sdmay18-27: Data Analytic Tools for Inconsistency Detection in Large Data Sets
Week 5 Report
October 7 - October 13

Team Members

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

Summary of Progress this Report

During this week, we have development initial documentation for the configuration parser and the raw data parser and begun some initial implementation on the configuration parser.

Configuration Parser

ConfigParser: The ConfigParser object will be used to read the configuration file and create the correct objects. Its main method will be parse which returns a Config object described below. It will also have several other methods for reading XML and creating methods. It will use C#'s built in XML parser to parse the XML file.

Config: Created by the ConfigParser. This object will hold lists of exports and key elements extracted from the configuration file

Export: Export is an object holding data on how the program should export data. This would generally fall into two categories database and files. If database is specified than login information must be provided.

Element: Element is an object representing different attributes that would be saved in either the database or the file. It will hold the attributes name and other optional names. It will also hold the desired output formats for that attribute.

Output Format: OutputFormat Objects will hold data on how an attribute should be outputed. This includes whether it should be raw data or hashed, how many bits, and which hash function should be used.

Raw Data Parser

rawDataParser: The rawDataParser object will take in a configuration object as an argument. It will then parse the files or files specified in by the file path within this object. It parses the specified relevant information, hashes if necessary, and then outputs the parsed information to the specified outputs.

See the full documentation of these for more information on methods and variables contained in all the classes.

Pending Issues

Prototype the configuration parser
    Development of the configuration parser class
    Development of the config object
    Development of the export object
    Development of the element object
    Development of the output format object
Prototype of the raw data parser
    Implementation of parsing raw data file
    Implementation of exporting data to output location
    Implementation of hashing class

Plans for Upcoming Reporting Period

Development of the configuration parser class
Development of the config object
Development of the export object
Development of the element object
Development of the output format object
Implementation of parsing raw data file
Implementation of exporting data to output location
Implementation of hashing class

Individual Contributions

Team Member	Contribution	Weekly Hours	Total Hours
Camden Voigt	Development of documentation for configuration parser. Started working on the prototype of configuration parser including working on the element and output format objects.	4	13
Christopher Konopka	Development of the raw data parser documentation and how the raw data parser will utilize the configuration object.	3	10
Logan Heitz	Development of documentation for configuration parser. Started working on the prototype of configuration parser including working on the config and export objects.	4	13.5
Timothy Rogers	Development of the raw data parser documentation and how the raw data parser will connect to the export location and send information.	3	11