Data-driven Neuroscience: Enabling Breakthroughs Via Innovative Data Management
Alexandros Stougiannis, Mirjana Pavlovic, Farhan Tauheed, Thomas Heinis, Anastasia Ailamaki

Indexing Dense Spatial Datasets

Dataset: Rat Neocortex
Model Size: 1692 Neurons
Dataset Size: 12.5 GB

GOAL: Simulate Human Brain
Model Size: 86 Billion Neurons
Expected Dataset Size: 606 PB

Spatial Index: R-Tree: Hierarchy of Minimum Bounding Rectangles (MBR)
Increase in Spatial Data Density => More Overlap

Prefetching for Spatial Sequences

Existing Approaches:
Trajectory Extrapolation Does Not Work Well With Irregular Paths in Neural Networks

SCOUT: Content Aware Prefetching Approach
Examine: Query Results
Identify: Guiding Path
Predict: Using Guiding Path

Effective Prefetching Requires Accurate Prediction

In-Memory Spatial Join

TOUCH: Hierarchical Data Oriented Partitioning
1) Avoid replication through data-oriented partitioning
2) Avoid overhead of overlap through direct assignment and batched join
3) Use filtering

Locating Synapses = Spatial Join!