Yandex Research

Graph Self-Supervised Learning for Node-Level Prediction

Eremeev Dmitry

Plan

Introduction and Problem Setup

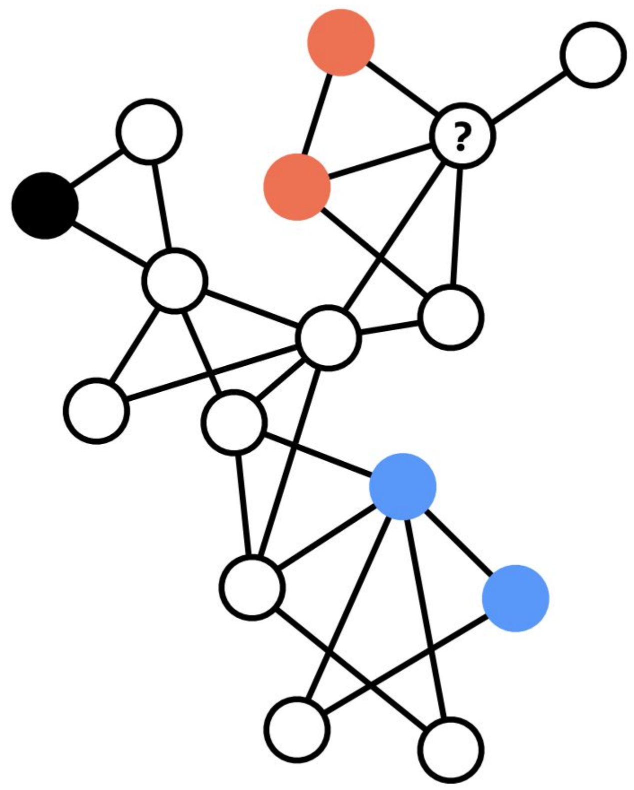
Graph SSL Methods

Our Current Results

Introduction and Problem Setup

Setup

- One graph
- All nodes have features
- Some nodes are labeled
- Need to predict labels for other labels
- Split: 10% train, 10% val, 80% test



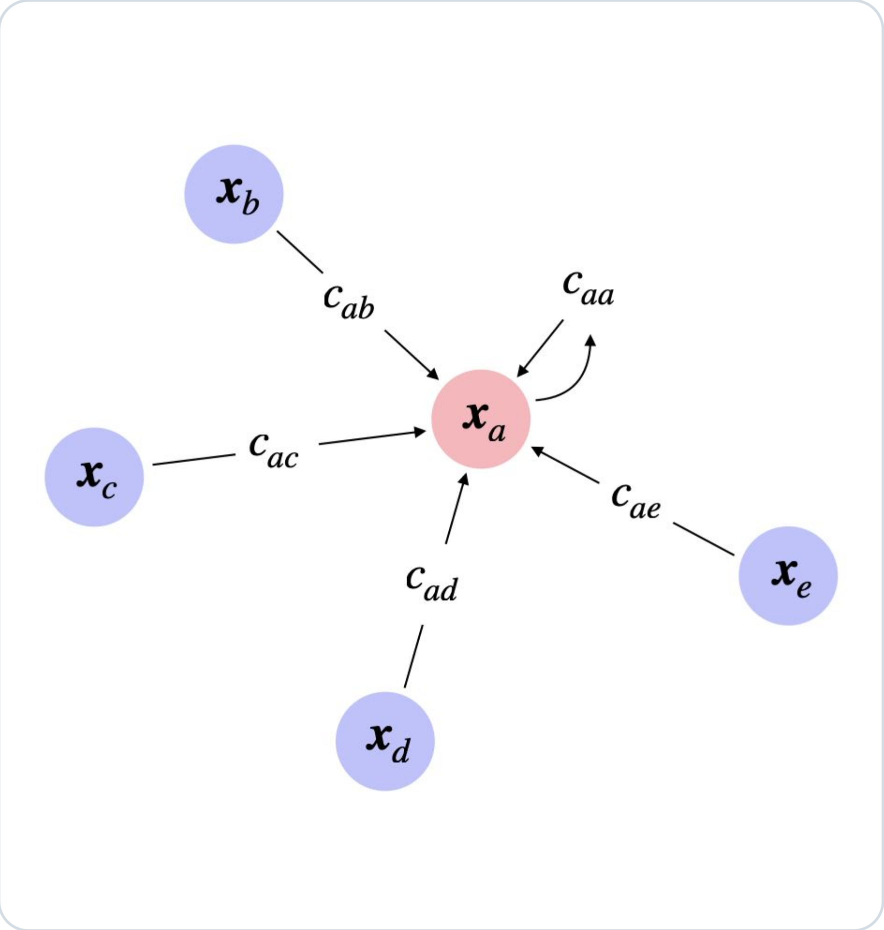
GNNs

- GCN

$$\boldsymbol{h}_v^l = \sigma\Big(\sum_{u \in \mathcal{N}(v) \cup \{v\}} \frac{1}{\sqrt{\hat{d}_u \hat{d}_v}} \boldsymbol{h}_u^{l-1} \boldsymbol{W}^l\Big)$$

