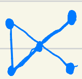
Note 11.22.

Metric on Graph:

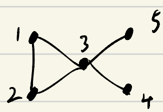
Diameter: $Diam(G) = \max_{ij} d_{ij}$ (Small World phenomenon)

↳ shortest distance between node i and node j.



=> $Diam(G) = 2$

Clustering Coefficient: $C = \#\text{ triangles} / \#\text{ triplets}$.



=> $C = \frac{(1+1+1+0+0)}{(1+1+6+0+0)} = \frac{3}{8}$

centered in 1   centered in 3   5

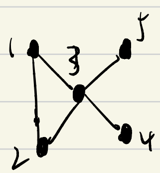
Density: $= 2M/N(N-1)$ ↳ Note: A complete graph has a density of 1.

M = Num of Edges N = Num of Nodes

Matrix on Nodes:

Degree Centrality: $C_{deg}(v) = Deg(v)$ The more central the node is, the higher its number of connections.

Closeness Centrality: $C_{close}(v) = \frac{1}{\sum_u d(u,v)}$ The more central the node is, the closer it is to other nodes



-> $C_{deg}(3) = 4$ (As Node 3 has 4 Nodes).

-> $C_{close}(3) = \frac{1}{4}$ (As Node 3 connects to 1, 2, 4, 5)

-> $C_{close}(5) = \frac{1}{2+2+1+2} = \frac{1}{7}$

$d(1,5)$ $d(2,5)$ ......

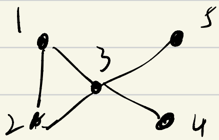
Harmonic Centrality : $C_h(v) = \sum_{u \neq v} \frac{1}{d(u,v)}$

$\frac{1}{d(u,v)} = 0$ if no path between u & v.

Betweenness Centrality : $C_b(v) = \frac{\sigma_{st}(v)}{\sigma_{st}}$

$\sigma_{st}(v)$ → shortest path go through v.

σ_{st} → total number of shortest path from s to t.



$C_h(3) = \frac{1}{1} + \frac{1}{1} + \frac{1}{1} + \frac{1}{1} = 4$

between centrality for Node 3 : 10

s	t	Pass Node 3 ?
1	2	Yes
1	3	No
1	4	Yes
1	5	No
2	3	
2	4	
2	5	
3	4	
3	5	
4	5	

10 Yes

Recommendation :

Ranking Aggregation :

Ranking : The ordered list of Node. (Ex. HW3).

Ex: If we have two rankings (ordered list) , w_1 , w_2 ,
How could we compare them ?

- Disagreement Distance.

Kendall τ distance: $d_\tau(w_1, w_2)$ = # pairs in different order

Ex:

Node	A	B	C	D
w_1	1	2	3	4
w_2	3	4	1	2

⇓

Pair	Agree or not
A B	✓
A C	×
A D	×
B C	×
B D	×
C D	✓

A B ✓ → A and B has same order for W_1 , W_2

$\therefore d_\tau(W_1, W_2) = \frac{4}{6}$

Thus, If we have $W_1, W_2, W_3 \cdots\cdots W_m$, we can generate an

Aggregate Ranking W: $W^* = \underset{W}{\arg\min} \sum_{i=1}^{m} d_\tau(W, W_i)$

$\Downarrow$

We want to minimize the sum of disagreement distance.