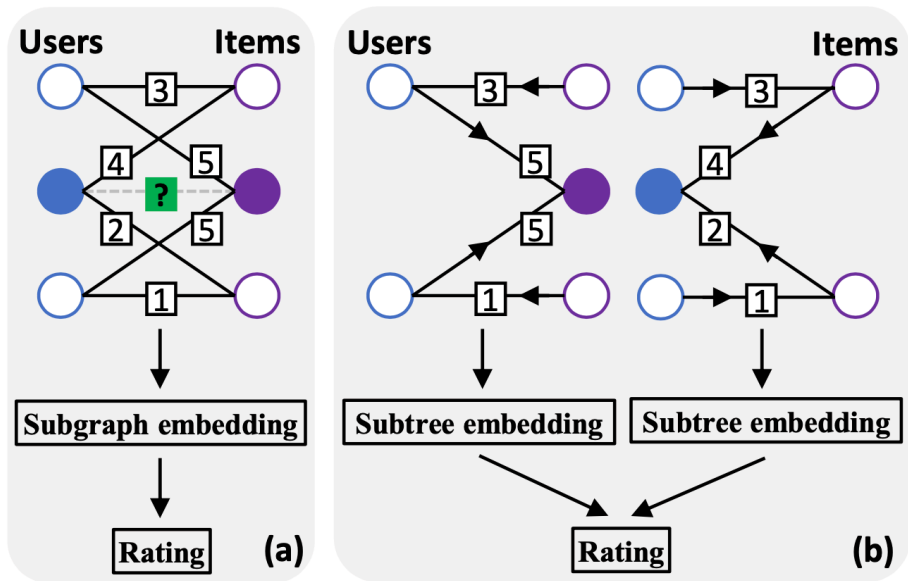


Subgraph multi-layer GNN

이 논문이 동일한 sub-graph에 대해서 node level GNN이 아닌 graph level의 GNN을 사용한 이유부터 설명해보겠습니다.



먼저 저자는 target node pair가 서로 밀접하게 연관성이 있는지 또는 서로 아무런 영향을 미치지 않는지 구분할 수 있어야 한다고 하면서, 위 그림과 같이 (a)와 (b)를 다른 그래프로 분류할 수 있는지로 설명을 합니다.

Node level GNN으로 예시를 든게 GC-MC인데 GC-MC는 node embedding을 하기 때문입니다. 이때 단순히 1-hop neighborhood만으로는 (a), (b)를 구분하지 못 한다고 합니다. 구분해내기 위해선 전체 그래프 정보를 다 사용해야한다고 하며, 그러기 위해선 여러 layers를 쌓아야하는데, 이는 node level에선 관련 없는 node까지 확장되고 또한 over-smoothing 문제를 일으키게 됩니다. [node level GNN에서는 sub-graph 추출을 하지 않고 whole graph 에 대해서 node embedding을 하는 것 같습니다.]

The reason is to leverage multiple hops' different graph structure features around each center node to achieve a maximum discriminative power that the Weisfeiler-Lehman algorithm has [1].

각 center node에 대하여 멀티 hop을 거치며 다양한 그래프 구조 특성을 사용하는 이유는 WL algorithm이 가지는 것과 같은 power의 분별력을 갖기 위해서 입니다. 다만 본 논문에서는 학습 시간을 효율적으로 가져가기 위해서 1hop만 취하고 있고 그래도 유의미한 성능을 보이고 있습니다.

Xu, K., Hu, W., Leskovec, J., & Jegelka, S. (2018). How powerful are graph neural networks? In ICLR-2019.

https://muhanzhang.github.io/papers/AAAI_2018_DGCNN.pdf

SEAL vs.

본 논문과 같은 저자이며, SEAL에서의 link prediction이 기본적으로 IGMCM과 같은 흐름을 가지고 있는걸로 보입니다.

조금 더 일반적인 경우에 link prediction 방법을 SEAL에서 제안하고 있으며, IGMCM은 bipartite graph에 좀 더 적합한 방법으로 SEAL을 알맞게 맞춘 framework라고 생각할 수 있겠습니다.

Subgraph sampling 방법.

제가 이해한 바로는 고정된 rating matrix라면 어떤 target user-item pair는 임의의 h - hop에서 동일한 sub-graph를 가지게 될 것입니다.

Algorithm 1 ENCLOSING SUBGRAPH EXTRACTION

```
1: input:  $h$ , target user-item pair  $(u, v)$ , the bipartite graph  $G$ 
2: output: the  $h$ -hop enclosing subgraph  $G_{u,v}^h$  for  $(u, v)$ 
3:  $U = U_{fringe} = \{u\}, V = V_{fringe} = \{v\}$ 
4: for  $i = 1, 2, \dots, h$  do
5:    $U'_{fringe} = \{u_i : u_i \sim V_{fringe}\} \setminus U$ 
6:    $V'_{fringe} = \{v_i : v_i \sim U_{fringe}\} \setminus V$ 
7:    $U_{fringe} = U'_{fringe}, V_{fringe} = V'_{fringe}$ 
8:    $U = U \cup U_{fringe}, V = V \cup V_{fringe}$ 
9:   Let  $G_{u,v}^h$  be the vertex-induced subgraph from  $G$  using vertices  $U, V$ 
10:  Remove edge  $(u, v)$  from  $G_{u,v}^h$ .
11: end for
12: return  $G_{u,v}^h$ 
```

Note: $\{u_i : u_i \sim V_{fringe}\}$ is the set of nodes that are adjacent to at least one node in V_{fringe} with any edge type.

아래 사진은 target link(또는 target user와 target item 쌍)를 기준으로 h -hop의 enclosing subgraph를 추출하는 algorithm입니다. 아래 방법으로 모든 labeled link(rating이 있는 user-item 쌍)에 대하여 subgraph를 추출할 수 있을 것입니다.

따라서 sub-sampling의 조건은 target user-item pair로 부터 몇 hop 떨어져 있는 노드들 까지 추출할지 일 것 같습니다.

그리고 이 위 조건으로는 하나의 rating matrix에서 구해진 bipartite graph에서 특정 edge(user-item node pair)는 항상 같은 sub-graph를 가질 것으로 생각됩니다.