- GraphSAGE

$$\boldsymbol{h}_v^l = \sigma(\boldsymbol{h}_v^{l-1} \boldsymbol{W}_1^l + (\text{mean}_{u \in \mathcal{N}(v)} \boldsymbol{h}_u^{l-1}) \boldsymbol{W}_2^l)$$



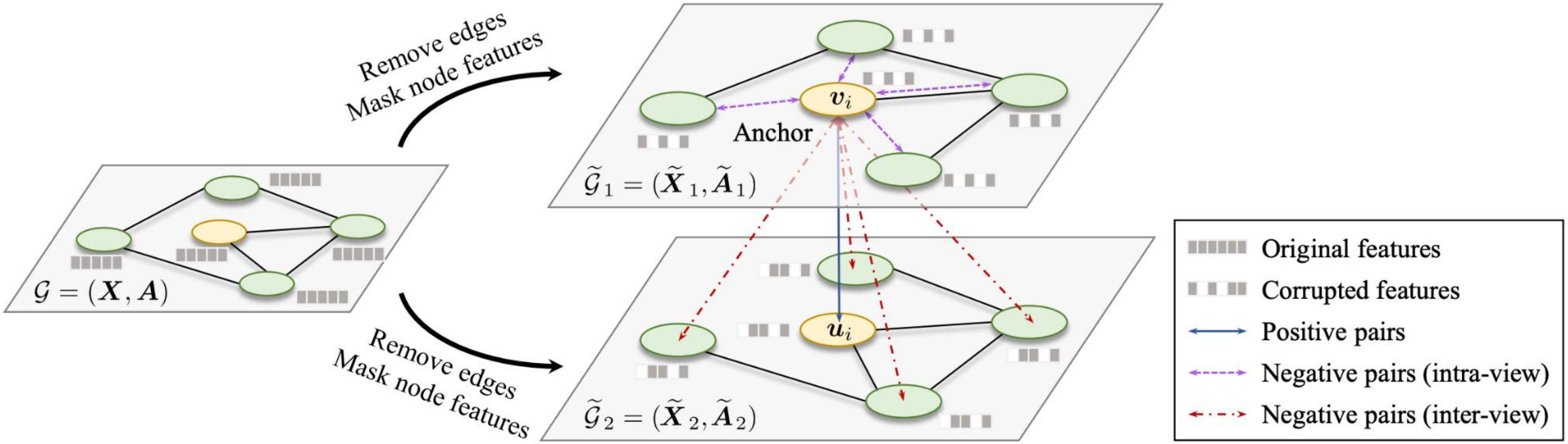
Graph SSL Methods

Key approaches

Contrastive methods

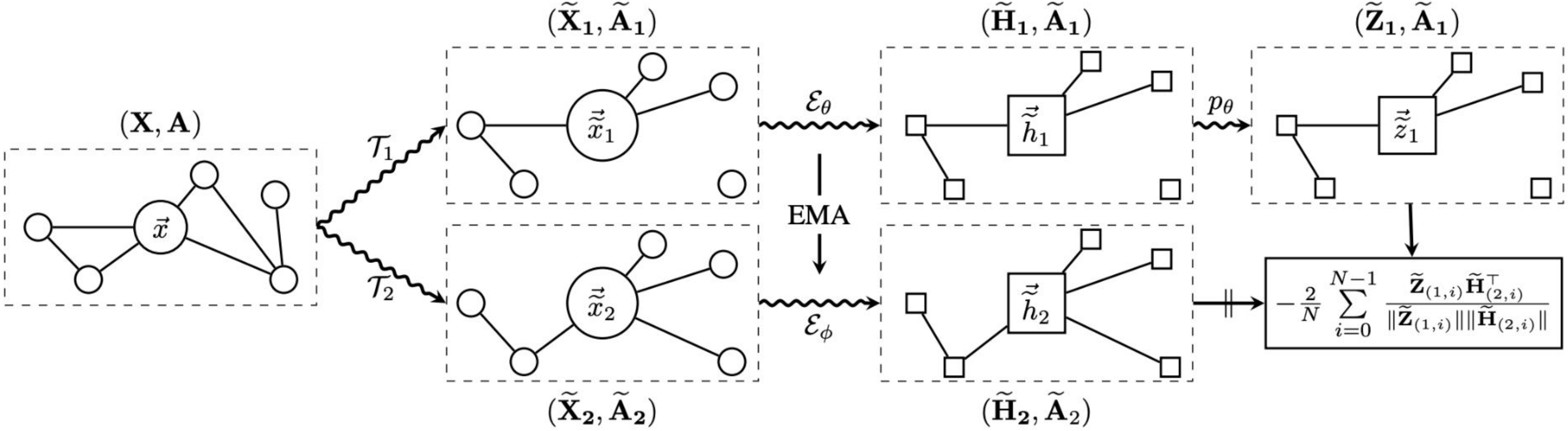
Generative (self-prediction) methods

