Heterogeneous Graph Representation Learning

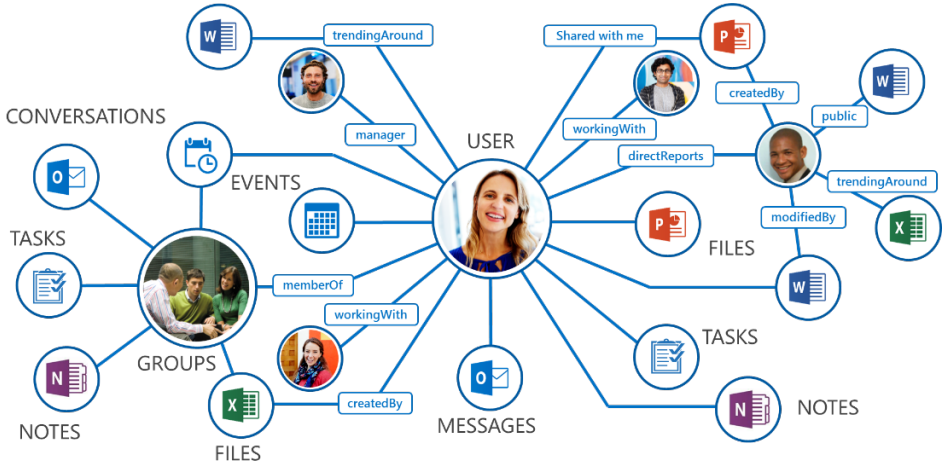
Yuxiao Dong, Ziniu Hu, Kuansan Wang, Yizhou Sun, Jie Tang

@IJCAI 2020

Can we get rid of the manual design of meta paths?

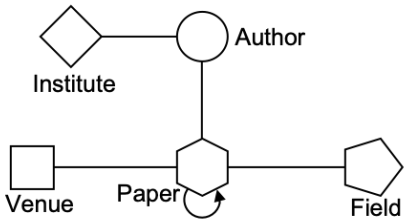
Heterogeneous Graphs

academic graph



office graph

Heterogeneous Graph **Mining**



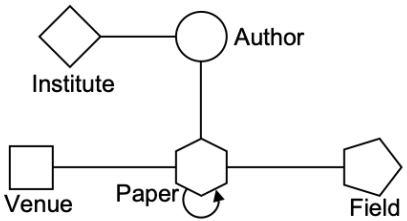
(a) The schema of heterogeneous academic networks

⇨ meta paths ⇨

- Classification
  - ✓ RankClass, …
- Clustering
  - ✓ RankClus, …
- Ranking
  - ✓ PathSim, …
- Link Prediction
  - ✓ PathPredict, …
- …

1. Yizhou Sun, Jiawei Han. Mining Heterogeneous Information Networks: Principles and Methodologies. M & C Publishers, 2012.

Heterogeneous Graph **Representation Learning**



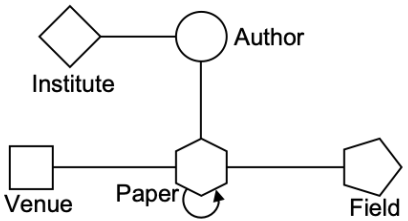
(a) The schema of heterogeneous academic networks

meta paths ✚ graph representation learning

heter. graph representation learning
- metapath2vec, PTE, …
- R-GCN, HetGNN, GEM, …

- Classification
- Clustering
- Ranking
- Link Prediction
- …

Heterogeneous Graph Representation Learning



(a) The schema of heterogeneous academic networks

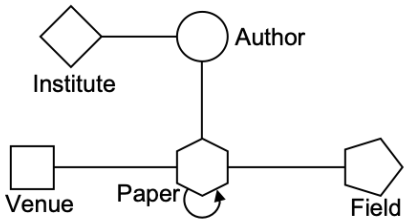
Can we get rid of meta paths?

meta paths (feature engineering) ✚ graph representation learning

heter. graph representation learning
- metapath2vec, PTE, …
- R-GCN, HetGNN, GEM, …

- Classification
- Clustering
- Ranking
- Link Prediction
- …

Heterogeneous Graph Representation Learning



(a) The schema of heterogeneous academic networks

Heterogeneous Graph Transformer (HGT)

- no manual design of meta paths!
- straightforward for (heterogeneous) graph pre-training, e.g., GPT-GNN

- Classification
- Clustering
- Ranking
- Link Prediction
- …

1. Hu, Dong, Wang, Sun. Heterogeneous graph transformer. In WWW 2020.
2. Hu, et al. GPT-GNN: Generative Pre-Training of Graph Neural Networks. In KDD 2020.

Heterogeneous Graph Representation Learning

https://ogb.stanford.edu/

(as of Dec. 13, 2020)

Leaderboard for ogbn-mag

Rank	Method	Test Accuracy	Validation Accuracy	Contact	References	#Params	Hardware
1	**HGT (LADIES Sample)**	0.5007 ± 0.0043	0.5124 ± 0.0039	Ziniu Hu	Paper, Code	21,173,389	Tesla K80 (12GB GPU)
2	GraphSAINT (R-GCN aggr)	0.4751 ± 0.0022	0.4837 ± 0.0026	Matthias Fey – OGB team	Paper, Code	154,366,772	GeForce RTX 2080 (11GB GPU)
3	**R-GCN+FLAG**	0.4737 ± 0.0048	0.4835 ± 0.0036	Kezhi Kong	Paper, Code	154,366,772	GeForce RTX 2080 Ti (11GB GPU)
4	NeighborSampling (R-GCN aggr)	0.4678 ± 0.0067	0.4761 ± 0.0068	Matthias Fey – OGB team	Paper, Code	154,366,772	GeForce RTX 2080 (11GB GPU)
5	SIGN	0.4046 ± 0.0012	0.4068 ± 0.0010	Lingfan Yu (DGL Team)	Paper, Code	3,724,645	Tesla T4 (15GB GPU)

1. Hu et al. Open Graph Benchmark: Datasets for Machine Learning on Graphs. In NeurIPS 2020.
2. Wang et al. Microsoft academic graph: When experts are not enough. Quantitative Science Studies, 2020

Heterogeneous Graph Representation Learning

Yuxiao Dong, Ziniu Hu, Kuansan Wang, Yizhou Sun, Jie Tang

ericdongyx@gmail.com, {bull, yzsun}@cs.ucla.edu,
kuansanw@Microsoft.com, jietang@tsinghua.edu.cn

- pdf: https://www.ijcai.org/Proceedings/2020/0677.pdf
- data&code: https://github.com/HeterogeneousGraph