PP Final Project:

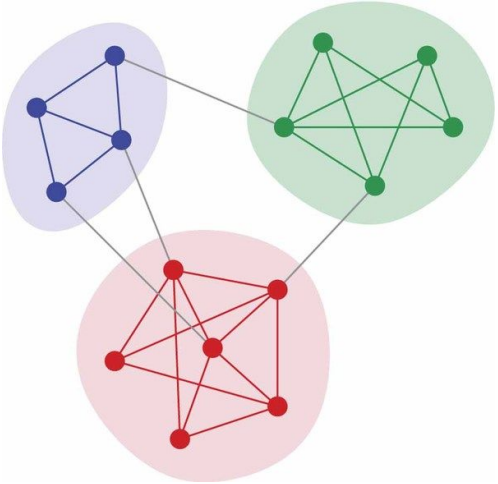
Large Graph Visualization

Team03
楊依辰、郭芳妤

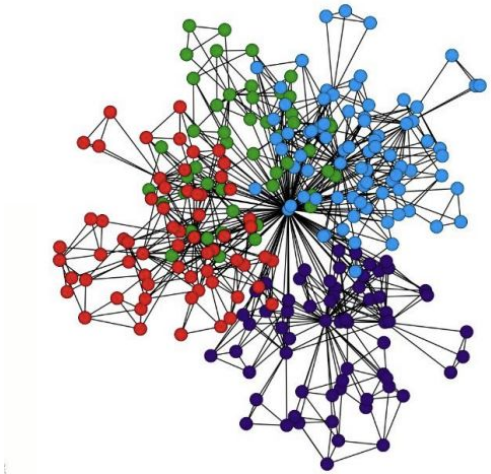
Problem Description: large graph is hard to visualization

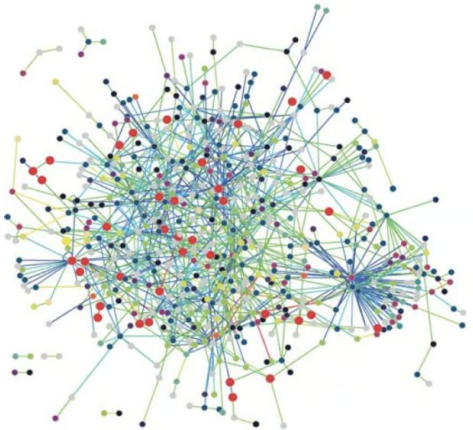
- Graph = (V, E) is a data structure composed by node and edge
- Large graph visualization take a lot of time to decide the layout.

