Tutorial Complex Systems

Freie Universität Berlin, Summer 2020
Martin Vingron · Piotr Sliwa

Assignment 4
Due date: 22.05.2020 10:00 AM before the lecture

You should upload a single PDF report (with all the important steps of your calculations), code files (with comments and clear function naming) and any additional files requested in the assignment on Whiteboard (KVV). Optimally you should generate your PDF report in Latex, we will also accept *legible* photos/scans grouped into a single PDF file. Form groups of 2 students to solve the problems but remember: everyone has to be able to explain the solutions.

Problem 1 (*70 Points; PageRank and Random Walks with Restart*)**.** Implement (using Python) the models introduced in the lecture. Use numpy / pandas and where appropriate the package NetworkX. When you are asked to implement, do not use a ready algorithm (e.g. PageRank from NetworkX). Use $\alpha = 0.85$.

(A) Implement the PageRank algorithm with the power iteration.

(B) Implement the Random Walk with Restart process. Implement both, the direct solution and the iterative solution.

(C) Create the following 5 random networks (remember to use a seed=42):

Erdős-Rényi (for $n = 100$, $p = \{0.01, 0.08, 0.4\}$),

Watts-Strogatz (for $n = 50$, $k = 7$, $p = 0.3$)

Barabási-Albert (for $n = 50$ and $m = 3$).

(D) Calculate the PageRank (using your above implementation) for all these networks using uniform distribution as your initial p. Plot the convergence (difference between k-th and $(k-1)$-st iterations) in the above PageRank calculations. How many iterations did you need to converge?

(E) Try different initial distributions. Does it change the end result?

(F) Plot the networks. Let the color (viridis colormap) of the node represent the PageRank value. Plot the colorbar next to the figure. Print the PageRank values.

(G) In the Barabási-Albert network randomly assign the probabilities $\{0.4, 0.1, 0.5\}$ to 3 nodes (the rest should have a 0 assigned) and propagate these scores in the network using the iterative approach in RWR. Plot how the PageRank value changes (X-axis - iteration, Y-axis - PageRank value) for all nodes (overlay the 50 lines in one figure).

(H) In which iterations can you observe the largest change in the scores? Create 4 plots - each in a different iteration to illustrate this change (the propagation).

(I) Repeat the above two steps for the Watts-Strogatz network.

(J) Now calculate the propagation using the direct solution. Is it the same as the converged iterative solution?

Problem 2 (*30 Points; Proofs and derivations*)**.** Go back to the lecture slides and:

(A) Prove that the convex combination of stochastic matrices is also stochastic.

(B) Derive the direct solution to the random walk with restart process.