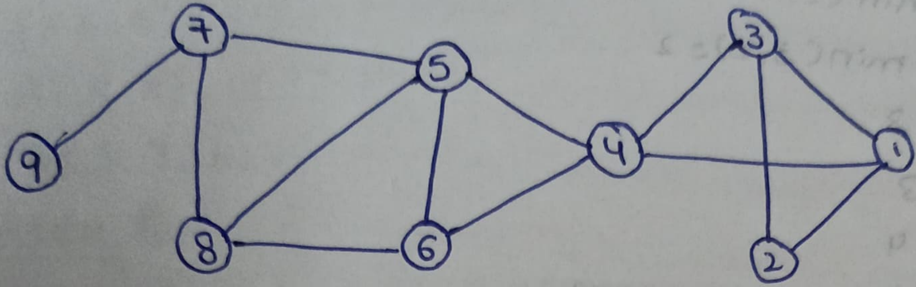
AP20110010559

Vyshnavi.Yakkanti

CSE-H

CSE 327 Assignment

1.



a] Users : 9

1, 2, 3, 4, 5, 6, 7, 8, 9

Edges : 14

[1 2] [1 3] [1 4]

[3 2] [3 4]

[4 5] [4 6]

[5 6] [5 7] [5 8]

[6 7] [6 8]

[7 8] [7 9]

b] 1st, lets find the distance for individual

D = 1, {2, 3, 4, 5, 6, 7, 8, 9}

d(1, 2) = 1

d(1, 3) = 1

d(1, 4) = min ((1, 3, 4), (1, 4)) = min (2, 1) = 1

d(1, 5) = 2 , d(1, 6) = 2 , d(1, 7) = 3, d(1, 8) = 3

d(1, 9) = 4

Max of these distances is eccentricity of vertex

$\boxed{e(1) = 4}$

for node 2

D = 2 , {1, 3, 4, 5, 6, 7, 8, 9}

d(2, 1) = 1

d(2, 3) = 1

d(2, 4) = min ((2, 1, 3, 4), (2, 3, 4)

