**RELAXED MODEL CHARACTERIZATION OF DEGREE-GUIDED TASK ASSIGNMENT WITH DATA LOCALITY CONSTRAINT**

By Junle Qian

[qian10@illinois.edu](mailto:qian10@illinois.edu)

Fall 2012

Advisor: Prof. Yi Lu

**ABSTRACT**

Cluster computing systems are popular in IT industries for data-intensive applications and services. In such systems, task file is divided in data chunks, which are replicated and distributed to a number of servers. Servers have to remotely acquire necessary file chunks that are not present locally before they proceed to relevant computation. Due to the fact that network connections have inevitable latencies, servers complete tasks with local data faster than those with remote data. Therefore, task assignment with data locality constraints plays an essential role in cluster computing systems.

In this research, we have proposed a degree-guided assignment algorithm that significantly outperforms Random Server Algorithm at light load. However, the performance of this algorithm drops sharply when system reaches queuing threshold. We have begun to look for enhancements that solve the problem of performance plunge at high loads. Specifically, my work in this thesis concentrates on characterizing the performance of degree-guided algorithm in a relaxed model, where leftover tasks are considered as fresh ones in a time slot. This will allow us to characterize degree-guided algorithm from an ideal perspective.