

Graph neural networks : A review of methods and applications

Jie Zhou, Ganqu Cui, Shengding Hu , Zhengyan Zhang , Cheng Yang , Zhiyuan Liu ,
Lifeng Wang , Changcheng Li , Maosong Sun

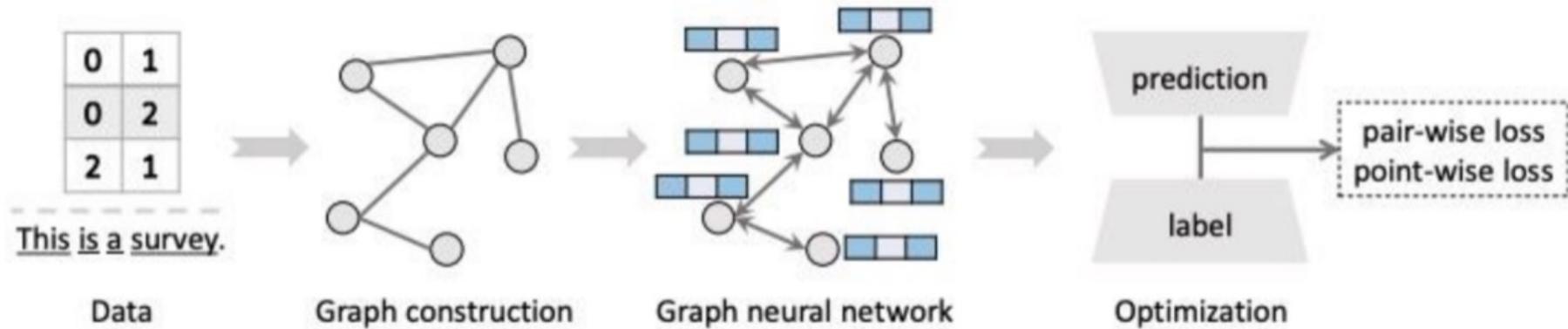
발표자 : 김수지

목차

1. Intro
2. General design pipeline of GNNs
3. Applications
4. Variants considering graph type and scale
5. Instantiations of computational modules
6. Variants for different training settings
7. Example of Design

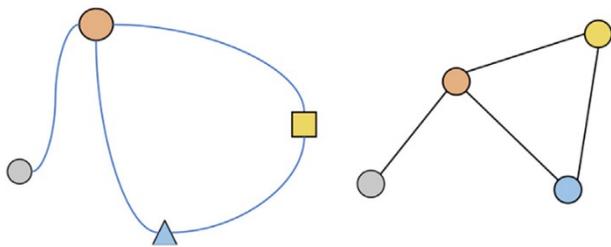
1. Intro

GNN 이란?

