The University of Sydney
School of Computer Science

THE UNIVERSITY OF SYDNEY

COMP5313/COMP4313—Large Scale Networks S1 2025
Week 6 - PageRank

The goal of this tutorial is to understand the definition and computation of PageRank.

Exercise 1: Computing the Basic PageRank Values

Compute the (equilibrium) basic PageRank values for nodes in the graph in Figure 1(a). After computing the PageRank values, verify that they are indeed equilibrium values. (Duration: 10 min)

What happens if we remove the edge from b to a? See the graph in Figure 1(b). (Duration: 5 min)

(a) (b)

Figure 1: Example graphs

Answer: For the graph in Figure 1(a), let r_a, r_b, r_c be the PageRank values of nodes a, b, c, respectively. Then, we have the following equations.

$$\begin{aligned} r_a &= \frac{r_a}{3} + \frac{r_b}{2} \\ r_b &= \frac{r_a}{3} + \frac{r_c}{2} \\ r_c &= \frac{r_a}{3} + \frac{r_b}{2} + \frac{r_c}{2} \\ 1 &= r_a + r_b + r_c \end{aligned}$$

Solving the system of linear equations, we get $r_a = \frac{3}{13}, r_b = \frac{4}{13}, r_c = \frac{6}{13}$.

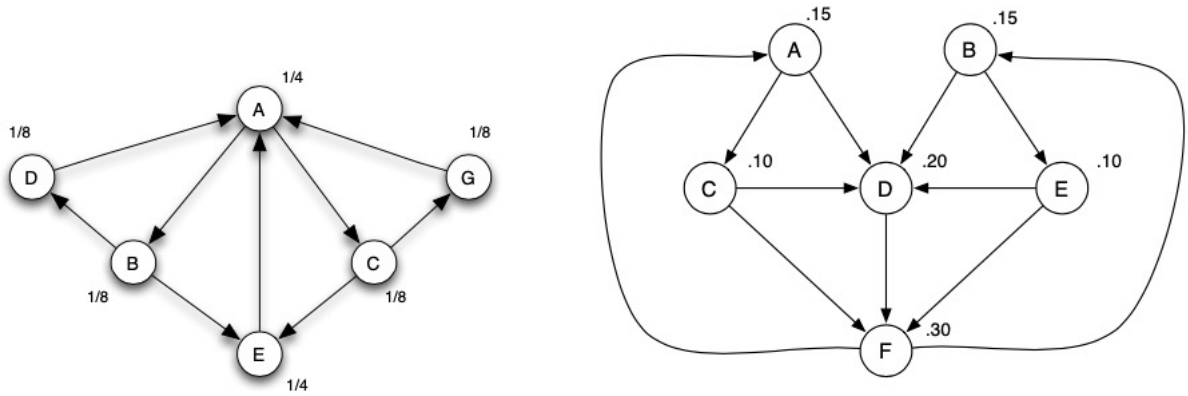
For the graph in Figure 1(b), the system of linear equations becomes:

$$
\begin{aligned}
r_a &= \frac{r_a}{3} \\
r_b &= \frac{r_a}{3} + \frac{r_c}{2} \\
r_c &= \frac{r_a}{3} + r_b + \frac{r_c}{2} \\
1 &= r_a + r_b + r_c
\end{aligned}
$$

Solving it obtains $r_a = 0, r_b = \frac{1}{3}, r_c = \frac{2}{3}$. Note that this graph is not strongly connected.

Exercise 2: Equilibrium PageRank Values

Consider the two graphs in Figure 2, where the proposed basic PageRank values are expressed as numbers next to the nodes. Are these correct equilibrium values? If not, what are the correct basic PageRank values?
For the graph in Figure 2(a). (Duration: 10 min)
For the graph in Figure 2(b). (Duration: 10 min)



(a) (b)

Figure 2: Example graphs

Answer: Let's first consider the graph in Figure 2(a). These are wrong equilibrium values. For example, after applying the basic PageRank update rule for one step, we would get the new PageRank value of node A to be $\frac{1}{2}$ which is not equal to $\frac{1}{4}$. To obtain the basic PageRank values, we let the PageRank value of node A be x, then the PageRank value of nodes B and C would be $\frac{x}{2}$ each as both B and C have only one incoming link (that is from A). Then, the PageRank values of nodes D, G, E are $\frac{x}{4}, \frac{x}{4}, \frac{x}{2}$, respectively. As the sum of the PageRank values of all nodes is 1, we get $x = \frac{1}{3}$ and the PageRank values of all nodes. Note that, you can also list the system of linear equations and then solve, as done in **Exercise 1**; but it is more tedious.

Now, let's consider the graph in Figure 2(b). These are wrong equilibrium values. For example, after applying the basic PageRank update rule for one step, we would get the new PageRank value of node C to be 0.075 which is not equal to 0.1. To obtain the basic PageRank values, we let the PageRank value of node F be x (Note that, letting the PageRank value of node A be x will not work for this

graph). Then, the PageRank value for nodes A and B would be $\frac{x}{2}$ and $\frac{x}{2}$, respectively. Then, the PageRank value for nodes C and E are $\frac{x}{4}$ and $\frac{x}{4}$, respectively. After that, the PageRank value of nodes D is $\frac{3x}{4}$. As the sum of the PageRank values of all nodes is 1, we get $x = \frac{4}{13}$ and the PageRank values of all nodes.

In general, if there is a node X such that (1) X can reach all other nodes and (2) the remaining graph has no cycle after removing X, then we can start the PageRank derivation from node X by letting its PageRank be x and then solving x. The derivation order should follow the topological ordering after removing node X; please refer to `https://en.wikipedia.org/wiki/Topological_sorting` for the definition of topological ordering. Note that, the graph in Figure 1(a) does not satisfy this condition, and thus cannot be solved by using this method.

Exercise 3: Scaled PageRank Values

Consider the graph in Figure 1(b) again. What are the scaled PageRank values for $\alpha = \frac{4}{5}$?
(Duration: 15 min)

Answer: For the graph in Figure 1(b), by using the scaled PageRank update rule, we have $\mathbf{r} = \alpha\mathbf{r}\mathbf{W} + \frac{1-\alpha}{n}\mathbf{e}$ (see Lecture slides for this equation, and recall that the equation for the basic PageRank update rule is $\mathbf{r} = \mathbf{r}\mathbf{W}$), and thus

$$
\begin{aligned}
r_a &= \frac{4}{5} \times \left(\frac{r_a}{3}\right) + \frac{1}{15} \\
r_b &= \frac{4}{5} \times \left(\frac{r_a}{3} + \frac{r_c}{2}\right) + \frac{1}{15} \\
r_c &= \frac{4}{5} \times \left(\frac{r_a}{3} + r_b + \frac{r_c}{2}\right) + \frac{1}{15} \\
1 &= r_a + r_b + r_c
\end{aligned}
$$

Solving it obtains $r_a = \frac{1}{11}, r_b = \frac{25}{77}, r_c = \frac{45}{77}$.