Continuation of Project 2

Grant Dawson | Ryan Rosiak | Project 3

Base Model Program (From Project 2)

Page Rank:

- Semi-working code
- Inefficient concepts
- Scatterbrain
- Too close to topic
- Not integrated to take input from files

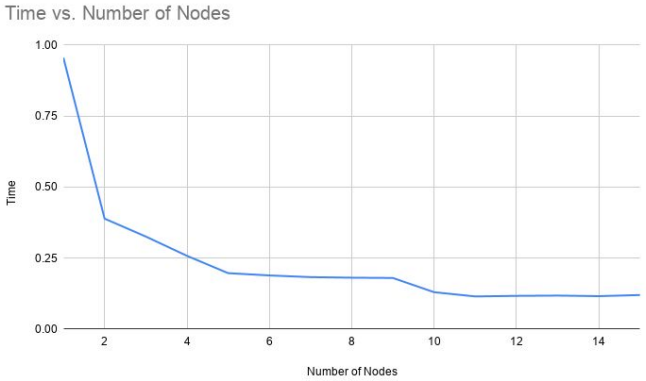
Files:

- Complete Meta-Data
  - Structs of metadata for papers
- Complete Index
  - Structs of paper ID and given unique index given by us
- Garbage Adjacency Matrix
- Binary Tree class

Upgraded Project 2/Project 3

Page Rank:

- Fully paralized code
- Takes input from file
- Equipt to take sparse matrices



Files:

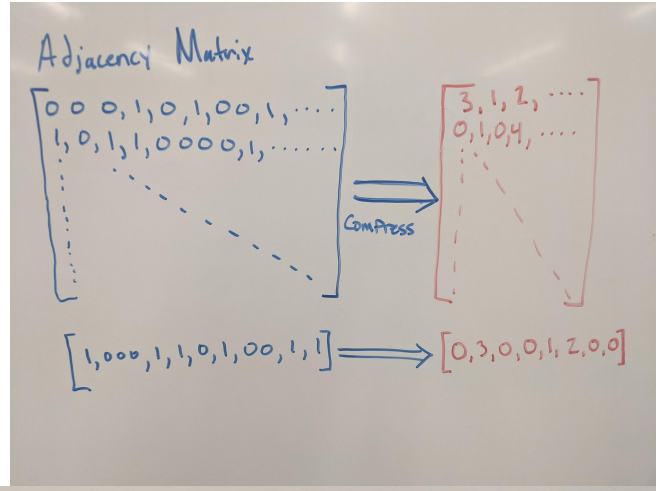
- Sparse Adjacency Matrix
- PageRank File
  - Structs of IDs and papers PageRank
- User Interface