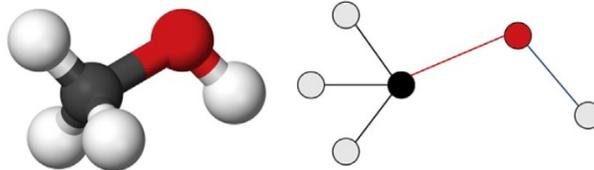


1. Intro

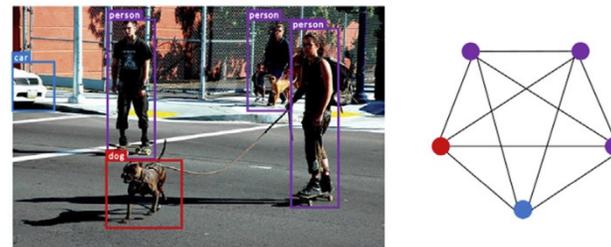
GNN을 사용한 task



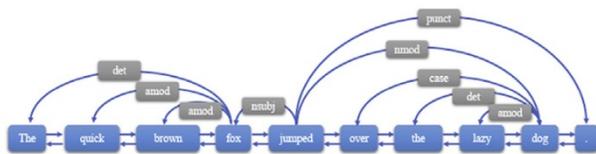
(a) Physics



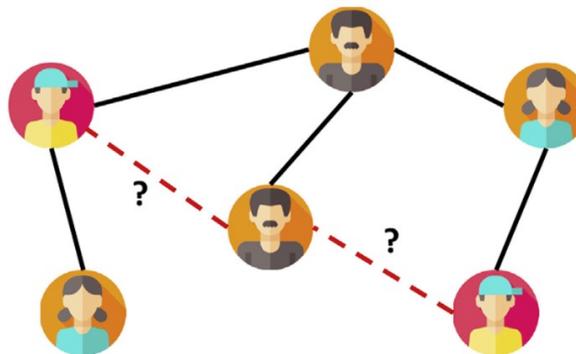
(b) Molecule



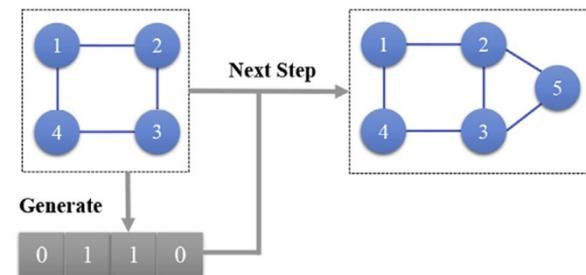
(c) Image



(d) Text



(e) Social Network



(f) Generation

2. General design pipeline of GNNs

(1) 그래프 구조 찾기

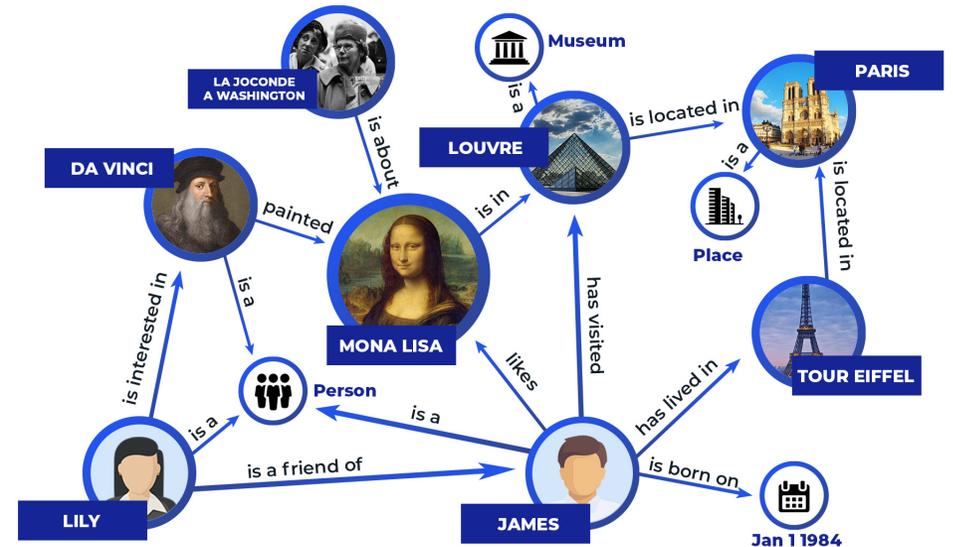
응용 분야에서 그래프 구조를 파악

- 구조적 시나리오 : 그래프 구조가 명시적으로 주어짐

Ex) 분자, 물리 시스템, 지식 그래프

- 비구조적 시나리오 : 그래프가 묵시적으로 주어짐

Ex) 이미지, 텍스트



Ex) 지식 그래프

2. General design pipeline of GNNs

(2) 그래프 유형과 규모 지정

그래프 유형과 규모를 파악

- 그래프 유형
 - Directed/Undirected Graphs
 - Homogeneous/Heterogeneous Graphs
 - Static/Dynamic Graphs

-> 그래프들이 제공하는 추가 정보를 고려

- 그래프 규모

그래프 크기 구분하는 명확한 기준 X

본 논문에서는,
그래프의 인접 행렬, 그래프 라플라시안을
장치에 저장하고 처리할 수 없는 경우

-> 대규모 그래프로 간주, 샘플링 방법 고려

2. General design pipeline of GNNs

(3) 손실 함수 설계

작업 유형과 학습 설정에 기반하여 손실 함수를 설계

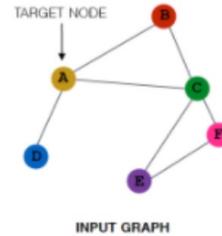
- Node level : node classification, node regression, node clustering
- Edge level : edge classification, link prediction
- Graph level : graph classification, graph regression, graph matching

- Supervised setting : train을 위해 label이 지정된 데이터 제공
- Semi-supervised setting : 작은 양의 label이 지정된 node와 많은 양의 label이 지정되지 않은 node 제공
- Unsupervised setting : label이 지정되지 않은 데이터만 제공

2. General design pipeline of GNNs

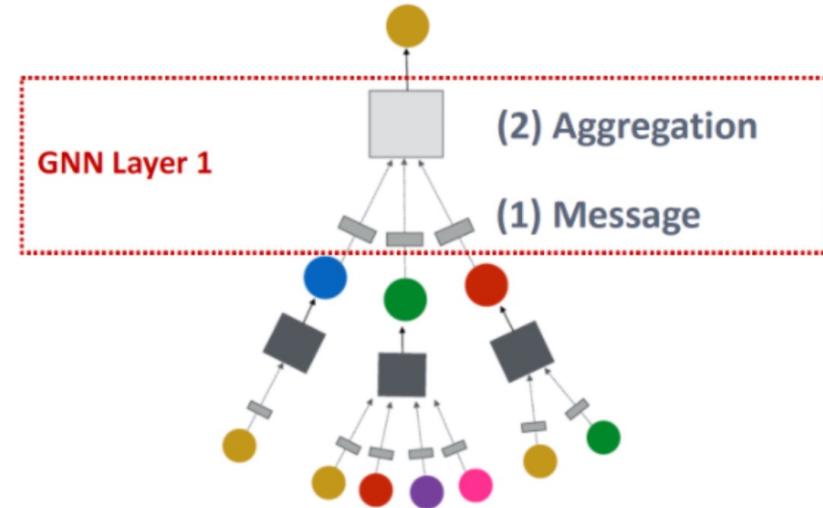
(4) 계산 모듈을 사용하여 모델 구축

- Propagation Module
- Sampling Module
- Pooling Module



GNN Layer = Message + Aggregation

- Different instantiations under this perspective
- GCN, GraphSAGE, GAT, ...



