Hybrid Parallelism for GPU-Accelerated Representation Learning on Geospatial and Graph-Structured Data

Abstract

Large-scale representation learning on geospatial and graph-structured data underpins digital twins, urban infrastructure monitoring, and spatial AI, yet existing parallel training frameworks are designed primarily for sequences, images, and text. We propose a hybrid parallelism framework that jointly exploits data distribution, graph topology, and hardware heterogeneity to accelerate representation learning on multi-GPU and cloud-native clusters. Our approach introduces (i) a graph-structure-aware parallelism taxonomy and scheduler that composes data, tensor, and pipeline parallelism with topology-guided partitioning, (ii) GPU-kernel and memory-layout optimizations for hybrid IVF–graph index construction and query processing during training, and (iii) adaptive communication and batch-sizing strategies that respond to data skew, graph connectivity, and device constraints. We instantiate this framework for geospatial and urban digital-twin workloads, demonstrating order-of-magnitude reductions in training time while preserving or improving representation quality. Our results establish a new systems foundation for scalable representation learning on graph-rich, spatially structured data and open pathways for real-time digital twins and large-scale spatial AI.

Keywords: representation learning, parallel training, graph neural networks, geospatial AI, GPU systems, digital twins, vector databases, distributed optimization

Reviewers: please read the Reviewer Guidelines (iclr.cc/Conferences/2027/ReviewerGuidelines) and the AI Policy for Reviewers (iclr.cc/Conferences/2027/AIPolicyForReviewers). If you used AI to expand, edit, or polish your review, please provide the input text to the LLM. Better still, consider skipping the LLM and submitting your original text: we, and the authors, are much more interested in your unedited thoughts than in what an LLM has to say. AI-assisted or not, you are putting your name and reputation behind your review: LLM-generated falsehoods, hallucinations or misrepresentations are subject to disciplinary action, which may include desk-rejecting all papers you have authored.