

README

← How to run →

Import the source file into eclipse and export into jar.

Depending on whether you want to use Simple or Blocking, set the Simple/Block Map Reduce class as your main class.

The jars take in 3 inputs:

the path to the input file

the path of the output the files (notice that the output will add the iteration number at the end of this path)

and the number of passes

← Precomputations →

netid used = hhc39

rejectMin = 0.9207

rejectLimit = 0.9307

number of edges selected = 7524242

Our input file is in the source code directory named *hadoopInput.txt* found at

<http://sweet-mapreduce.s3.amazonaws.com/pagerank/hadoopInput.txt>

key\tabvalue

nodeID current_pagerank num_Outgoing_edges list_of_outNodes_if_exists

← Computations →

Simple PageRank:

- Mapper:
 - Parses the value and sends its rank to its outNodes or to the counter for no outNodes
- Reducer:
 - Adds its residual to the counter (multiple the number by 10E12)
 - Output the files in the exact same format as mapper receives them

Block PageRank:

- Mapper:

- Parses the value and sends the value to its node
- If it does have any out edges, add its pagerank to the counter for no out edges
- Calculates its outrank to each outNode
 - If its outNode is in the same block, send its nodeID to it.
 - Else send its outrank to that node's block.
- Reducer:
 - Process all edges to create various structures and pageranks.
 - Iterate at most 10 times or until the block residual is less than 0.01. (Keep the iteration average)
 - Find the biggest nodeID in each block and pass it to its corresponding counter

← Simple PageRank Residuals →

Iteration 0: 2.3386718

Iteration 1: 0.322979956

Iteration 2: 0.1918914

Iteration 3: 0.0940770

Iteration 4: 0.06284983

← Block PageRank Results →

For 20 maximum iterations:

Iteration 0 -> avg residual 2.816369003250558

Has average block passes of 20.0

Iteration 1 -> avg residual 0.03952213406068036

Has average block passes of 19.455882352941178

Iteration 2 -> avg residual 0.02515115893901318

Has average block passes of 14.411764705882353

Iteration 3 -> avg residual 0.011081515876333494

Has average block passes of 10.779411764705882

Iteration 4 -> avg residual 0.004767019523342527

Has average block passes of 8.088235294117647

Iteration 5 -> avg residual 0.0012005201820118792

Has average block passes of 4.352941176470588

Iteration 6 -> avg residual 6.356268238109831E-4

Has average block passes of 2.676470588235294

And these are the node pagerank values below for the biggest nodes in a block:

10327 1.894608508720783E-6

20372 5.144655138415415E-7

30628 3.0659370092288054E-7

40644 2.809275717642252E-7

50461 2.998581431986872E-7

60840 2.1890460137472094E-7
70590 3.0039254814045086E-7
80117 2.1890460137472094E-7
90496 8.678147372711481E-4
100500 2.6004293165929734E-7
110566 2.066903634883556E-6
120944 4.207894575473791E-7
130998 2.3475209735695868E-7
140573 6.24627634596471E-7
150952 2.1890460137472094E-7
161331 2.1890460137472094E-7
171153 5.622050208636744E-7
181513 3.529455517939439E-7
191624 5.027447400911657E-6
202003 4.101231620853567E-6
212382 2.6241452671835857E-7
222761 0.0012296187423283694
232592 1.1623395590934494E-4
242877 5.469335976991001E-7
252937 2.1890460137472094E-7
263148 3.0489269801220846E-7
273209 2.1890460137472094E-7
283472 3.3682959383192096E-7
293254 2.624144658753177E-7
303042 5.033521932074791E-6
313369 8.719168394302445E-7
323521 2.1890460137472094E-7
333882 5.609964678674138E-7
343662 2.5594038513883153E-7
353644 3.595100817283461E-7
363928 3.5564578445364246E-7
374235 3.840988942466275E-6
384553 4.123516215246791E-7
394928 6.704061996436505E-7
404711 2.471322840888531E-7
414616 3.116695738274327E-7
424746 1.3800460168039234E-6
434706 2.8142783774753487E-6
444488 3.8057415526741863E-6
454284 4.975162372075287E-7
464397 1.1191790566942556E-5
474195 9.869203662571346E-7
484049 5.827251748805271E-7

493967 2.1890460137472094E-7
503751 6.916575304080437E-6
514130 6.165920506325054E-7
524509 9.884126202737804E-4
534708 2.57305188469195E-5
545087 0.0026419954798941147
555466 0.0011913855931278384
565845 1.634285145604442E-6
576224 1.1053308062662258E-6
586603 9.740401207655695E-7
596584 4.368801127608822E-6
606366 4.267736144364642E-7
616147 6.12217506194149E-7
626447 1.4459894301271158E-5
636239 1.0720388107813924E-6
646021 3.1193905695897734E-7
655803 2.1890460137472094E-7
665665 9.976892307945333E-7
675447 1.1023447496272205E-6
685229 3.5859316143098615E-7