Ex) Propagation module

2. General design pipeline of GNNs

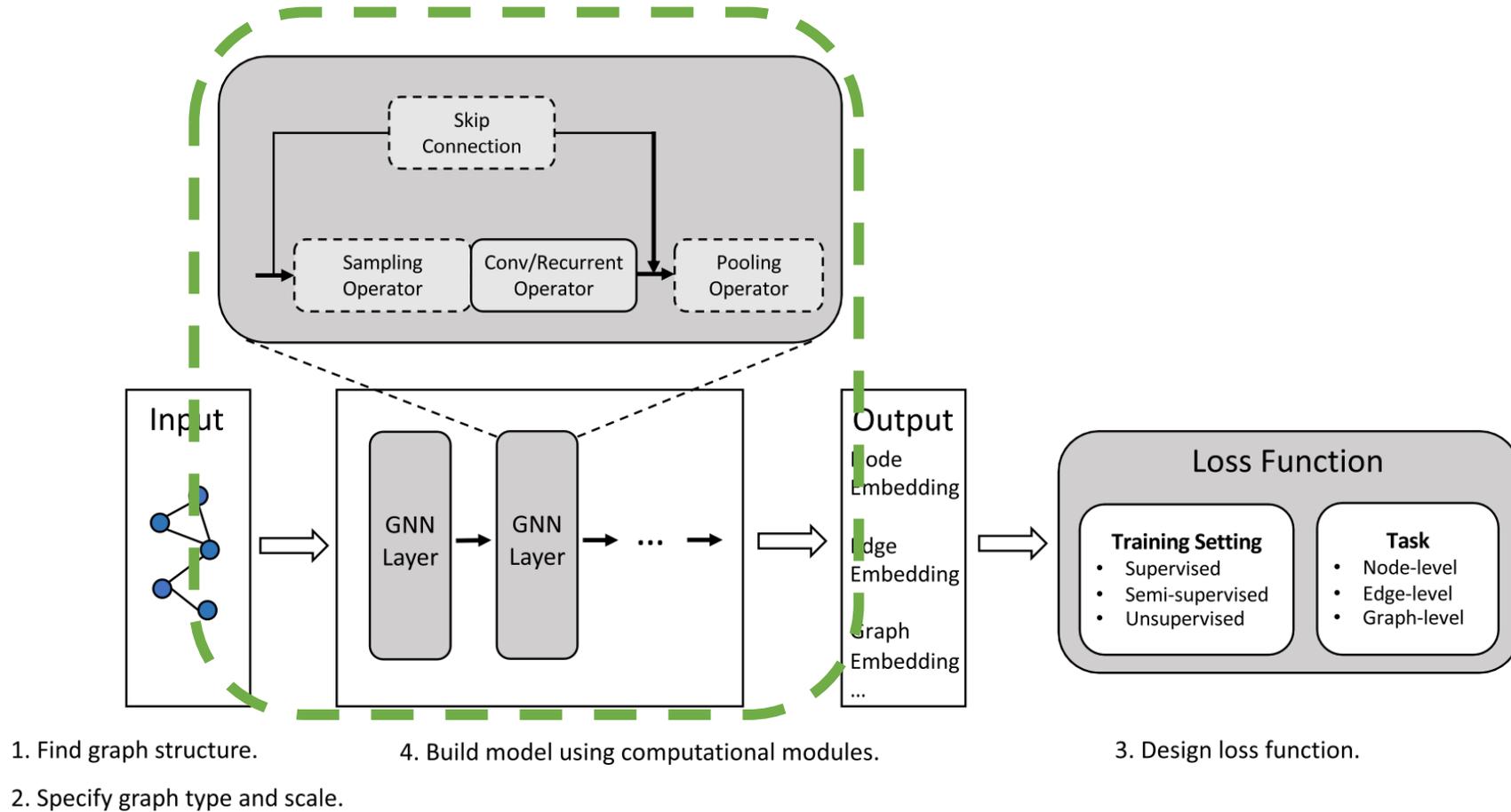


Fig. 2. The general design pipeline for a GNN model.

3. Applications

- 구조적 시나리오

: 데이터에 명시적인 관계 구조가 있는 경우

: 그래프 마이닝, 물리 시스템 및 화학 시스템 모델링, 지식 그래프, 교통 네트워크 및 추천 시스템 등

EX)

지식 그래프 ex) 저차원 임베딩 학습, 링크 예측, 다중 홉 추론

- R-GCN : 지식 그래프 임베딩과 GNN을 통합하는 첫번째 메소드
- Structure-Aware Convolutional Network

추천시스템

- GC-MC : 사용자-아이템 평점 그래프에 GCN을 처음 적용
- PinSage : 이분 그래프에 가중 샘플링 전략 사용
- GraphRec : 아이템 측면과 사용자 측면에서 사용자 임베딩 학습

3. Applications

- 비구조적 시나리오

: 관계 구조가 암묵적이거나 없는 경우

: 이미지 (컴퓨터 비전)와 텍스트 (자연어 처리)가 포함

EX)

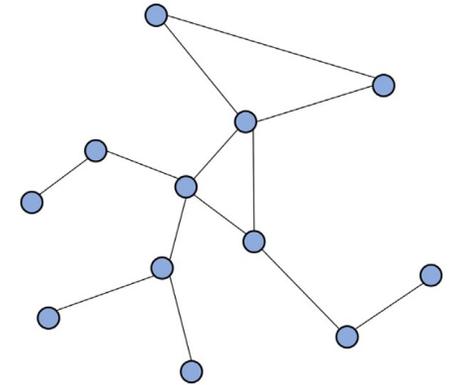
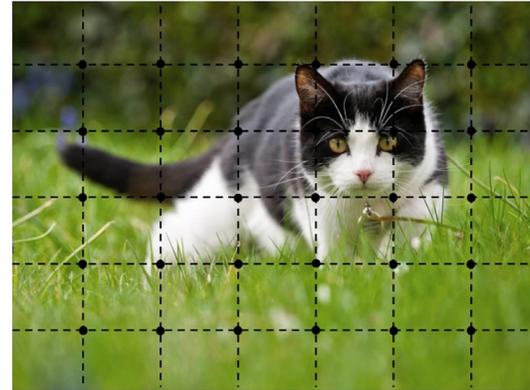


Fig. 1. Left: image in Euclidean space. Right: graph in non-Euclidean space.

