COMP 4601A
Fall 2023 - Lab #5

Objectives

The goal for this lab is to calculate the PageRank values for the fruit pages example.

Demonstrating/Submitting

There will be two ways to receive credit for completed labs, outlined below:
1. Attend an in-person lab or office hours and demonstrate your completed lab before the deadline. You will have to show that the goals of the lab have been completed and answer some questions about the lab and your code (see the lab reflection questions for some examples). Your grade will depend on the level of completion, as well as the quality of your design and answers. Only one partner is required for demonstration, though all partners are encouraged to take part. **If you demonstrate your lab this way, you don't need to submit anything on Brightspace.**
2. Record a video demonstration that is <10 minutes long. Ensure that your discussion in the video makes it clear that you have understood the content that the lab covers and that you demonstrate all the required functionality. Submit a ZIP containing a copy of your code (don't include database files, etc.), your answers to the lab reflection questions, and a copy of your demonstration video (either link to a public URL in your README or include the video file directly) to Brightspace. **If you are working with a partner, only one of you should make a submission. Include the names of all group members in the README file.**

Lab Description

For this lab, you must use the network structure data you have produced from crawling the fruit website example (root for crawling: https://people.scs.carleton.ca/~davidmckenney/fruitgraph) to calculate PageRank values for each page within that example. To receive credit for the lab, you must show that you have calculated the PageRank values correctly using the approach outlined in the PageRank lecture (i.e., power iteration). You will also be required to explain the approach and code used in the calculation. You do not need a web-based frontend for this lab and can simply print out the PageRank values after calculation. The expected PageRank values for the 25 highest ranked pages are listed below. These results were calculated using an alpha value of 0.1 and stopping iteration after the Euclidean distance between the PageRank vectors at t_x and t_{x-1} was less than 0.0001.

Expected PageRank Values:

#1. (0.0209949245) https://people.scs.carleton.ca/~davidmckenney/fruitgraph/N-1.html
#2. (0.0156983765) https://people.scs.carleton.ca/~davidmckenney/fruitgraph/N-4.html
#3. (0.0133858899) https://people.scs.carleton.ca/~davidmckenney/fruitgraph/N-3.html
#4. (0.0104631400) https://people.scs.carleton.ca/~davidmckenney/fruitgraph/N-0.html
#5. (0.0088755402) https://people.scs.carleton.ca/~davidmckenney/fruitgraph/N-9.html
#6. (0.0083186698) https://people.scs.carleton.ca/~davidmckenney/fruitgraph/N-12.html
#7. (0.0080328201) https://people.scs.carleton.ca/~davidmckenney/fruitgraph/N-15.html
#8. (0.0078924853) https://people.scs.carleton.ca/~davidmckenney/fruitgraph/N-20.html
#9. (0.0075545652) https://people.scs.carleton.ca/~davidmckenney/fruitgraph/N-6.html
#10. (0.006761578) https://people.scs.carleton.ca/~davidmckenney/fruitgraph/N-42.html
#11. (0.006161805) https://people.scs.carleton.ca/~davidmckenney/fruitgraph/N-10.html
#12. (0.005498782) https://people.scs.carleton.ca/~davidmckenney/fruitgraph/N-5.html
#13. (0.005452235) https://people.scs.carleton.ca/~davidmckenney/fruitgraph/N-21.html
#14. (0.004497487) https://people.scs.carleton.ca/~davidmckenney/fruitgraph/N-11.html
#15. (0.004464576) https://people.scs.carleton.ca/~davidmckenney/fruitgraph/N-62.html
#16. (0.004253380) https://people.scs.carleton.ca/~davidmckenney/fruitgraph/N-67.html
#17. (0.004245869) https://people.scs.carleton.ca/~davidmckenney/fruitgraph/N-13.html
#18. (0.004159161) https://people.scs.carleton.ca/~davidmckenney/fruitgraph/N-96.html
#19. (0.004094785) https://people.scs.carleton.ca/~davidmckenney/fruitgraph/N-33.html
#20. (0.004025139) https://people.scs.carleton.ca/~davidmckenney/fruitgraph/N-365.html
#21. (0.003878140) https://people.scs.carleton.ca/~davidmckenney/fruitgraph/N-7.html
#22. (0.003736079) https://people.scs.carleton.ca/~davidmckenney/fruitgraph/N-14.html
#23. (0.003722789) https://people.scs.carleton.ca/~davidmckenney/fruitgraph/N-175.html
#24. (0.003605079) https://people.scs.carleton.ca/~davidmckenney/fruitgraph/N-32.html
#25. (0.003478277) https://people.scs.carleton.ca/~davidmckenney/fruitgraph/N-34.html

Lab Reflection Questions

1. How did you generate the adjacency matrix from the crawled data?
2. What effect would lower/higher alpha values have on the results?