when #node is small



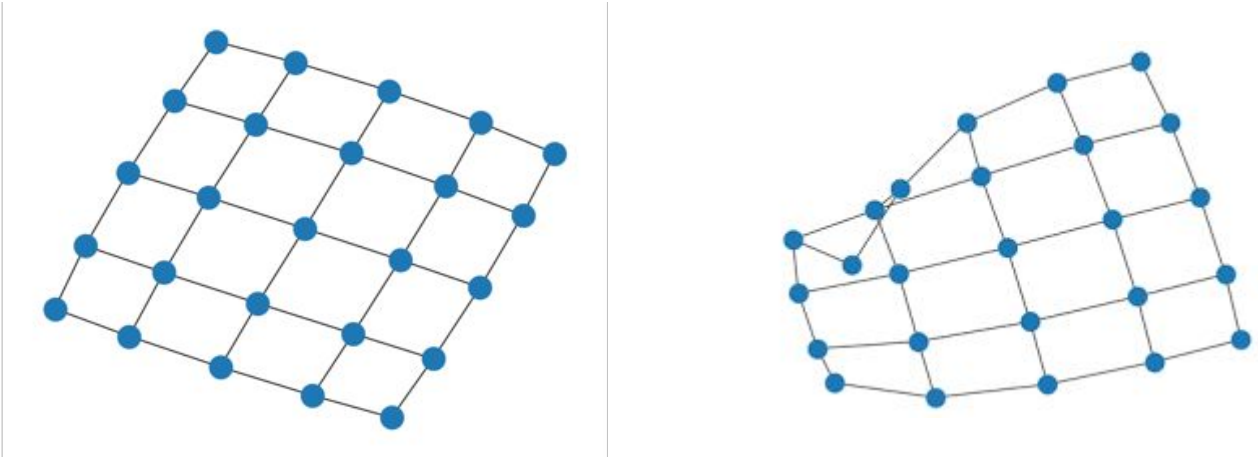
when #node is large





Current Tools and challenge

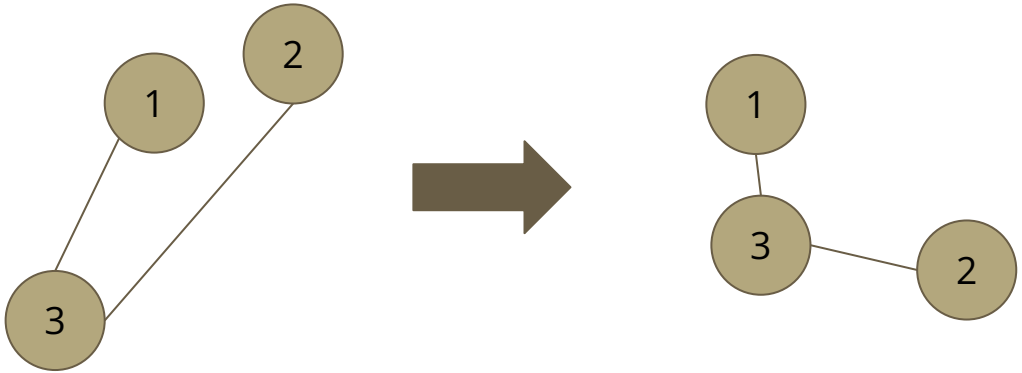
- Current Graph Visualization tools (e.g. networkX, graphvis, echart) provide layout decision algorithm if user did not specify the node position beforehand.



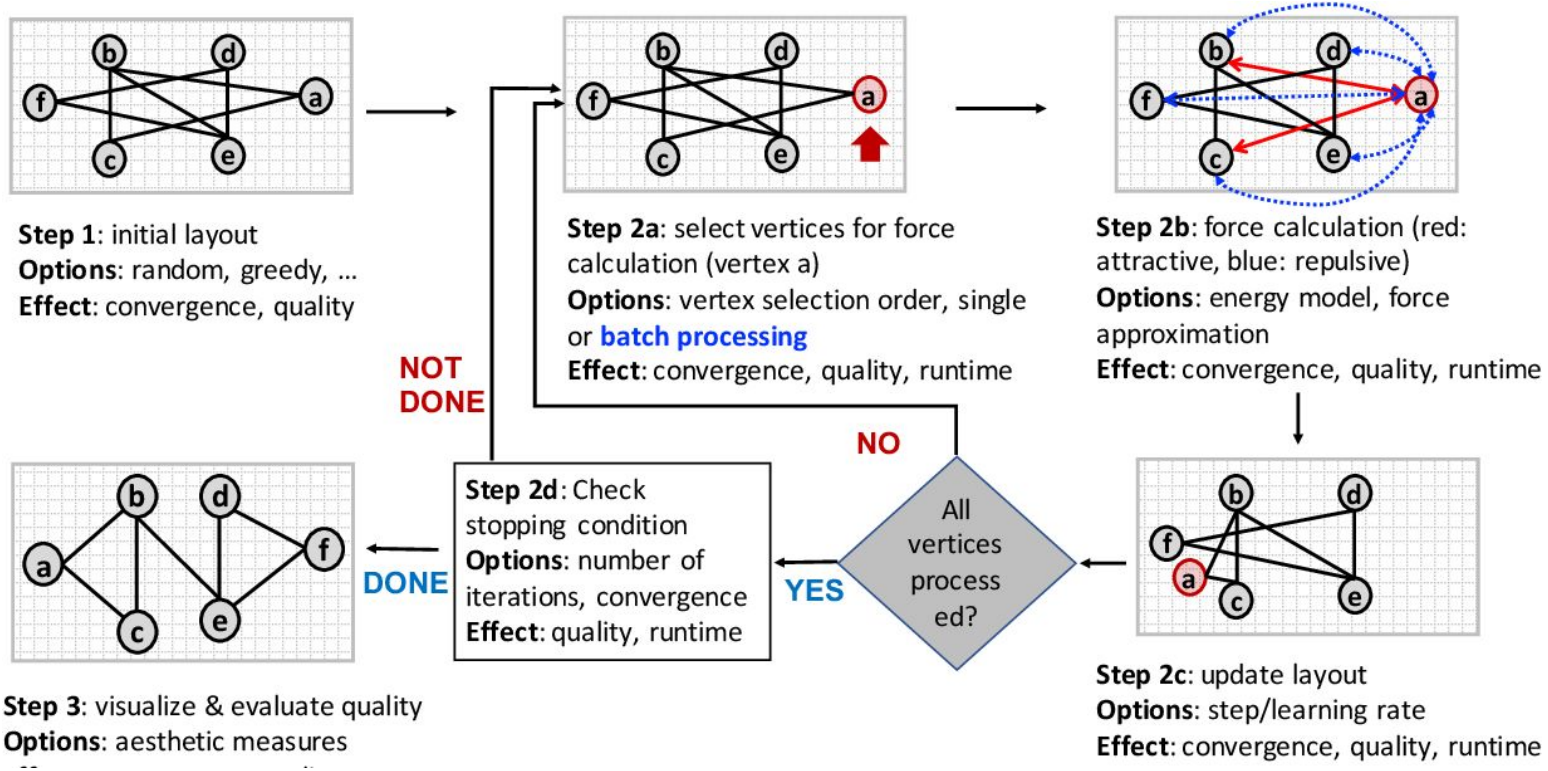
Existing Algorithm: Force-Directed Algorithm