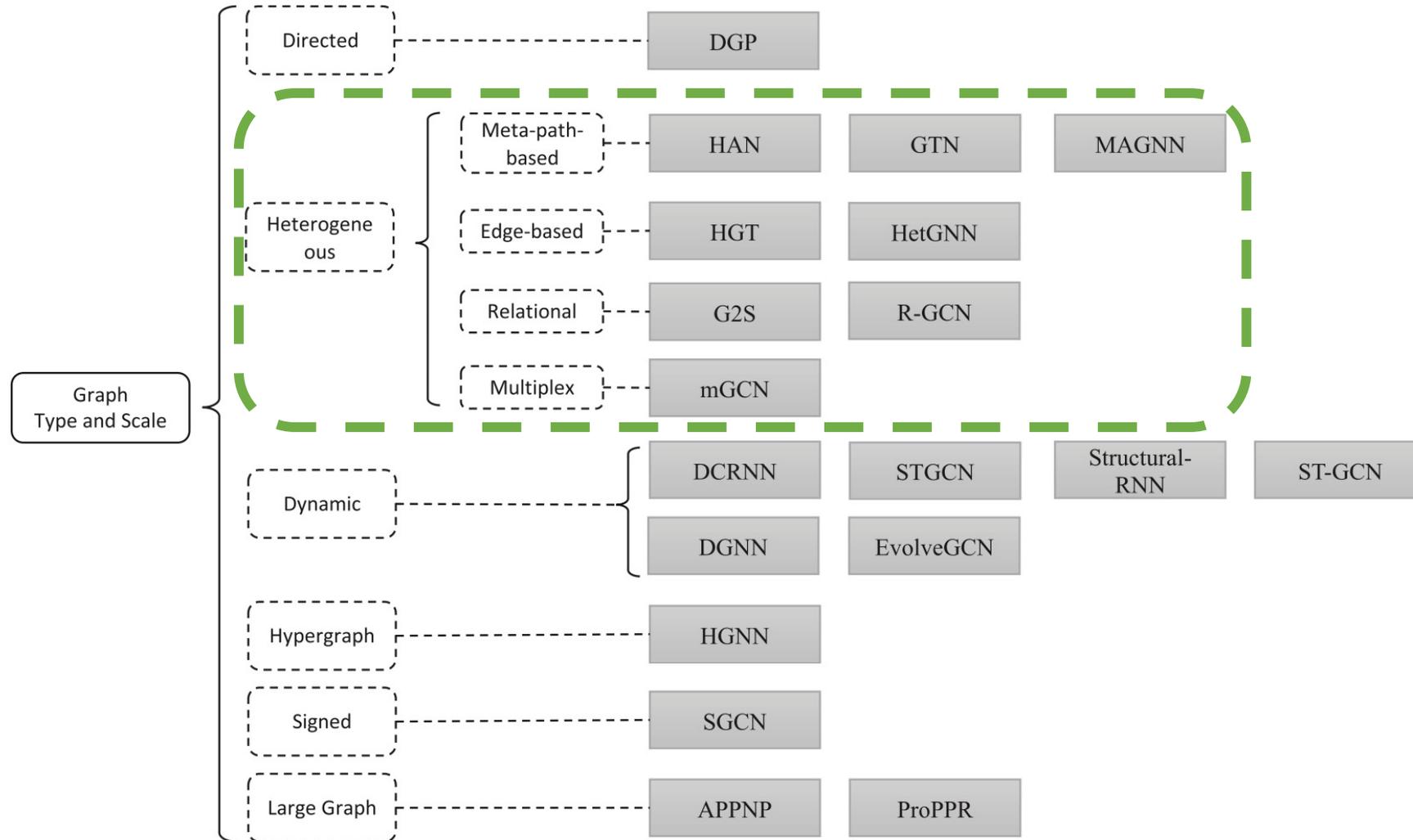
이미지

(1) ; N-shot 이미지 분류를 쉽게 하기 위해 knowledge graph 사용

텍스트

(2) ; 텍스트를 단어 그래프로 나타내 비연속적이고 먼 거리의 단어들 간 의미 포착

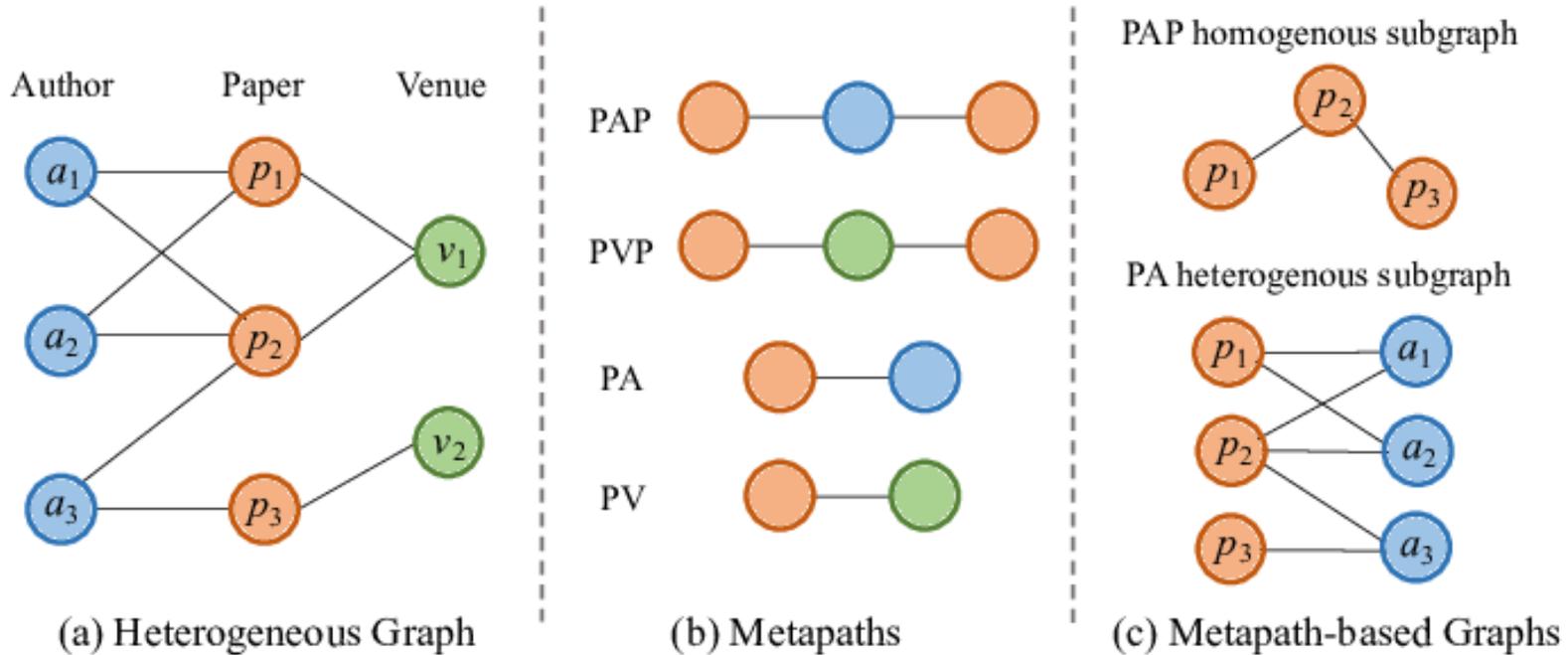
4. Variants considering graph type and scale