GRACE



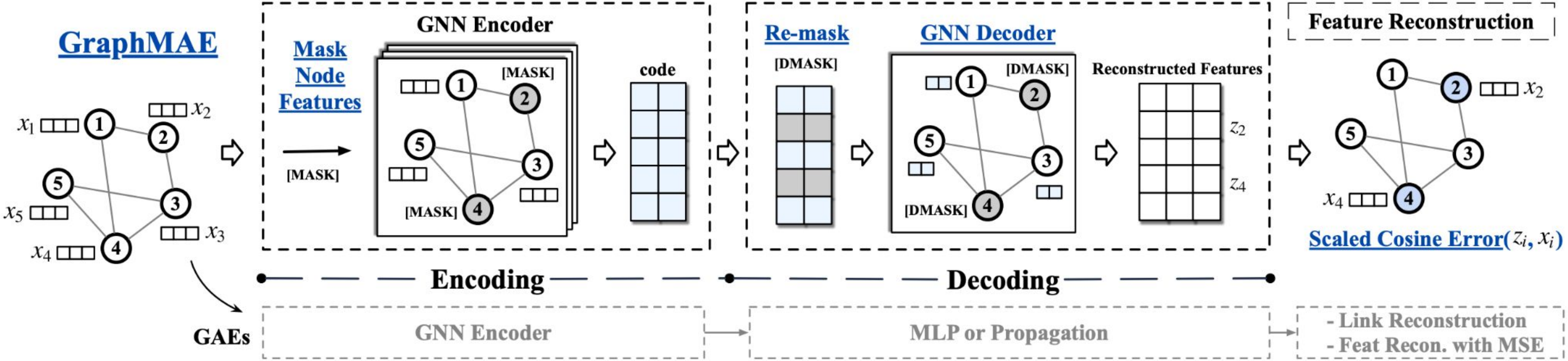
Zhu, Yanqiao, et al. "Deep graph contrastive representation learning." arXiv preprint arXiv:2006.04131 (2020).

BGRL



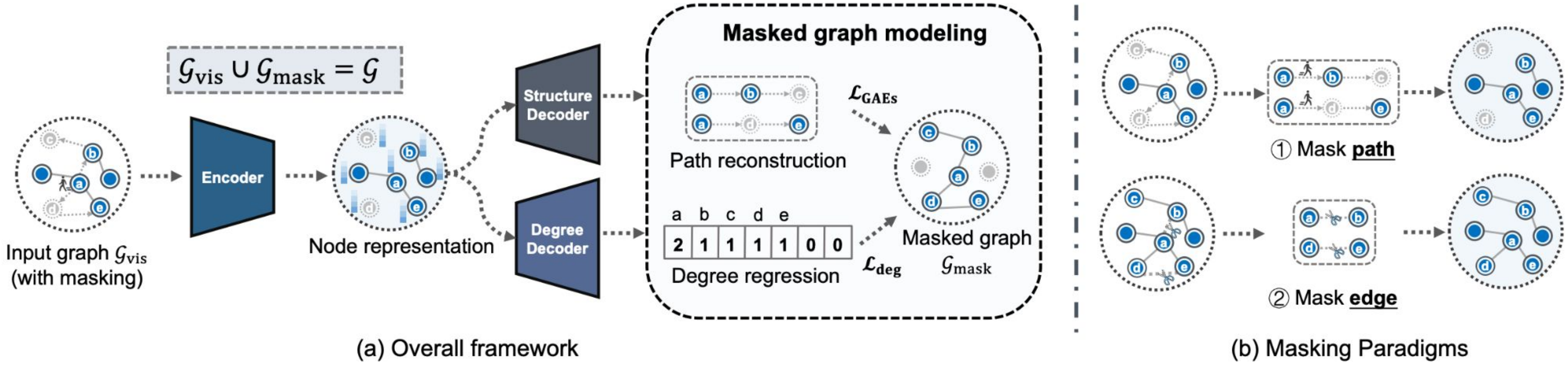
Thakoor, Shantanu, et al. "Large-scale representation learning on graphs via bootstrapping." arXiv preprint arXiv:2102.06514 (2021).