Decide the node position by **force computing**

1. the attractive forces between adjacent node pairs
2. the repulsive forces between nonadjacent vertices

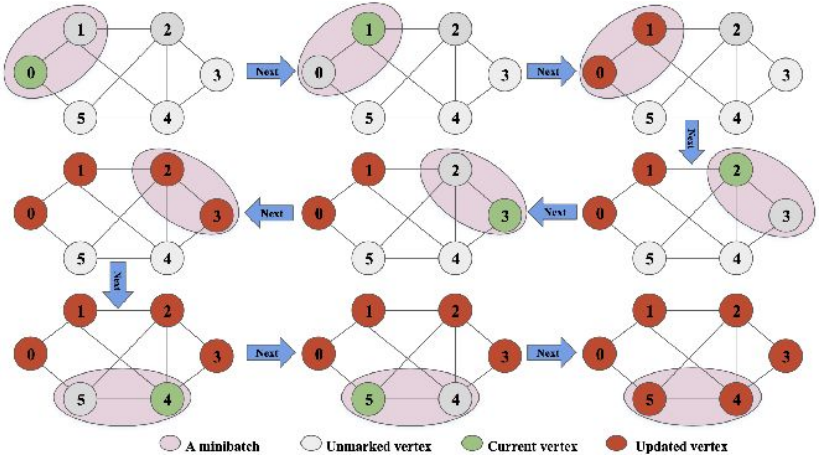


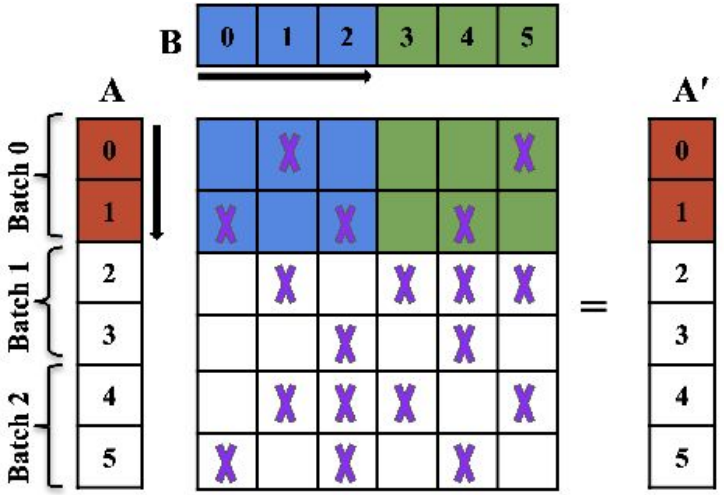
Force-Direceted Algorithm



Optimized Version: Batch-Parallel FD Algorithm

- BatchLayout: A Batch-Parallel Force-Directed Graph Layout Algorithm in Shared Memory (with code)