4. Variants considering graph type and scale

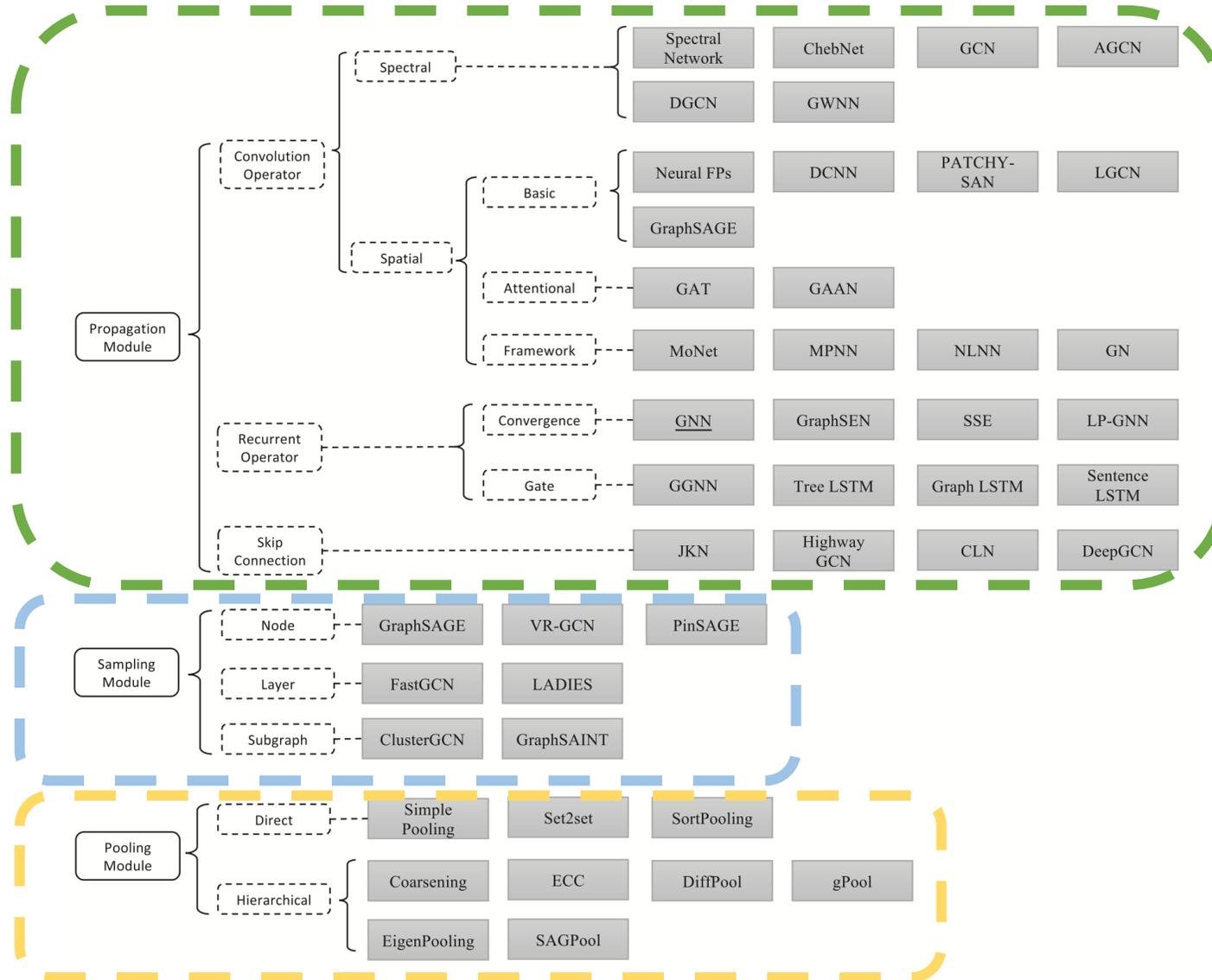
Heterogeneous graph

- Meta-path-based : 메타 패스의 두 끝 노드를 연결함으로써, 직접적으로 연결되지 않은 두 노드의 유사성을 포착



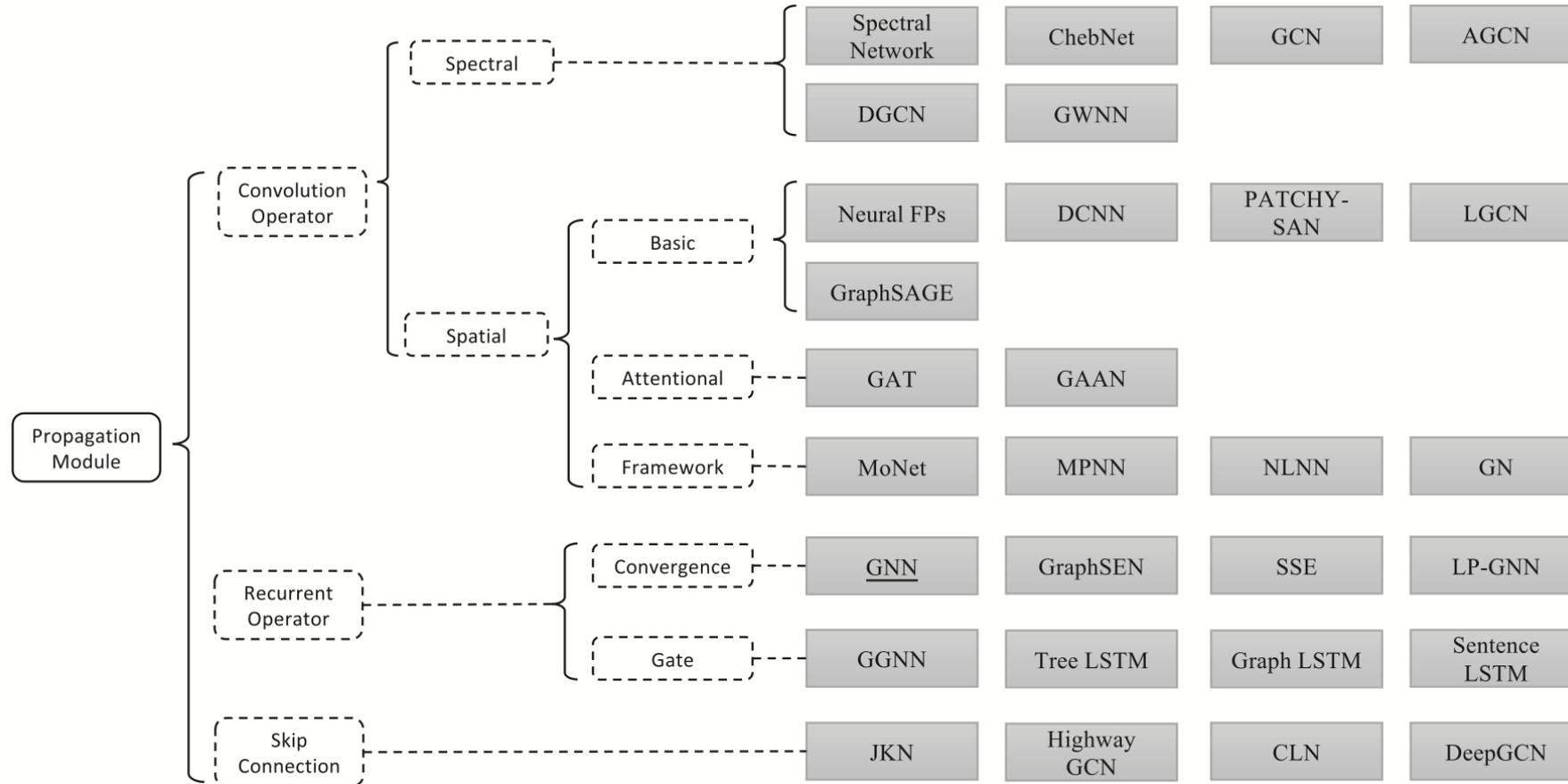
- 이외에도 Edge-based, Relational, Multiplex 등이 존재

