

HW 4 Recitation Problems

Exercise 1.1 (Leskovec)**Exercise 3.1.1**

$$J(\{1, 2, 3, 4\}, \{2, 3, 5, 7\}) = |\{2, 3\}| / |\{1, 2, 3, 4, 5, 7\}| = (2 / 6)$$

$$J(\{1, 2, 3, 4\}, \{2, 4, 6\}) = |\{2, 4\}| / |\{1, 2, 3, 4, 6\}| = (2 / 5)$$

$$J(\{2, 3, 5, 7\}, \{2, 4, 6\}) = |\{2\}| / |\{2, 3, 4, 5, 6, 7\}| = (1 / 6)$$

Exercise 3.2.1

Sentence: “The most effective way to represent documents as sets, for the purpose of identifying lexically similar documents is to construct from the document the set of short strings that appear within it.”

Shingles: {“The”, “he “, “e m”, “ mo”, “mos”, “ost”, “st “, “t e”, “ ef”, “eff”}

Exercise 3.3.3

a)

Element	S1	S2	S3	S4	$2x+1 \% 6$	$3x+2 \% 6$	$5x+2 \% 6$
0	0	1	0	1	1	2	2
1	0	1	0	0	3	5	1
2	1	0	0	1	5	2	0
3	0	0	1	0	1	5	5
4	0	0	1	1	3	2	4
5	1	0	0	0	5	5	3

Hash	S1	S2	S3	S4
$2x+1 \% 6$	5	1	1	1
$3x+2 \% 6$	2	2	2	2
$5x+2 \% 6$	0	1	2	0

b) $5x+2 \pmod 6$ is a outputs a true permutation

c) Column-Column = Below Diagonal || Signature-Signature = Above Diagonal

I don't think the numbers are very close but the relationship of the number is pretty good.

The 0/4 correspond to the 1/3 and the 1/4 correspond to the 2/3.

J. Similarity	S1	S2	S3	S4
S1		1/3	1/3	2/3
S2	0/4		2/3	2/3
S3	0/4	0/4		2/3
S4	1/4	1/4	1/4	

Exercise 3.4.1

S Curse $(1 - (1-s^r)^b)$	$r = 3$ and $b = 10$	$r = 6$ and $b = 20$
0.1	0.010	0.000
0.2	0.077	0.001
0.3	0.239	0.014
0.4	0.484	0.079
0.5	0.737	0.270
0.6	0.912	0.615
0.7	0.985	0.918
0.8	0.999	0.998
0.9	1.000	1.000

Exercise 3.5.4

a) $1 - (3/5) = 2/5$

b) $1 - (0/6) = 6/6$

Exercise 3.5.5

a) $-7 / 14$

- b) Scalar multiple, angle = 0 => value is 1
- c) $-13 / 41$
- d) $1 / 3$

Exercise 1.2 (Leskovec)

Exercise 9.2.1

- a) $A = \{3.06(1), 500\alpha, 6\beta\}$
 $B = \{2.68(1), 320\alpha, 4\beta\}$
 $C = \{2.92(1), 640\alpha, 6\beta\}$
 $A \text{ and } B = (8.20 + 160,000\alpha^2 + 24\beta^2) /$
 $(\sqrt{9.36 + 250,000\alpha^2 + 36\beta^2}) * \sqrt{7.18 + 102,400\alpha^2 + 16\beta^2})$
 $A \text{ and } C = (8.94 + 320,000\alpha^2 + 36\beta^2) /$
 $(\sqrt{9.36 + 250,000\alpha^2 + 36\beta^2}) * \sqrt{8.53 + 409,600\alpha^2 + 36\beta^2})$
 $B \text{ and } C = (7.83 + 204,800\alpha^2 + 24\beta^2) /$
 $(\sqrt{7.18 + 102,400\alpha^2 + 16\beta^2}) * \sqrt{8.53 + 409,600\alpha^2 + 36\beta^2})$
- b) $A \text{ and } B = 160,032.2 / (500.05 * 320.04) = (160,032.2 / 160,036) \Rightarrow \Theta = 0.00707 \text{ deg}$
 $A \text{ and } C = 320,044.94 / (500.05 * 640.03) = (320,044.94 / 320,047) \Rightarrow \Theta = 0.00434 \text{ deg}$
 $B \text{ and } C = 204,831.83 / (320.04 * 640.03) = (204,831.83 / 204,835.2) \Rightarrow \Theta = 0.00574 \text{ deg}$
- c) $A \text{ and } B = 30.2 / (6.58 * 4.63) = (30.2 / 30.47) \Rightarrow \Theta = 0.136 \text{ deg}$
 $A \text{ and } C = 49.94 / (6.58 * 7.65) = (49.94 / 50.34) \Rightarrow \Theta = 0.126 \text{ deg}$
 $B \text{ and } C = 34.31 / (4.63 * 7.65) = (34.31 / 35.42) \Rightarrow \Theta = 0.250 \text{ deg}$

