Parallel & Distributed Computing
CSE525

Assignment **#8** - to be submitted to **Dr. Masroor Hussain**

Map/Reduce using Hadoop

Submitted by,
Quswar Mahmood Abid, CS2003

Matrix-Matrix Multiplication using Map-Reduce (Hadoop)

Code sourced from Khushali Dave's repo on matrix multiplication using Hadoop. Code also available at a forked library [https://github.com/quswarabid/Matrix-Multiplication-Hadoop-Map-Reduce] or with the files attached. The input matrices in these files are of sizes 4x3 and 3x5, respectively. We provided these matrices in flattened format as in txt files as we cannot provide data as 2D Array. So the first element in Matrix A which refers to 1st row and 1st column which is 5 as follows : a,0,0,5 where "a" represents matrix, "0" represents row number, "0" represents column number which contains element "5". To run these files, execute:

```
hadoop jar /usr/lib/hadoop-mapreduce/hadoop-streaming.jar

        -input /user/cloudera/input

        -output /user/cloudera/MatrixMulOutput

        -mapper /home/cloudera/Matrix/Matrix_Mapper.py

        -reducer /home/cloudera/Matrix/Matrix_Reducer.py
```

The output of the program is a list of matrix elements in style of key-value pairs:

```
(0,0) 30

(0,1) 51

(0,2) -5

(0,3) 15

(0,4) 14

(1,0) 15

(1,1) -12

(1,2) -25

(1,3) 12

(1,4) 28

(2,0) 50

(2,1) 65

(2,2) 5

(2,3) 33

(2,4) -26
```

(3,0) -5

(3,1) 2

(3,2) -3

(3,3) -6

(3,4) 16