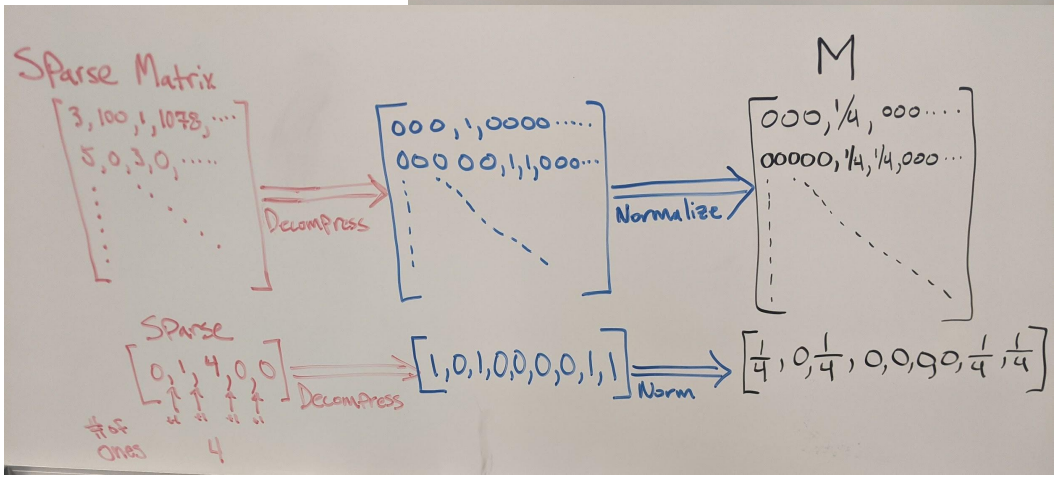
Issues Along the Way

- Memory space
  - Solution: Only load in what we need as we need it. Put load on hard drive
- Then hard drive space
  - Solution: Double -> short -> char -> Bit masking -> Sparse Matrices
    - 30T -> 8T -> 3T -> 43GB -> 35GB
- Change code
  - We have this new solution but it is now using a new protocol of Sparse Matrices
- Unclean data
  - Missing papers from citation file
- Main module not complete
  - The tree's are implemented but the structure is not fully fleshed out. There are certain problems with reading in files in parallel that are causing issues. (MPI and structs!)

