Database for predicting the probability of flight delay

Group member: Weiqing Wang, Xiangyu Cao, Xiaoyi Xie

**Project Introduction**

The purpose of this project is to create a database included a website which will provide a prediction of the probability of the flight delayed in the next week. In our daily life, traveling by air is one of the domain transportation ways in the US. Based on the searching results through the internet, there is barely information about predicting of the possibility of flight delayed on the certain flight. In our project, we will create a new database based on three existing databases: flight delay database[1], weather report database[2], calendar database[3] .Hope this new database will provide convincing evidence to help the prediction.

**Project Specification**

* Database Server

We will choose Amazon Cloud as the database server. And there are several possible database management systems like PSQL, MySQL. We tend to use MySQL for its popularity and various information.

* Website

The website we provided will be based on the PHP, we will use this opportunity to learn the language and understanding how it works with the MySQL.

* Source database details

When comes to a flight delay, the possible reason will be weather related, for example, storming, lighting, snowing, winding, heavy fog and so on, and the man-made reason such as flight schedule and takeoff priority. So we find this three database.

The first resource is the flight information database provided by United States department of transportation, this is our main resource. it has the date, flight number, scheduled time, delay time of the flight. Consider the scope of our project, we will focus on the flights that take off or landing on the St. Louis airport.

The second resource is about weather reports of St. Louis (including the history reports and weather forecast). It recorded the temperature, snow depth, wind speed and visibility values which may affect the takeoff and launch of planes. And we try to follow the recent information about the weather data, may using the API to make the records real-time update.

The third resource from the calendar of holidays national holidays and some peak weeks like thanksgiving and Christmas.

**Project Goal**

We plan to build a database based on the website which provides flight prediction service based on the data gathered from these database weather information database, previous flight delay information database and calendar database. Using our website, customs can decide whether they should leave more time space between their travel plans (because the weather forecasts are only valid within a week in future, so we can only predict the probability of flight delay in the future one week). By choosing the date and flight number of their flight, the web will return a probability of the flight delay. Also, we may give customs some advice to arrange their agenda and selection of airline.

**Time Schedule**

|  |  |
| --- | --- |
| Task | Date |
| Project description | 03/11/2016 |
| Database build up and analysis | 03/20/2016 |
| ER/UML Design | 04/01/2016 |
| Project SQL Statements and Initial Setup | 04/15/2016 |
| Website construction, test and evaluation | 04/20/2016 |
| Final Project Deliverable | 04/29/2016 |

**Scope of work**

* Project description
* Requirement collection
* Database resource searching
* Database server selection
* Database build up and analysis
* Data import
* Database resource searching
* Database server selection
* ER/UML Design
* Entity, attribute types define
* Relationships and constraints define
* UML transformation
* Project SQL Statements and Initial Setup
* Database implementation using SQL statements
* Website construction, test and evaluation
* Database modification and reconstruction
* Website build up
* Result evaluation
* Final Project Deliverable

**Source Database**

[1] weather information database:

<https://www.wunderground.com/history/airport/KSTL/2015/7/28/MonthlyHistory.html?req_city=&req_state=&req_statename=&reqdb.zip=&reqdb.magic=&reqdb.wmo>

[2]. previous flight delay information database:

<http://apps.bts.gov/xml/ontimesummarystatistics/src/dstat/OntimeSummaryArrivalsData.xml>

[3] calendar database:

http://nationalholidaydates.com/UnitedStates/2105.aspx