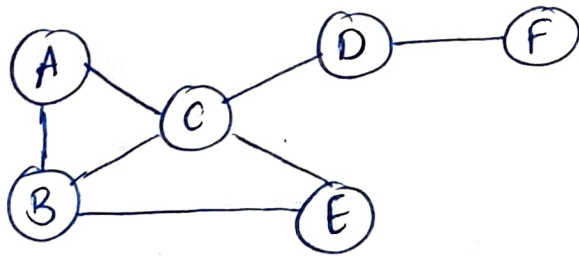


CIS600 principles of social media and Data mining

Homework 1

Adishree Sane
adsane@syr.edu

①



(a) degree of each node

A - 2

B - 3

C - 4

D - 2

E - 2

F - 1

- degree of nodes is number of edges connected to it.
- diameter is longest shortest path between any two nodes in network.
- average path length: mean shortest path length between all pairs of node.

(b) Diameter table

Node pair	shortest distance
A, B	1
A, C	1
A, D	2
A, E	2
A, F	3
B, C	1
B, E	1
B, D	2
B, F	3
C, D	1
C, E	1
C, F	2
D, E	2
D, F	1
E, F	3

hence, longest shortest path is
i.e diameter is 3

b) ② average path of network.

	A	B	C	D	E	F
A	0	1	1	2	2	3
B	1	0	1	2	1	3
C	1	1	0	1	1	2
D	2	2	1	0	2	1
E	2	1	1	2	0	3
F	3	3	2	1	3	0

average length of network
 $= \frac{\text{Sum of all paths}}{\text{no. of path}}$

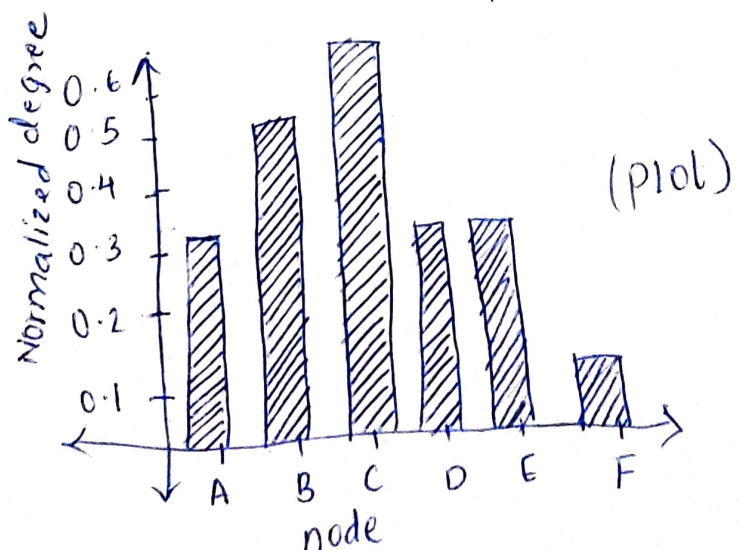
$$= \frac{54}{30} = \underline{\underline{1.8}}$$

a) ② Normalized degree

formula: Normalized degree = $\frac{\text{Degree of Node}}{\text{Total no. of nodes}}$

Total nodes = 6

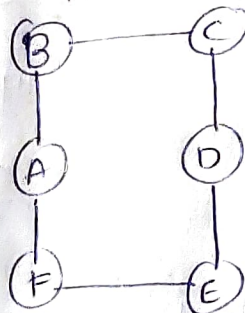
Normalised degrees	degree	element/node
$2/6 = 0.333$	2	A
$3/6 = 0.50$	3	B
$4/6 = 0.6667$	4	C
$2/6 = 0.333$	2	D
$2/6 = 0.333$	2	E
$1/6 = 0.1667$	1	F