5. Instantiations of computational modules



5. Instantiations of computational modules

Propagation module : 각 노드가 인접한 노드들(이웃)로부터 정보를 수집하고 전파



5. Instantiations of computational modules

- Spectral

: 그래프 신경망에서 라플라시안(Laplacian) 행렬의 스펙트럴 분해를 기반

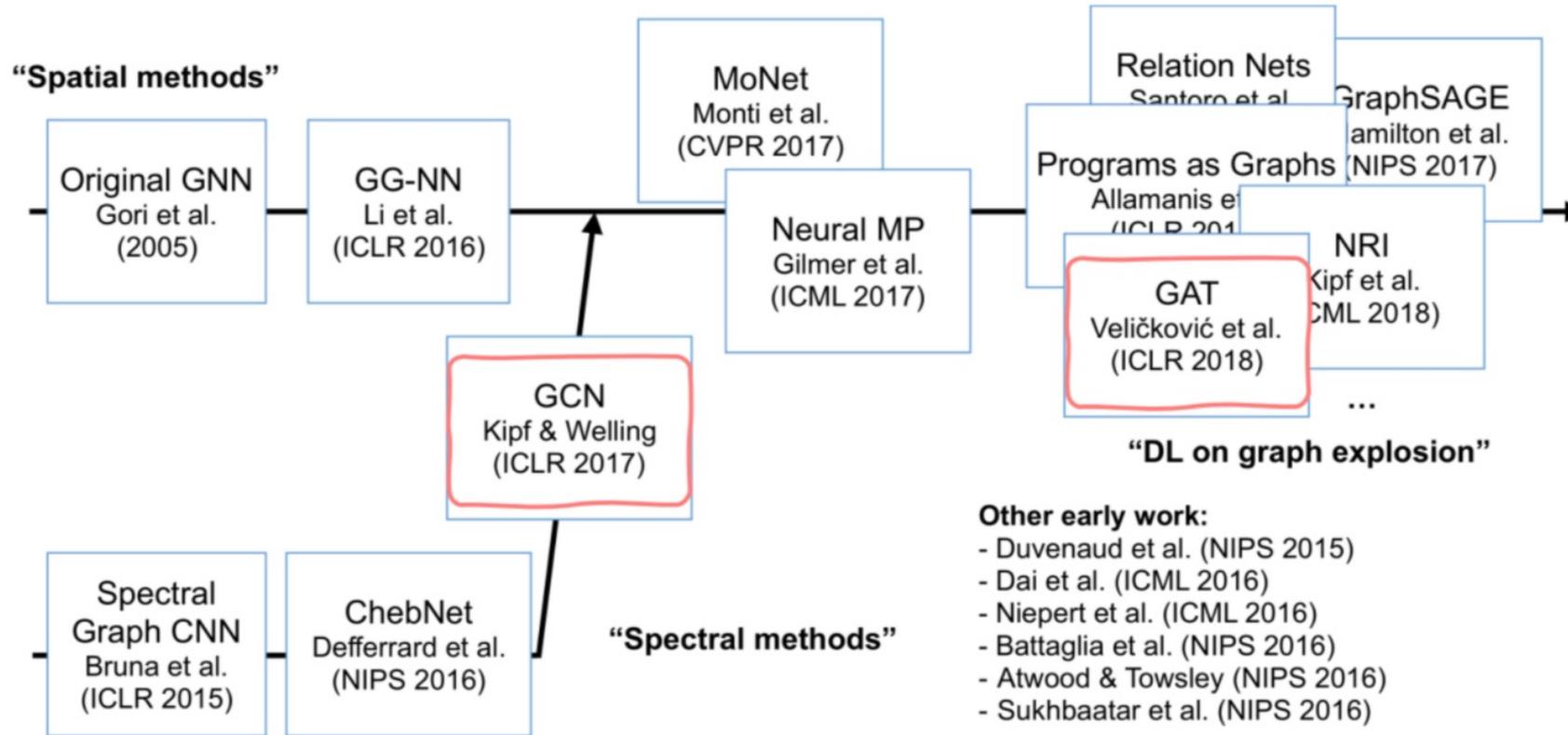
: 라플라시안 행렬의 고유값과 고유벡터를 계산하여 그래프 신경망의 컨볼루션 연산 수행

