**DESIGN DOCUMENT**

Created by: Jian Xian, Marshall Chang, Tina Do

**GENERAL OVERVIEW OF OUR SYSTEM:**

Our program is written in python3 bsddb3 module for Berkeley database and sqlite3.

DIAGRAM

**A description of your algorithm for efficiently evaluating queries, in particular evaluating queries with multiple conditions and wild cards and range searches and an analysis of the efficiency of your algorithm**

**USER GUIDE:**

**Phase 1:** Phase1.py is called in the command line by typing: python3 Phase1.py

The user will be prompted for a file name. If a file is not found, it will continue to ask for a valid input. Once a valid input is received, the program will run for a short time depending on the length of the file. After completion, it will output the time it completed the process

**Phase 2:** Phase2.py is called in the command line by typing: python3 Phase2.py

The program will use the four text files created from Phase1 and sort, remove duplicates, and reformat to be used by db\_load shell command to generate four index files.

Using the bash shell and script way:

In the linux command line or bash window. Navigate to the folder containing the files created from Phase 1. Type “./Phase2\_sort.txt” in the command line. It will sort the files and remove any duplicate lines. The files remain the same name

**Phase 3:**

**DESIGN OF SOFTWARE:**

**OUR TESTING STRATEGY:**

**GROUP WORK BREAK-DOWN AND DETAILS**

**Group Meetings:**

**Software Development- Includes Design, Development, and self-QA time of functions**

**Marshal:**

**Jianxiang:**

**Phase 2:**

**Tina:**

**Phase 1: ~2hrs**

**Phase 2: ~3hrs**

**QUALITY ASSURANCE TESTING:**