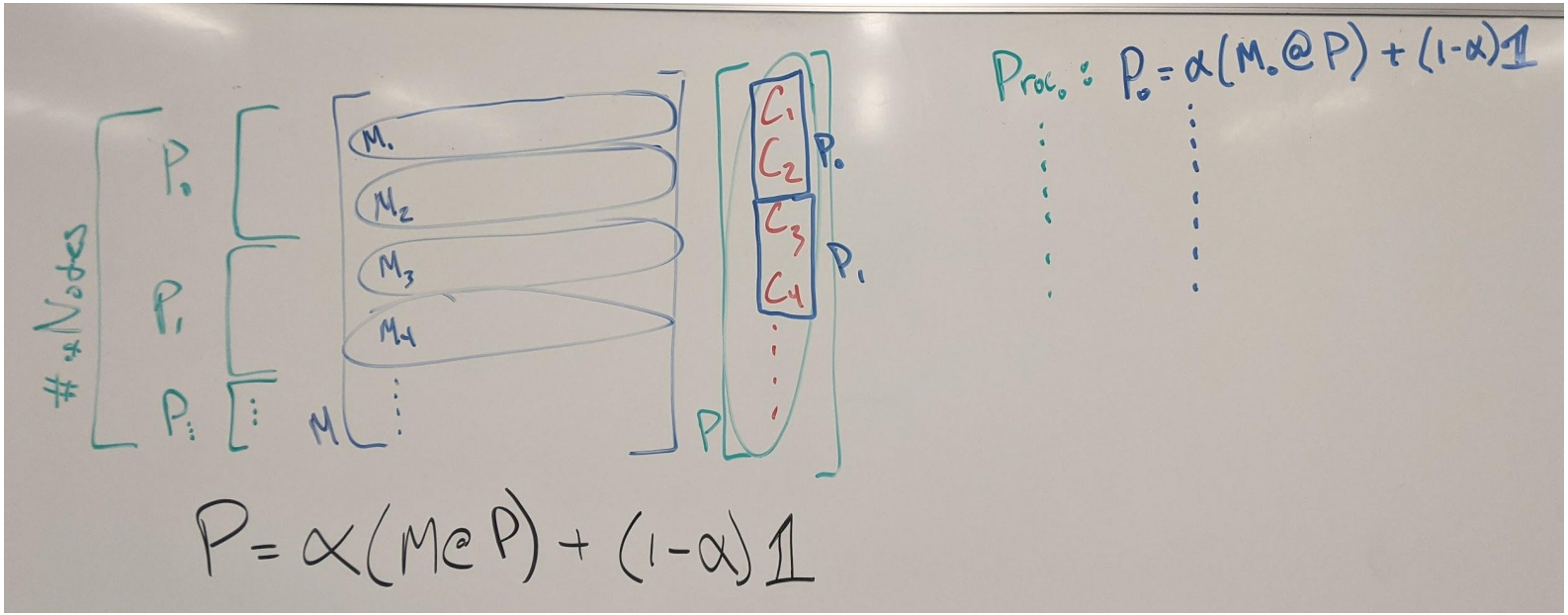
Sparse Matrices

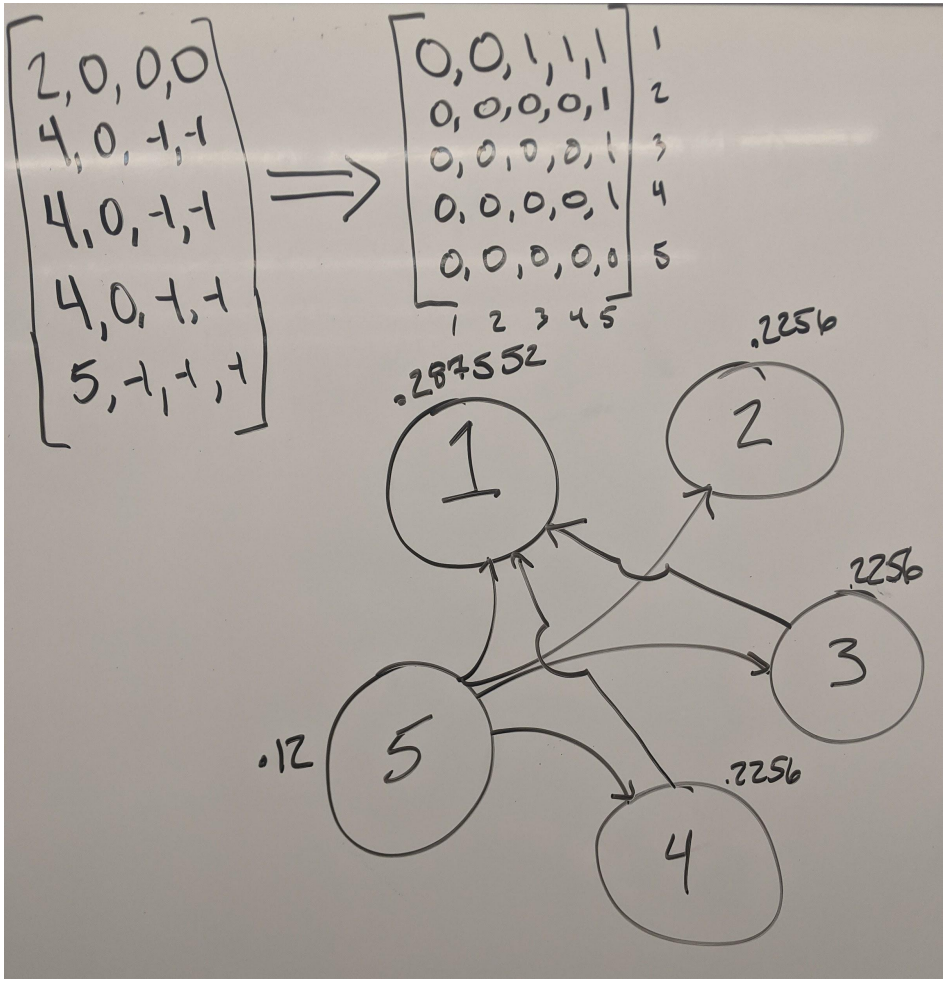
- Only shows benefits for sparse matrices hence the name, Sparse Matrix
- The 1's are implied
- Numbers represent number of 0's





PageRank Concepts




2,0,0,0
4,0,-1,-1
4,0,-1,-1
4,0,-1,-1
5,-1,-1,-1
0,0,1,1,1
0,0,0,0,1
0,0,0,0,1
0,0,0,0,1
0,0,0,0,0
1 2 3 4 5
.287552
.2256
.2256
.12
.2256
1
2
3
4
5

Where to next?

- We plan to finish this code. We plan to have a fully functioning program.
- Make the code travel size
  - This involves us making the code both more portable and condensed. This allows for an easier user interface, changes in data it would read in so one day it would be able to work on bigger and smaller data sets with raw data files with the same protocols.
- Refining UI
  - No user input is allowed (yet)

Any Questions?