- Spatial

: 그래프 신경망에서 노드 자체와 이웃 노드들의 정보를 활용하여 컨볼루션 연산을 수행

: 각 노드는 그 자체와 주변 이웃의 정보를 종합하여 새로운 표현을 계산하며,
이 정보를 통해 그래프 상의 특징을 추출

5. Instantiations of computational modules



University Amsterdam, Thomas Kipf

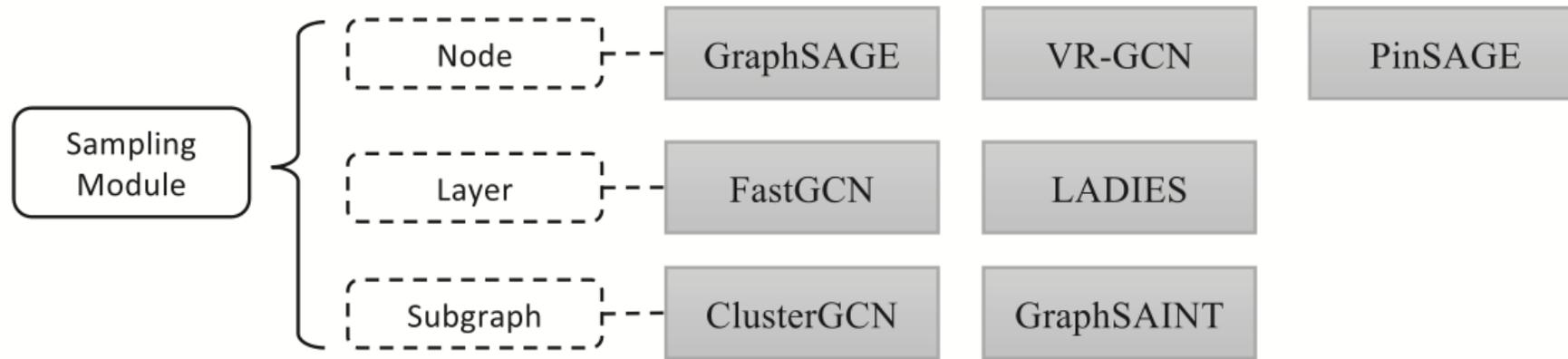
5. Instantiations of computational modules

Sampling module : 그래프(노드, 레이어, 서브그래프)를 샘플링하여 이웃 폭발 문제 완화

Node sampling : 각 노드의 이웃에서 하위 집합을 선택

Layer sampling : 각 계층에서 aggregation을 위해 작은 노드 집합 유지

Subgraph sampling : 다른 방식으로 여러 서브 그래프를 샘플링하고 서브 그래프 내에서 이웃 검색을 제한

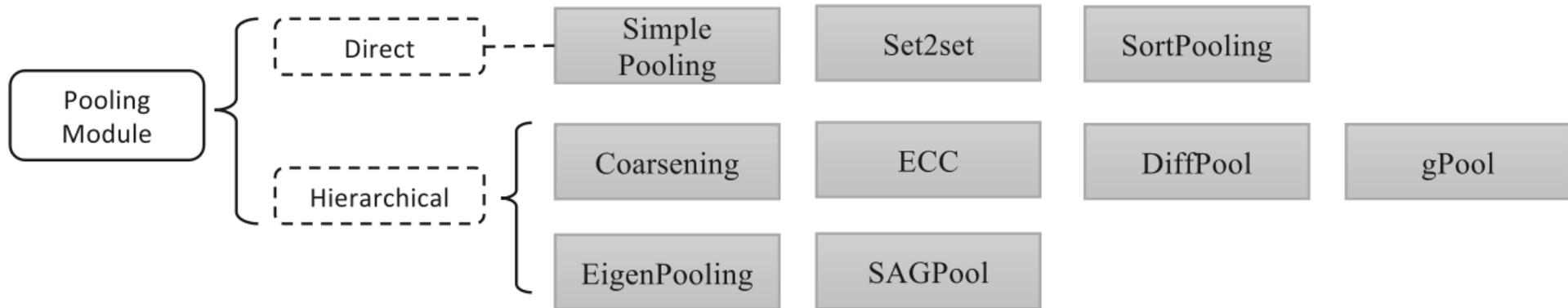


5. Instantiations of computational modules

Pooling module : 그래프의 노드를 그룹화하거나 집계하여 그룹 단위로 정보를 요약하는 과정

Direct pooling : 그래프의 노드를 직접적으로 그룹화하거나 집계하여 전역적인 정보를 추출

Hierarchical pooling : 그래프의 노드를 여러 단계로 나누어 계층 구조를 형성하고, 각 계층에서 그룹화 및 집계를 수행

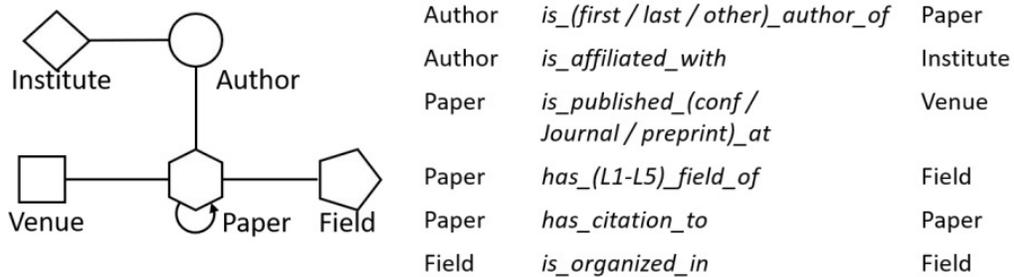


7. Example of Design : GPT-GNN

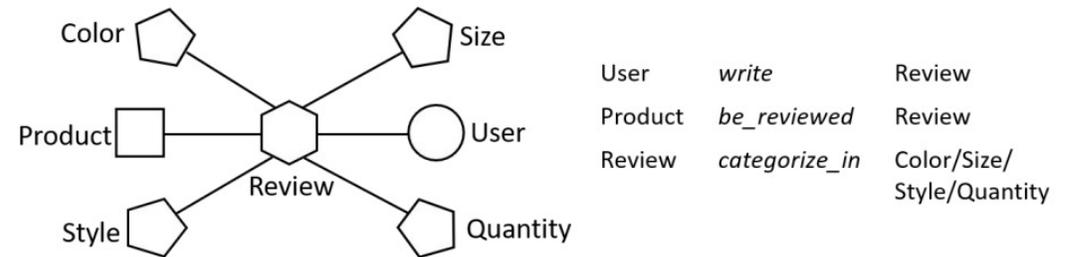
(1) 그래프 구조 찾기

학술적 지식 그래프 및 추천 시스템 응용에 중점

추천 시스템에서는 사용자, 항목 및 리뷰를 node로 간주하고 이들 사이의 상호작용을 edge로 간주



(a) Schema and meta relations of Open Academic Graph (OAG)



