

マルチ GPU 環境における GNN の高速・高精度学習手法の開発

Development of Fast and Accurate GNN Training Methods in Multi-GPU Environments

小林 諒平
東京科学大学

1. 研究目的

グラフニューラルネットワーク (Graph Neural Network, GNN) は, SNS, 交通網, 購買履歴, 化合物・材料構造, 知識グラフなど, 多様なグラフ構造データを扱うための基盤的な機械学習手法である. 一方で, 大規模グラフに対する GNN 学習では, 汎化精度の向上と学習時間の短縮を同時に達成することが難しい.

汎化精度の面では, 多くのメッセージパッシング型 GNN が近傍集約に基づいてノード表現を更新するため, 遠方ノードの情報や大域的構造を十分に取り込むことが難しい. また, 実行性能の面では, GNN 学習が疎なグラフ構造に基づく近傍探索, 特徴量取得, 疎行列-密行列演算, GPU 間通信を多く含むため, 画像や言語モデルに比べてメモリアクセスおよび通信の不規則性が大きい. 特にマルチ GPU 環境では, 分割されたグラフの境界をまたぐ特徴量取得や中間表現の通信が性能ボトルネックとなる.

本研究では, このような GNN 学習における「汎化精度」と「学習時間」の二つの課題を同時に扱うことを目的とした. 具体的には, 汎化精度向上の手段としてノードに Random Features (RF) を付加する手法を検討し, 学習時間短縮の手段として通信回避型 GNN 学習手法である CAGNET を活用した. 2025 年度は, RF と CAGNET を統合したマルチ GPU 向け GNN 学習コードを実装し, PyTorch Geometric (PyG) を基盤とする分散ミニバッチ学習実装との比較を通じて, 提案方針の有効性と課題を明らかにすることを目標とした.

2. 研究成果の内容

2025 年度は, RF と CAGNET を組み合わせた GNN 学習手法の実装, 分散ミニバッチ学習基盤の整備, および評価実験を実施した.

まず, 既存の CAGNET 分散サンプリング実装を基盤として, ノード特徴量に RF を付加する機能を追加した. RF の次元数, 分布, スケール, 乱数シード, 再サンプリング頻度を変更できるようにし, 静的 RF, エポックごとの RF 再サンプリング, ミニバッチごとの RF 再サンプリングを比較可能な構成とした. これにより, RF が GNN の汎化精度に与える影響を, 分散学習環境上で評価できるようになった.

次に, PyG を基盤とした分散 NeighborLoader ベースラインを整備した. フルグラフを各 GPU に複製するのではなく, 分割済みグラフストアを利用して各 GPU が担

当するパーティション上で近傍サンプリングを行う構成を採用し、通常の分散ミニバッチ学習、CAGNET 系の通信回避を組み込んだ分散学習、RF を付加した分散学習を比較可能な評価基盤を構築した。

評価を通じて、RF については重要な知見が得られた。RF は GNN の表現力を補強する手段として有望であり、一部のデータセットや構造的条件下では汎化精度向上に寄与する可能性がある。一方で、その効果は対象データセットの構造、タスク設定、入力特徴量の性質、RF の次元数や再サンプリング方式に強く依存することが分かった。また、RF を付加すると入力特徴次元が増加するため、特徴量取得、GPU 間通信、疎行列-密行列演算のコストも増加する。したがって、RF は常に用いるべき汎用的な精度向上手段というよりも、データセットやタスクの性質に応じて選択的に利用すべき拡張手法であると整理した。

一方で、CAGNET および分散ミニバッチ学習基盤については、今後の GNN 学習高速化研究の中核となる実験環境を整備できた。GNN 学習では、近傍サンプリング、特徴量取得、サブグラフ構築、GPU 間通信、勾配同期が複雑に絡み合うため、単純なデータ並列だけでは性能が頭打ちになりやすい。本研究により、RF の有無、サンプリング条件、プロセスグリッド構成、混合精度実行などを比較可能な実装基盤が得られた。これにより、当初想定した RF+CAGNET 統合手法の有効性と限界を明確化し、今後は RF に過度に依存せず、より汎用性の高い分散ミニバッチ GNN 学習の高速化へ研究方針を発展させる見通しを得た。