Exercise 9.2.3

- a) Average rating = $11/3 \Rightarrow 3.67$ ||| Normalized Rating = $\{A: 0.33, B: -1.67, C: 1.33\}$
- b) $\{2.92, 640, 6\}$

Exercise 9.3.1

- a) $J(A, B) = 4 / 8$
 $J(A, C) = 4 / 8$
 $J(A, C) = 4 / 8$

b) A and B = 4 / 6

A and C = 4 / 6

B and C = 4 / 6

c) $J(A, B) = 2 / 5$

$J(A, C) = 2 / 6$

$J(B, C) = 1 / 5$

d) A and B = $2 / (2 * \sqrt{3})$

A and C = 2 / 4

B and C = $1 / (\sqrt{3} * 2)$

e)

	AVG	a	b	c	d	e	f	g	h
A	3.33	0.67	1.67		1.67	-2.33		-0.33	-1.33
B	2.33		0.67	1.67	0.67	-1.33	-0.33	-1.33	
C	3	-1		-2	0		1	2	0

f) A and B = $5.78 / \sqrt{13.34} + \sqrt{5} = 5.78 / (3.65 * 2.24) = 0.71$

A and C = $-1.33 / \sqrt{13.34} + \sqrt{10} = -1.33 / (3.65 * 3.16) = -0.12$

B and C = $-6.33 / \sqrt{5} * \sqrt{10} = -6.33 / (2.24 * 3.16) = -0.89$

Exercise 1.3 (Leskovec)

Exercise 5.1.1

Node	a	b	c
a	1/3	1/2	0
b	1/3	0	1/2
c	1/3	1/2	1/2

a = 0.231, b = 0.307, c = .462 (from the power code)

Exercise 5.1.2

0.8	a	b	c
a	1/3	1/2	0
b	1/3	0	1/2
c	1/3	1/2	1/2

0.2			
	1/3	1/3	1/3
	1/3	1/3	1/3
	1/3	1/3	1/3

Taxed	a	b	c
a	5/15	7/15	1/15
b	5/15	1/15	7/15
c	5/15	7/15	7/15

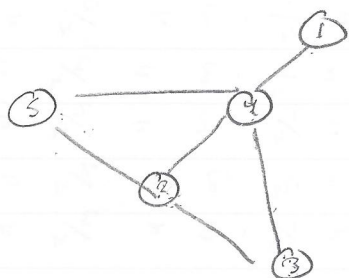
$a = 0.259$, $b = 0.309$, $c = 0.432$ (from the power iteration code)

Exercise 5.1.6

Each node (except the tail node and the head node) have one input and one output. The tail has only one input and the self-looped head has two outs and one in. All the nodes after the head (including the tail) would be half the page rank of the head - This is because they receive the same page rank as the nodes before it (one input from a node with one output). They all have a page rank of half of the head because the head goes down the list one half of the time and goes to itself the other half.

Exercise 1.4 (Centrality Measures)

a)



$$\begin{aligned} a) \quad C_P^*(1) &= \frac{1}{4} \cdot 1 = \frac{1}{4} = 0.25 \\ C_D^*(2) &= \frac{1}{4} \cdot 3 = \frac{3}{4} = 0.75 \\ C_D^*(3) &= \frac{1}{4} \cdot 2 = \frac{2}{4} = 0.5 \\ C_D^*(4) &= \frac{1}{4} \cdot 4 = \frac{4}{4} = 1.0 \\ C_D^*(5) &= \frac{1}{4} \cdot 2 = \frac{2}{4} = 0.5 \end{aligned}$$

$$\begin{aligned} b) \quad C_C^*(1) &= 4 \cdot (1 / (2+2+1+2)) = 4 \cdot \frac{1}{7} = \frac{4}{7} \\ C_C^*(2) &= 4 \cdot (1 / (2+1+1+1)) = 4 \cdot \frac{1}{5} = \frac{4}{5} \\ C_C^*(3) &= 4 \cdot (1 / (2+1+1+2)) = 4 \cdot \frac{1}{6} = \frac{4}{6} \\ C_C^*(4) &= 4 \cdot (1 / (1+1+1+1)) = 4 \cdot \frac{1}{4} = \frac{4}{4} \\ C_C^*(5) &= 4 \cdot (1 / (2+1+1+2)) = 4 \cdot \frac{1}{6} = \frac{4}{6} \end{aligned}$$

$$\begin{aligned} c) \quad 2-3 &= 0/1 & C_B(1) &= 2 \cdot 0 \\ 2-4 &= 0/1 & C_B^*(1) &= \frac{0}{12} \\ 2-5 &= 0/1 & \\ 3-4 &= 0/1 & \\ 3-5 &= 0/2 & \\ 4-5 &= 0/1 & \end{aligned}$$