GraphMAE



Hou, Zhenyu, et al. "Graphmae: Self-supervised masked graph autoencoders." Proceedings of the 28th ACM SIGKDD conference on knowledge discovery and data mining. 2022.

MaskGAE



Li, Jintang, et al. "What's behind the mask: Understanding masked graph modeling for graph autoencoders." Proceedings of the 29th ACM SIGKDD Conference on Knowledge Discovery and Data Mining. 2023.

Our Current Results

Evaluation Setup

Diverse datasets

Architecture enhancements

Hyperparameter tuning

Unified setup

Luo, Yuankai, Lei Shi, and Xiao-Ming Wu. "Classic gnns are strong baselines: Reassessing gnns for node classification." arXiv preprint arXiv:2406.08993 (2024).

Evaluation Results

Metrics	AR	cora	citeseer	pubmed	lastfm-asia	facebook	amazon-photo	amazon-computers	tolokers-tab	questions-tab	amazon-ratings
GT-sep Baseline	6.50 ± 1.36	80.99 ± 0.60	70.16 ± 0.57	85.92 ± 0.22	82.83 ± 0.57	92.93 ± 0.18	94.00 ± 0.23	88.80 ± 0.34	55.44 ± 3.38	81.15 ± 1.69	40.89 ± 0.45
MaskGAE LP	6.95 ± 1.85	78.42 ± 1.13	70.08 ± 0.98	86.78 ± 0.22	85.55 ± 0.15	91.40 ± 0.21	93.35 ± 0.16	88.45 ± 0.66	58.59 ± 0.79	76.01 ± 16.68	40.80 ± 0.42
GraphMAE LP	8.30 ± 1.00	80.48 ± 0.72	69.06 ± 0.54	85.15 ± 0.29	80.87 ± 0.39	89.67 ± 0.39	93.31 ± 0.29	90.34 ± 0.21	49.29 ± 1.67	80.97 ± 1.23	39.96 ± 0.52
BGRL LP	4.95 ± 2.33	79.46 ± 1.36	69.52 ± 1.24	86.78 ± 0.20	84.36 ± 0.26	93.22 ± 0.31	94.34 ± 0.14	90.94 ± 0.25	59.62 ± 0.89	84.87 ± 0.60	40.20 ± 0.32
GRACE LP	3.00 ± 2.28	82.25 ± 0.63	71.05 ± 0.68	87.93 ± 0.23	84.08 ± 0.29	93.69 ± 0.10	94.44 ± 0.19	91.29 ± 0.22	60.44 ± 1.46	84.60 ± 0.53	40.49 ± 0.23
MaskGAE FullFT	4.30 ± 2.33	80.60 ± 1.16	68.80 ± 0.58	87.40 ± 0.30	86.07 ± 0.29	93.02 ± 0.24	94.20 ± 0.27	90.70 ± 0.27	60.20 ± 2.59	83.94 ± 0.62	42.94 ± 0.36
GraphMAE FullFT	4.40 ± 1.56	81.01 ± 0.99	71.15 ± 0.55	86.99 ± 0.13	84.23 ± 0.55	93.74 ± 0.16	94.32 ± 0.22	90.58 ± 0.09	53.64 ± 4.48	81.47 ± 1.79	42.58 ± 0.31
BGRL FullFT	3.80 ± 2.04	80.72 ± 0.78	71.84 ± 0.61	87.52 ± 0.12	85.78 ± 0.29	94.02 ± 0.18	94.35 ± 0.25	89.90 ± 0.32	52.16 ± 1.54	84.57 ± 0.63	40.98 ± 0.49
GRACE FullFT	2.80 ± 1.72	81.35 ± 1.16	72.19 ± 0.48	87.46 ± 0.17	86.44 ± 0.42	94.35 ± 0.18	94.81 ± 0.26	90.49 ± 0.33	56.23 ± 2.90	82.41 ± 1.91	40.93 ± 0.41

Analysis Setup

Real graph with synthetic features
(features are possibly independent of graph)

MLP-probing on top of GNN embeddings

Measure quality of reconstruction of features, edges
and probably other characteristics

Analysis Results

GRACE and BGRL with StrucAug improve feature reconstruction and vice versa for FeatAug

GraphMAE improves feature reconstruction

MaskGAE improves edge reconstruction

GraphRec (Working Title)

Use both features and structure in generative method

Simply combining MaskGAE and GraphMAE already outperforms them

With further improvements becomes SOTA

Summary

- Graph SSL is a useful approach for tasks with low label ratio
- Always remember to properly tune baselines
- Combining both features and graph structure in SSL objective can be crucial