(a) ABCDED - walk

(b) DEFA - path

(c) ABCDEFA - cycle

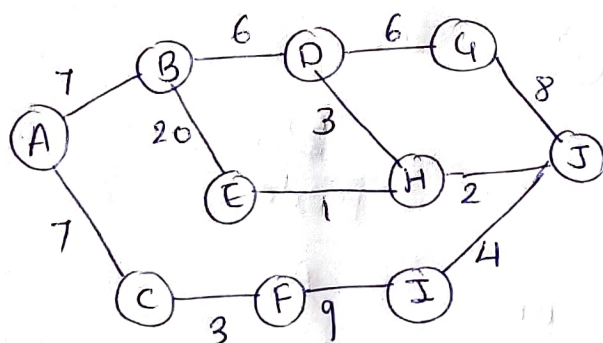


Walk: A sequence of vertices, where consecutive vertices are adjacent and vertices/edges may be repeated.

Path: A walk where no vertex is repeated.

Cycle: A path that starts and ends at the same vertex.

(3) Dijkstra's



(a) shortest path from A to J

initially every distance of node is ∞ .

node	explored
A = 0	-
B = 7, C = 7	A
C = 7, D = 13, E = 21	A, B
D = 13, E = 21, F = 10	A, B, C
D = 13, E = 21, I = 19	A, B, C, F
E = 21, I = 19, H = 16, G = 19	A, B, C, F, D
E = 21, I = 19, J = 17	A, B, C, F, D, H
E = 21, I = 19, J = 17	A, B, C, F, D, H, J

Shortest path from A to J is from parent and children.

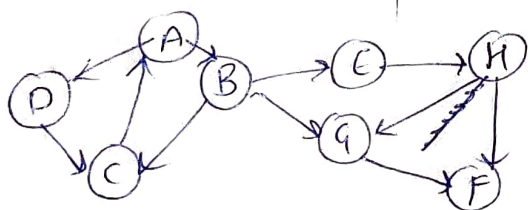
A-B-D-H-J by backtracking with cost (18)

⑤ Shortest path between B and E

node	explored.
B=0	-
D=6, E=20	B
E=20, G=12, H=9	B, D
E=20, G=12, H=9	B, D, H
E=20, G=12, H=9	B, D, H, J
E=20, G=12, H=9	<div>path is B-D-H-E(10)</div>

Path is
B-D-H-J

node	explored
B=0	-
D=6, E=20	B
E=20, H=9, G=12	B, D
E=20, H=9, G=12	B, D, H
E=20, H=9, G=12	B, D, H, E



⑥ BFS

node	Explored
A	-
D, B	A
B, C	A, D
C, E, G	A, D, B
E, G	A, D, B, C
G, H	A, D, B, C, E
H, F	A, D, B, C, E, G
F	A, D, B, C, E, G, H
-	A, D, B, C, E, G, H, F

or

BFS

node	Explored
A	-
B, D	A, B
D, C, E, G	A, B
C, E, G,	A, B, D
E, G,	A, B, D, C
G, H	A, B, D, C, E
H, F	A, B, D, C, E, G
F	A, B, D, C, E, G, H
-	A, B, D, C, E, G, H, F

I choose D first

Here I choose B first

path - A-B-D-C-E-G-H-F

④ ⑥ DFS

node

visited

A
~~A~~, B^A, D^A
D^A, C^B, E^B, G^B
D^A, E^B, G^B
D^A, G^B, H^E
D^A, G^B, F^H
D^A, G^B
D^A
—

—
A
A, B
A, B, C
A, B, C, E
A, B, C, E, H
A, B, C, E, H, F
A, B, C, E, H, F, G
A, B, C, E, H, F, G, D

path is

A - B - C - E - H - F - G - D

If we consider the graph as undirected.

⑤ • 2-clique:

number of 2-cliques = 11

(A, B), (A, D), (B, C), (B, E), (B, G), (D, E), (E, H), (G, F)

• 3-clique

number of 3-cliques = 3

(A, B, D)

(B, E, D)

• 4-clique = 2

For directed all cliques would be 0.