

CSCI-4974/6971: Assignment 6 (100 pts)

Various Graph Processing Methods and Frameworks

Due Date: Thursday 8 Dec. 2016, 16:00

This assignment will just involve some short response questions based on Lectures 21–24.

1 Response Questions

In a single file, submit your responses to the following questions:

1. We calculated betweenness centrality in Assignment 3 for a static graph. How would you calculate it for a temporal graph, where the graph uses the “discrete” window model (i.e. static snapshots of the graph for various windows in time). Describe the algorithm and any assumptions you need to make.
2. You’re using label propagation to detect community structure of a changing graph. Assume you have an initial static graph, and then future updates in terms of edge additions/deletions (e.g. “edge between vertex 10 and vertex 55 is added”) are streamed in. Describe an algorithm that processes these updates. Make any assumptions you need to.
3. GraphBLAS is a push towards formulating graph analytics as linear algebraic operations. Describe at least two motivating factors.
4. There are many existing parallel graph processing frameworks. How each framework stores and processes a graph can be described as falling into one of the following classes: shared-memory, distributed-memory, or external-memory. Into which class does each of the following fit: CombBLAS, Giraph, GraphX, GraphLab, Ligra, FlashGraph?
5. As talked about in class, PowerGraph is a GraphLab derivative. What are the primary differences/improvements?
6. What is a 2D distributed graph layout? What is its primary benefit relative to a 1D layout?
7. Give at least three challenges with processing graphs on manycores such as GPUs and Xeon Phis.

This homework will be evaluated on your answers to these questions (14.2857... pts each).