805 Proposal: **Bohan Zhang** bohanzhang@cmu.edu

Subgraph enumeration is difficult task. It needs to list all subgraphs in the large graph, and it has numerous subgraph candidates. Existing parallel solutions are mainly based on two different models: Mapreduce join, and Pregel-like model(vertex-centric model). These two methods were both proposed and published in SIGMOD. However, for the Pregel model, it needs to store all subgraph instances in memory, which cannot be scaled. For the mapReduce join methods, it generates lots of invalid subgraph instances, and it also be limited to the memory. Thus, I would like to compare the existing solutions, and propose my new efficient parallel algorithm.

I will work with two 605 students. I plan to work with the data on SNAP , and begin with a small graph Twitter, which has 80 thousand nodes and 2 million edges. Also, I will also experiment on larger graphs like millions of nodes.