3. 学際共同利用プログラムが果たした役割と意義

本研究では、マルチ GPU 環境における GNN 学習の性能評価が不可欠であった。GNN 学習の性能ボトルネックは、単一 GPU 環境では十分に観測できない場合が多い。特に、GPU 間通信、分散サンプリング、リモート特徴量取得、パーティション間の負荷不均衡、勾配同期の影響は、実際のマルチ GPU 環境で評価する必要がある。

筑波大学計算科学研究センターの学際共同利用プログラムにより、NVIDIA H100 GPU を搭載した Pegasus 等の計算資源を利用できたことで、RF+CAGNET 統合実装、PyG 分散ミニバッチ学習実装、分散 NeighborLoader ベースラインの動作確認と比較評価を進めることができた。2025 年度の実利用は主に Pegasus 上で行い、ローカル環境では困難なマルチ GPU 実行、分割グラフを用いたサンプリング、通信を含む性能評価を実施した。

また、本課題で整備した GNN 学習評価環境は、ウェアスケールプロセッサにおける GNN 学習の処理能力と学習効率を評価する研究にも活用された。具体的には、東京科学大学小林研究室の綿貫晃雅氏による卒業論文発表「ウェアスケールプロセッサにおける GNN 学習の処理能力と学習効率に関する研究」における評価環境として利用された。GPU クラスタ上の GNN 学習と、Cerebras CS-3 をはじめとする AI アク

セラレータ上の GNN 学習を比較するための基盤としても、本課題で得られた知見は有用である。

4. 今後の展望

2025 年度の研究を通じて、RF は GNN の表現力を補強する有望な手段ではあるものの、汎用的な精度向上手法として常に有効とは限らないことが分かった。今後は、RF を提案手法の必須構成要素として扱うのではなく、データセットやタスクの構造的性質に応じて選択的に利用する拡張手法として位置づける。

一方で、研究の主軸は、分散ミニバッチ GNN 学習におけるサンプリング、特徴量取得、GPU 間通信、データ配置の最適化へ移す。フルグラフ学習は大規模 GPU クラスタ上で高い性能を達成できる可能性があるものの、利用可能な計算資源に対する要求が大きい。より広い利用環境で実用的な GNN 学習基盤を実現するためには、ミニバッチ学習を前提に、time-to-accuracy を改善する方向が重要である。

具体的には、CAGNET 系の通信回避アルゴリズムを分散ミニバッチ学習により自然に適用する手法を検討する。複数ミニバッチをまとめて扱う bulk sampling, サンプリングされたノード集合の局所性を利用した特徴量取得、分割グラフ上でのリモート特徴フェッチ削減、GPU 間通信と計算のオーバーラップ、プロセスグリッド構成の自動選択などが重要な検討課題となる。また、RF に代わる、あるいは RF を補完する精度向上手段として、構造特徴量、位置エンコーディング、サンプリング戦略の改善、正則化、データ拡張なども比較対象に加える。

さらに、2026 年度以降は、NVIDIA GPU 環境に加えて、AMD GPU や Cerebras CS-3 をはじめとする AI アクセラレータシステムとの比較も進める。GNN 学習は疎で不規則なメモリアクセスを多く含むため、アーキテクチャ差が学習効率に与える影響を明らかにすることは、今後の GNN 向け計算基盤設計にとって重要である。

5. 成果発表

(1) 学術論文

なし

(2) 学会発表

なし

(3) その他

本課題を通じて、RF+CAGNET 統合実装、PyG 分散 NeighborLoader ベースライン実装、分散ミニバッチ GNN 学習評価コードを整備した。これにより、RF の有無、サンプリング方式、CAGNET 系通信回避、混合精度実行、GPU 数、分割数を変えた比較評価が可能となった。

また、本課題で整備した GNN 学習評価環境は、東京科学大学小林研究室の綿貫晃雅氏による卒業論文発表「ウェハスケールプロセッサにおける GNN 学習の処理能力と学習効率に関する研究」における評価環境としても活用された。

筑波大学計算科学研究センター 2025 年度学際共同プログラム利用報告書

使用計算機	使用計算機に○	配分リソース※		
		当初配分	移行*	一般利用による追加
Pegasus	○	4000		0
Miyabi-G	○	13500	0	0
Miyabi-C				
※配分リソースについてはノード時間積をご記入ください。 *バジェット移行を行った場合、「+2000」「-1000」のように記入				