

Data Analytic Tools for Inconsistency Detection in Large Data Sets

Week 4 Report

02/24/2018 – 03/09/2018

Client/Advisor: Kingland Systems, Ying Cai

Team Members/Role:

Logan Heitz – Project Lead

Christopher Konopka – Communication Lead

Camden Voigt – Technical Lead

Timothy Rogers – Quality Lead

Weekly Summary

Past week accomplishments

- Logan:
 - Fixed issue with unclosed jdbc statement in the storage module
 - Implemented threading module and added into existing modules to achieve multithreading for solution
 - Started work on utilizing AWS for better parallelization of the solution
- Christopher:
 - Updated our usage of tuples storing records so that keys with multiple values are properly utilized
 - Added in command line options that will better allow us to start the program in AWS
 - Started testing performance on AWS
- Camden:
 - Worked with advisor to get AWS instance setup and accessible by

- team
 - Finished front end prototype view for reporters
 - Finished REST API to fetch inconsistency information
 - Added documentation on getting AWS setup and running to Gitlab wiki
- Timothy:
 - Started removing SQL dependencies from tests
 - Started work on improving testing for multithreading

Pending issues

- Logan:
 - Continue parallelization of the solution using AWS
- Christopher:
 - Continue to gather performance metrics on AWS using current solution
- Camden:
 - Deploy rest API to AWS
 - Integrate AWS with front end prototype
- Timothy:
 - Continue improvements to testing module
 - Validate the current raw data parser using the XML schema that has been received from Kingland

Individual contributions

Team Member	Weekly Hours	Total Hours
Logan Heitz	12	106
Christopher Konopka	12	90.5
Camden Voigt	15	99
Timothy Rogers	15	102

Plan for coming week

- Logan:
 - Continue parallelization of the solution using AWS
- Christopher:
 - Continue to gather performance metrics on AWS using current solution
- Camden:
 - Deploy rest API to AWS
 - Integrate AWS with front end prototype
- Timothy:
 - Continue improvements to testing module
 - Validate the current raw data parser using the XML schema that has been received from Kingland

Summary of weekly advisor meeting

03/02/2018 Advisor Meeting

- Worked on getting all team members set up with AWS and ensuring we all can log in and know how to run the project.
- Christopher will begin doing the testing on AWS so we can figure out how AWS will affect our performance.
- Logan will look into how we can take advantage of AWS to better parallelize our solution.