(b) Schema and meta relations of Amazon Review Recommendation Dataset (Amazon)

7. Example of Design : GPT-GNN

(2) 그래프 유형과 규모 지정

해당 논문은 Heterogeneous graph에 집중

Academic graph 및 추천 graph에는 수백만개의 노드 포함

-> 대규모 heterogeneous graph에 초점

7. Example of Design : GPT-GNN

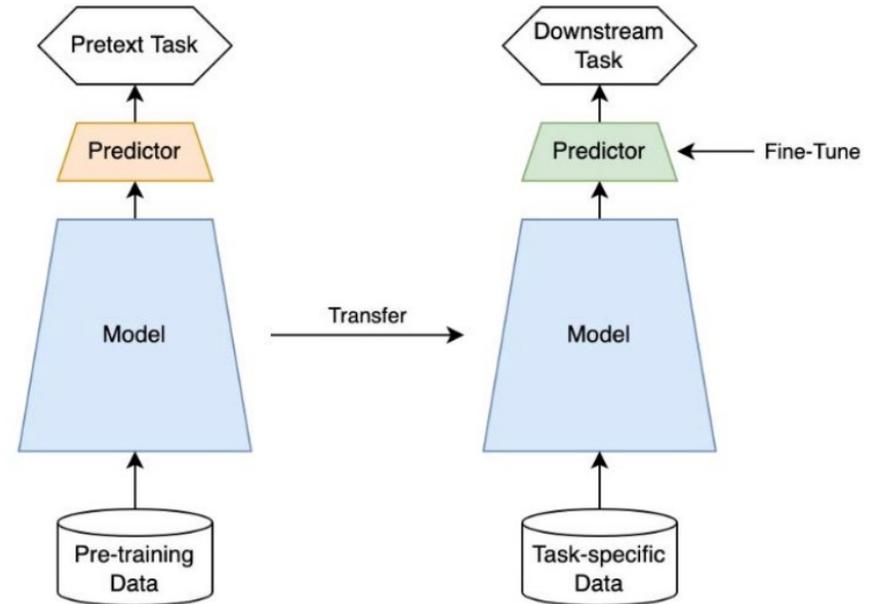
(3) 손실 함수 설계

해당 논문은 모두 노드 수준의 task (Paper-field 예측 등)

-> Pretrain 단계에서 노드 representation 학습

Label 지정된 데이터 x

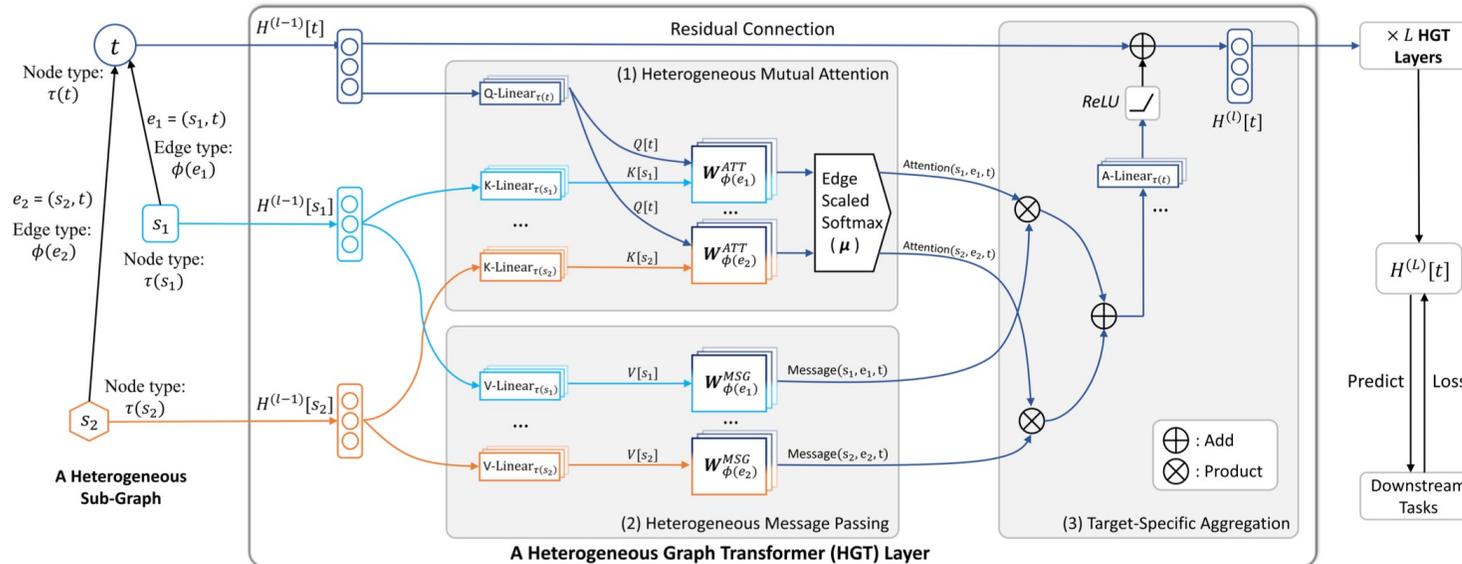
-> Self-supervised graph generation task 설계



7. Example of Design : GPT-GNN

(4) 계산 모듈을 사용하여 모델 구축

- Propagation Module : HGT
- Sampling Module : HGSampling
- Pooling Module : 사용 X



감사합니다