

Thesis: Facilitating Big Data Processing for Main Stream Object-Oriented Software

Abstract: With the increasing demand for big data processing abilities of modern software systems, more techniques for facilitating big data processing are being integrated into mainstream Object-Oriented programming languages, frameworks, and platforms. In this dissertation proposal, an approach that can automatically suggest streams code to run in parallel or sequentially will be proposed. Stream APIs integrate with MapReduce-like operations so that developers could easily use them to process big data. Using streams efficiently and properly needs many subtle considerations. The use and misuse patterns for stream codes will also be proposed and discussed. Besides, modern software systems, especially for highly transactional software systems, generate a huge volume of data every day. The vast information in these logs can bring information overload to developers, which prevents them from receiving useful information effectively. The log level is used to filter run-time information. I propose an approach to empower developers to filter useful information by rejuvenating log levels from the code change history.

Committee:

- Professor Raffi Khatchadourian, Mentor, Hunter College
- Professor Subash Shankar, Hunter College
- Professor Ashwin Satyanarayana, NYCCT

Outside Member:

- Mehdi Bagherzadeh, Oakland University