Weilin Cong

congweilin95@gmail.com CongWeilin Website

Ph.D. candidate at the Department of Computer Science and Engineering, Pennsylvania State University. Research focuses on both the fundamental problems in graph representation learning (including optimization, generalization, and expressive power) and model architecture design. Published in the top-tier AI conferences NeurIPS, KDD, ICLR, AISTATS, and SDM on graph representation learning.

Experience

- **Pennsylvania State University** — **State College, PA**
  Graduate Assistant — *Jan 2018–Present*
  - CSE 597 Large-scale Machine Learning TA
  - D.A.T.A. Lab RA (Spring & Summer Semester)

- **Meta AI (Facebook AI Research)** — **Menlo Park, CA, United States**
  Research intern — *May 2022 - Present*

- **Meta AI (Facebook AI Research)** — **Remote, United States**
  Research intern — *Jun 2021 - Aug 2021*

- **China Mobile Research Institute** — **Beijing, China**
  Deep Learning Software Engineer — *Sep 2016 - Jun 2017*
  - Mobile face recognition application development.

Education

- **Pennsylvania State University** — **State College, PA**
  Ph.D. (changed degree from M.S.) in Computer Science Engineering — *Aug 2020 - Present*
- **Pennsylvania State University** — **State College, PA**
  M.S. in Computer Science Engineering — *Aug 2017 - Aug 2020*
- **Beijing Institute of Technology** — **Beijing, China**
  B.S. in Computer Science — *Sep 2013 - Jun 2017*

Most Recent Research Projects

- **Graph representation unlearning (AISTATS23)** — *Dec 2021 - Present*
  - We propose algorithms to efficiently remove the effect of a set of nodes on the pre-trained GNN model.
  - Algorithm GraphEditor: We first formulate GNNs as an alternative problem with closed-form solution, then we can edit the weight parameters based on the change of graph structure. Paper
  - Algorithm Projector: Based on our observation that the linear GNN's weight parameters (trained with logistic regression) are in the span of all node features, we propose to unlearn by projecting the original weight paramters into a subspace that is irrelevant to the deleted node features. The paper is accepted by AISTAT23 with very high rating. Paper
- **Dynamic graph learning (SDM23, ICLR23-Oral)** — *Jun 2021 - Present*
  - Propose a Transformer-based dynamic graph learning method named Dynamic Graph Transformer (DGT). To improve the generalization ability, we introduce two complementary self-supervised pre-training tasks and show that jointly optimizing the two pre-training tasks results in a smaller Bayesian error rate via an information-theoretic analysis. We also propose a temporal-union graph structure and a target-context node sampling strategy for an efficient and scalable training. This work is accepted by SDM23. Paper
  - We propose a conceptually and technically simple architecture that consists of three components: (1) a *link-encoder* that is only based on multi-layer perceptrons (MLP) to summarize the information from temporal links, (2) a *node-encoder* that is only based on neighbor mean-pooling to summarize node

information, and (3) an MLP-based *link classifier* that performs link prediction based on the outputs of the encoders. Despite its simplicity, our proposal attains an outstanding performance on temporal link prediction benchmarks with faster convergence and better generalization performance. This work is accepted by ICLR as Oral paper (notable-top-5% paper). Paper

- **Distributed GNN training (ICLR22)** *Dec 2020 - Oct 2021*
  - Propose a communication-efficient distributed GNN training technique. Provide theoritical analysis on the convergence of the proposed distributed GNN training method, and shows its benefit over naively parameter average type of methods. Paper
- **Understand performance degradation in deeper GNN (NeurIPS21)** *Dec 2020 - Jun 2020*
  - Reveal the discrepancy between the theoretical understanding of over-smoothing and the practical capabilities of GCNs. Analyze the generalization capability of GCNs, and show that the training strategies to achieve high training accuracy significantly deteriorate the generalization capability of GCNs. Propose a decoupled architecture that enjoys a provable better expressive and generalization power. Paper
- **Minimal variance sampling graph neural networks (KDD20)** *Aug 2019 - Fed 2020*
  - Propose a decoupled variance reduction strategy that employs gradient information to adaptive sample nodes with minimal variance, while explicitly reducing the variance introduced by node embedding approximation. Accepted as a full paper by KDD20 research track. Paper
- **Efficient training graph neural network by lazy sampling (NeurIPS20)** *Apr 2020 - Oct 2020*
  - Propose a framework that efficiently train GCN by lazy sampling, i.e., sampling periodically and effectively recycle the sampled nodes to mitigate data preparation overhead. The proposed framework enjoys superior speedup without compromising the accuracy. Accepted by NeuralIPS20. Paper
- **Encrypted rich-data steganography using GAN (MiseML20)** *Aug 2019 - Aug 2020*
  - Propose an end-to-end generative adversarial network model to encode data types of different modalities, e.g., text, audio and image, and is able to hide message deeply into a cover image without being detected and decoded by a third-party adversary who is not given permission to access the message. Paper.
- **Doubly variance reduced graph neural networks** *Aug 2019 - May 2022*
  - Provide the first convergence analysis of sampling-based GCN training under weak assumptions. Propose and analyze a general *doubly variance reduction* schema that can accelerate any sampling-based GCN training method under the memory budget for GCN training. Paper is under submission. Paper

Technical and Personal skills

- **Related Fields:** Machine Learning Engineering, Theoretical Machine Learning
- **Programming Languages:** Python
- **Machine Learning Tools:** PyTorch