sdmay18-27: Data Analytic Tools for Inconsistency Detection in Large Data Sets
Week 3 Report
September 23 - September 29

Team Members

Camden Voigt — *Technical Lead*
Christopher Konopka — *Communication Lead*
Logan Heitz — *Project Lead*
Timothy Rogers — *Quality Lead*

Summary of Progress this Report

Completed first version of the project plan and website.
Meet with advisor to discuss technical details of the project and next steps. We discussed two approaches we can take from here.
The first approach is to put our records into a standard SQL database after hashing them and index them on the attributes we will need to detect inconsistencies for. This method will be easier to implement, but could be slower.
The second approach is to develop our own binary file format for storing the records and use that to detect inconsistencies. This method will be significantly more difficult to implement and may not end up being efficient, but the optimizations may lead to faster comparisons.
We will move forward with method 1 and determine if it is fast enough or if we need to implement method 2. In either case we first need to create a parser that will take the daily reports and convert them to the appropriate format for our inconsistency detection. We will use a config file for the parser so that we can quickly change what attributes of the reports we want to get and which should be hashed.

Pending Issues

Develop format of the configuration file
Create outline of configuration parser
Create outline of raw data parser
Prototype configuration parser
Prototype raw data parser
Create system for doing comparisons in the SQL table and flagging inconsistencies
Update proposed solution with new design plans
Update project schedule with new information

Plans for Upcoming Reporting Period

Creation of configuration file structure
Update proposed solution with new design plans
Update project schedule with new information

Individual Contributions

Team Member	Contribution	Weekly Hours	Total Hours
Camden Voigt	Finished our initial proposed design and block diagram and added them to the project plan.	1.5	7
Christopher Konopka	Finished the initial problem statement and the literature review for our project plan.	2.5	5.5
Logan Heitz	Finished the project timeline and deliverables.	2.5	7.5
Timothy Rogers	Finished validation and acceptance testing and challenges for project plan. Updated team website with project plan.	2.5	7