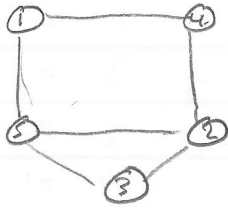
$$\begin{aligned} 1-3 &= 0/1 & C_B(2) &= 2 \cdot \frac{1}{2} \\ 1-4 &= 0/1 & C_B^*(2) &= \frac{1}{12} \\ 1-5 &= 0/1 & \\ 3-4 &= 0/1 & \\ 3-5 &= 1/2 & \\ 4-5 &= 0/1 & \end{aligned}$$

$$\begin{aligned} 1-2 &= 0/1 & C_B(3) &= 2 \cdot 0 \\ 1-4 &= 0/1 & C_B^*(3) &= \frac{0}{12} \\ 1-5 &= 0/1 & \\ 2-4 &= 0/1 & \\ 2-5 &= 0/1 & \\ 4-5 &= 0/1 & \end{aligned}$$

$$\begin{aligned} 1-2 &= 1/1 & C_B(4) &= 2 \cdot \frac{2}{2} \\ 1-3 &= 1/1 & C_B^*(4) &= \frac{7}{12} \\ 1-5 &= 1/1 & \\ 2-3 &= 0/1 & \\ 2-5 &= 0/1 & \\ 3-5 &= 1/2 & \end{aligned}$$

$$\begin{aligned} 1-2 &= 0/1 & C_B(5) &= 2 \cdot 0 \\ 1-3 &= 0/1 & C_B^*(5) &= \frac{0}{12} \\ 1-4 &= 0/1 & \\ 2-3 &= 0/1 & \\ 2-4 &= 0/1 & \\ 3-4 &= 0/1 & \end{aligned}$$

b)



$$\begin{aligned}
 a) \quad C_D^*(1) &= \frac{1}{4} \cdot 2 = \frac{2}{4} = 0.5 \\
 C_D^*(2) &= \frac{1}{4} \cdot 3 = \frac{3}{4} = 0.75 \\
 C_D^*(3) &= \frac{1}{4} \cdot 2 = \frac{2}{4} = 0.5 \\
 C_D^*(4) &= \frac{1}{4} \cdot 2 = \frac{2}{4} = 0.5 \\
 C_D^*(5) &= \frac{1}{4} \cdot 3 = \frac{3}{4} = 0.75
 \end{aligned}$$

$$\begin{aligned}
 b) \quad C_C^*(1) &= 4 \cdot (1 / (2 + 2 + 1 + 1)) = 4/6 \\
 C_C^*(2) &= 4 \cdot (1 / (2 + 1 + 1 + 1)) = 4/5 \\
 C_C^*(3) &= 4 \cdot (1 / (2 + 1 + 2 + 1)) = 4/6 \\
 C_C^*(4) &= 4 \cdot (1 / (1 + 1 + 2 + 2)) = 4/6 \\
 C_C^*(5) &= 4 \cdot (1 / (1 + 1 + 1 + 2)) = 4/5
 \end{aligned}$$

$$\begin{aligned}
 c) \quad 2-3 &= 0/1 & C_B(1) &= 2 \cdot \frac{1}{2} \\
 2-4 &= 0/1 & & \\
 2-5 &= 0/1 & C_B^*(1) &= \frac{1}{12} \\
 3-4 &= 0/1 & & \\
 3-5 &= 0/1 & & \\
 4-5 &= 1/2 & &
 \end{aligned}$$

$$\begin{aligned}
 1-3 &= 0/1 & C_B(2) &= 2 \cdot \frac{3}{2} \\
 1-4 &= 0/1 & & \\
 1-5 &= 0/1 & C_B^*(2) &= \frac{3}{12} \\
 3-4 &= 1/1 & & \\
 3-5 &= 0/1 & & \\
 4-5 &= 1/2 & &
 \end{aligned}$$

$$\begin{aligned}
 1-2 &= 0/2 & C_B(3) &= 2 \cdot 0 \\
 1-4 &= 0/1 & & \\
 1-5 &= 0/1 & C_B^*(3) &= \frac{0}{12} \\
 2-4 &= 0/1 & & \\
 2-5 &= 0/1 & & \\
 4-5 &= 0/2 & &
 \end{aligned}$$

$$\begin{aligned}
 1-2 &= 1/2 & C_B(4) &= 2 \cdot \frac{1}{2} \\
 1-3 &= 0/1 & & \\
 1-5 &= 0/1 & C_B^*(4) &= \frac{1}{12} \\
 2-3 &= 0/1 & & \\
 2-5 &= 0/1 & & \\
 3-5 &= 0/1 & &
 \end{aligned}$$

$$\begin{aligned}
 1-2 &= 1/2 & C_B(5) &= 2 \cdot \frac{3}{2} \\
 1-3 &= 1/1 & & \\
 1-4 &= 0/1 & C_B^*(5) &= \frac{3}{12} \\
 2-3 &= 0/1 & & \\
 2-4 &= 0/1 & & \\
 3-4 &= 0/1 & &
 \end{aligned}$$