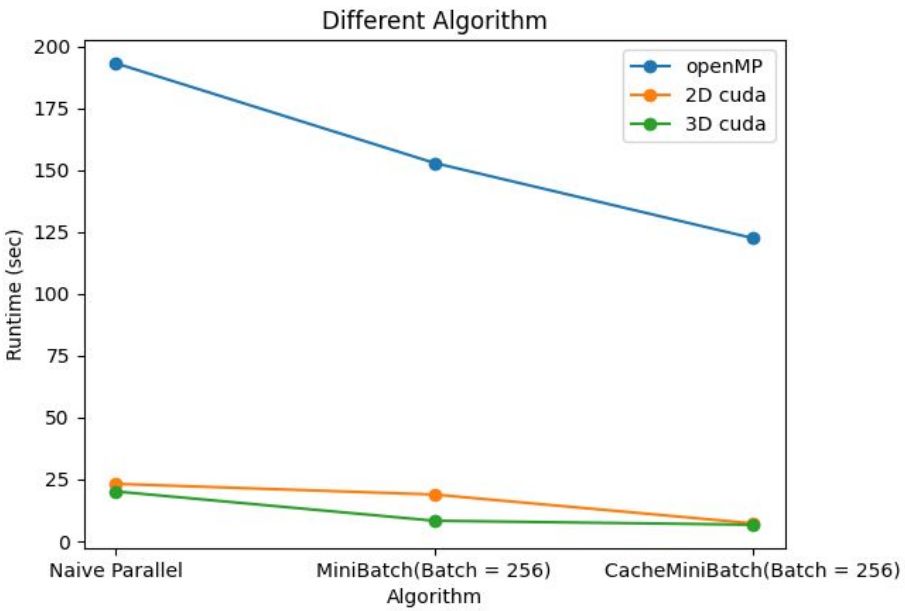


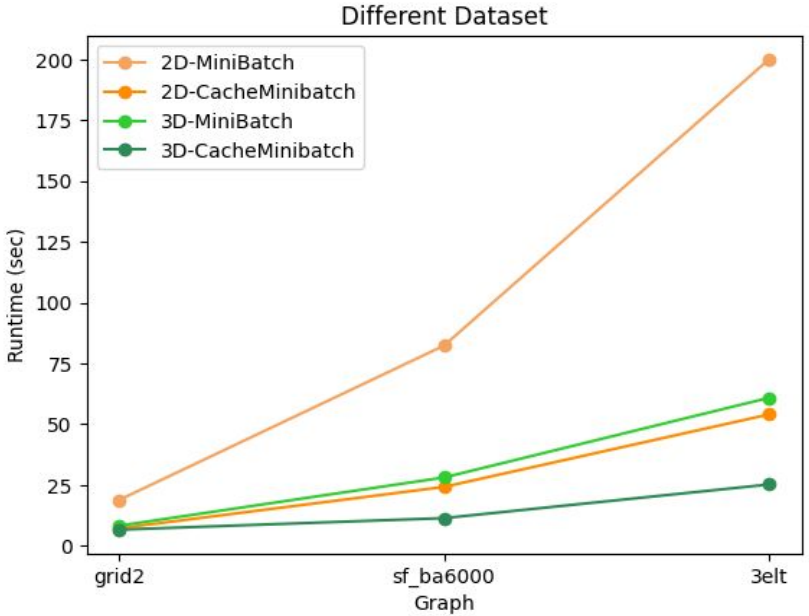
https://ieeexplore.ieee.org/document/9086287

Implementation items in Final Project

1. openmp (in source code) → **cuda version**
2. 2D → **3D implementation**
3. Experiments
    a. cuda vs. original (speedup)
    b. scalability with different datasize
    c. visualizatio quality with different batchsize

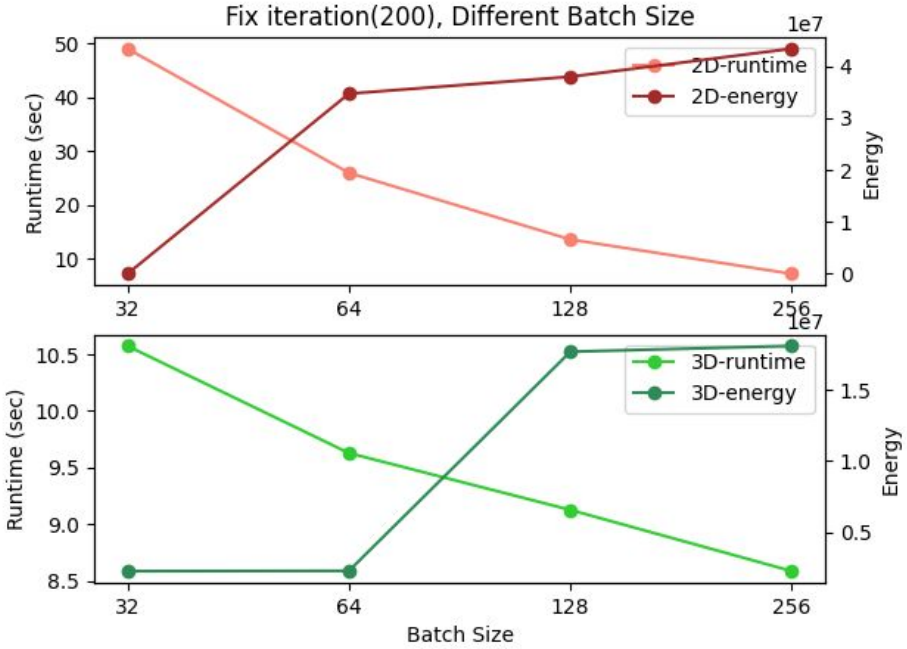
Experiment


Different Algorithm
openMP
2D cuda
3D cuda
Runtime (sec)
200
175
150
125
100
75
50
25
0
Naive Parallel
MiniBatch(Batch = 256)
CacheMiniBatch(Batch = 256)
Algorithm


