Assignment 1: Basics and Map Reduce

Group 1 Jin Yin a1761810. Yingyao Lu a1784870

Exercise1, Suspected Pairs (10 points)

1, The number of days of observation is 5000; and

2, The number of people observed was raised to 5 billion (and there were therefore 500, 000 hotels); and

3, We only reported a pair as suspect if they were at the same hotel at the same time on four different days.

In this case, the chance that they will visit the same hotel on one given day is:

Four days the, the possibility is

The number of pairs in the people is

The number of pairs of days is

So, the amount of the people is

Exercise 2, Hadoop (20 points)

Standalone mode:

|  |
| --- |
|  |

Exercise 3, Friend Recommendation System (Stanford) (35 points)

Exercise 4, MapReduce (15 points)

Exercise 5, Summary of 2.4 and 2.5 (10 +10 points) (Postgraduate Students (COMP SCI 7306) only)