Different Dataset
2D-MiniBatch
2D-CacheMinibatch
3D-MiniBatch
3D-CacheMinibatch
Runtime (sec)
200
175
150
125
100
75
50
25
0
grid2
sf_ba6000
3elt
Graph

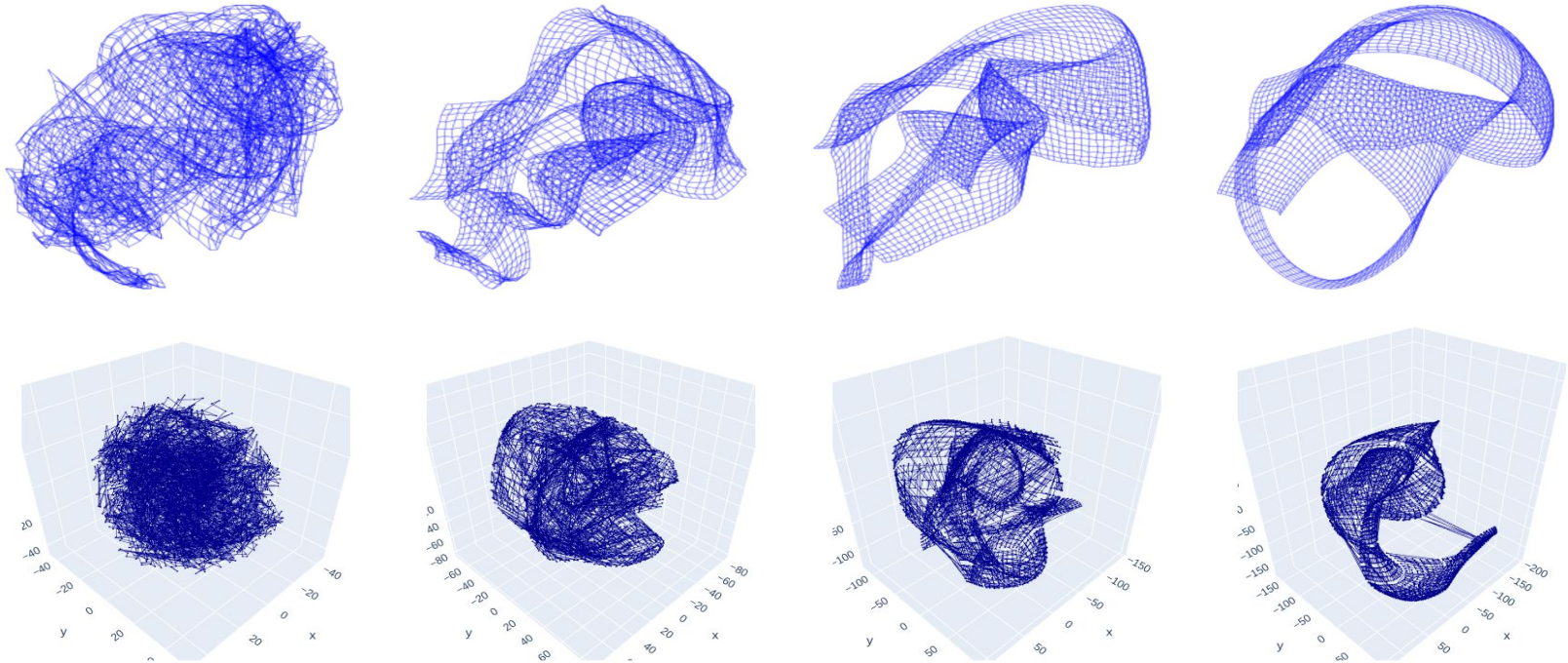
Experiment

隨著 batch size 增加，可以觀察到所需要的時間下降，但同時也意味著更新位置的頻率下降，影響最終 output 的精準度。



Visualization (different iteration)

50 → 100 → 200 → 400



Visualization (Demo)