= min(3, 2) = 2

d(2, 5) = 3

d(2, 6) = 3

d(2, 7) = 4

d(2, 8) = 4

d(2, 9) = 5 $\boxed{e(2) = 5}$

for node 3

D = 3, {1, 2, 4, 5, 6, 7, 8, 9}

d(3, 1) = 1

d(3, 2) = 1

d(3, 4) = 1

d(3, 5) = 2

d(3, 6) = 2 $\boxed{e(3) = 4}$

d(3, 7) = 3

d(3, 8) = 3

d(3, 9) = 4

for node 4

D = 4, {1, 2, 3, 5, 6, 7, 8, 9}

d(4, 1) = 1

d(4, 2) = 2

d(4, 3) = 1

d(4, 5) = 1 $\boxed{e(4) = 3}$

d(4, 6) = 1

d(4, 7) = 2

d(4, 8) = 2

d(4, 9) = 3

for node 5

$D = 5, \{1, 2, 3, 4, 6, 7, 8, 9\}$

$d(5,1) = 2$

$d(5,2) = 3$

$d(5,3) = 2$

$d(5,4) = 1$

$d(5,6) = 1$

$d(5,7) = 1$

$d(5,8) = 1$

$d(5,9) = 2$

$\boxed{e(5) = 3}$

for node 6

$D = 6, \{1, 2, 3, 4, 5, 7, 8, 9\}$

$d(6,1) = 2$

$d(6,2) = 3$

$d(6,3) = 2$

$d(6,4) = 1$

$d(6,5) = 1$

$d(6,7) = 1$

$d(6,8) = 1$

$d(6,9) = 2$

$\boxed{e(6) = 3}$

for node 7

$D = 7, \{1, 2, 3, 4, 5, 6, 8, 9\}$

$d(7,1) = 3$

$d(7,2) = 4$

$d(7,3) = 3$

$d(7,4) = 2$

$d(7,5) = 1$

$d(9,6) = 1$

$d(7,8) = 1$

$d(7,9) = 1$

$\boxed{e(7) = 4}$

For node 8

$D = 8, \{1, 2, 3, 4, 5, 6, 7, 9\}$

$d(8,1) = 3$

$d(8,2) = 4$

$d(8,3) = 3$

$d(8,4) = 2$

$d(8,5) = 1$

$d(8,6) = 1$ $\boxed{e(8) = 4}$

$d(8,7) = 1$

$d(8,9) = 2$

for node 9

$D = 9, \{1, 2, 3, 4, 5, 6, 7, 8\}$

$d(9,1) = 4$

$d(9,2) = 5$

$d(9,3) = 4$

$d(9,4) = 3$

$d(9,5) = 2$ $\boxed{e(9) = 5}$

$d(9,6) = 2$

$d(9,7) = 1$

$d(9,8) = 2$

Diameter of the graph = Max of all the Calculated eccentricities

$e(1) = 4$

$e(2) = 5$

$e(3) = 4$

$e(4) = 3$ $\quad$ Max $= \{e(2), e(9)\}$

$e(5) = 3$ $\quad = 5$

$e(6) = 3$ $\quad$ Diameter $= 5$

$e(7) = 4$

$e(8) = 4$

$e(9) = 5$

c) Density of the graph :

$$\text{Density}(D) = \frac{2|E|}{|V|(|V|-1)}$$

$|E| = 14$, $|V| = 9$

$$D = \frac{2 \times 14}{9(9-1)} = \frac{2 \times 14}{9 \times 8} = \frac{28}{72}$$

$= 0.388$

Density = 0.388

e) periphery of the graph =
nodes who's ecentricity is equal to the diameter of the graph

Node '2' and Node '9'

f) Clustering coefficients of each node :

$CC(v)$: $v \to$ a node
$k_v \to$ its degree
$N_v \to$ No. of links b/w neighbours of v

for Node 1 :

$k_v = 3$
$N_v = 2$

$$CC(1) = \frac{2N_v}{k_v(k_v-1)}$$

$$= \frac{2 \times 2}{3 \times 2} = \frac{2}{3}$$

$\boxed{CC(1) = 0.667}$

for Node ②:

$K_v = 2$

$N_v = 1$

$$CC(2) = \frac{2N_v}{K_v(K_v-1)} = \frac{2\times 1}{2\times 1}$$

$$\boxed{CC(2) = 1}$$

for Node ③:

$K_v = 3$

$N_v = 2$

$$CC(3) = \frac{2N_v}{K_v(K_v-1)} = \frac{2\times 2}{3\times 2}$$

$$\boxed{CC(3) = 0.667}$$

for Node ④:

$K_v = 4$

$N_v = 2$

$$CC(4) = \frac{2N_v}{K_v(K_v-1)} = \frac{2\times 2}{4(3)} = 1/3$$

$$\boxed{CC(4) = 0.33}$$

for Node ⑤:

$K_v = 4$

$N_v = 4$

$$CC(5) = \frac{2\times 4}{4\times 3} = 2/3$$

$$\boxed{CC(5) = 0.667}$$

for Node ⑥:

$K_v = 4$

$N_v = 4$

$$CC(6) = \frac{2\times 4}{4(3)} = 2/3$$

$$\boxed{CC(6) = 0.667}$$

for Node ⑦:

$K_v = 4$

$N_v = 3$

$$CC(7) = \frac{2\times 3}{4\times 3} = 1/2$$

$$\boxed{CC(7) = 0.5}$$

for Node ⑧ :

$K_v = 3$

$N_v = 3$

$$CC(8) = \frac{2N_v}{K_v(K_v-1)} = \frac{2\times3}{3\times2} = 1$$

$$\boxed{CC(8) = 1}$$

for Node ⑨ :

$K_v = 1$

$N_v = 0$

$$\boxed{CC(9) = 0}$$

9) Average clustering Coefficient :

$$\sum_{i=1}^{n} \frac{CC(i)}{n}$$

So,

$$\frac{\sum_{i=1}^{9} CC(i)}{9}$$

$$= \frac{0.667 + 1 + 0.667 + 0.33 + 0.667 + 0.667 + 0.5 + 1 + 0}{9}$$

$$= \frac{5.498}{9} = 0.6108$$

$$\boxed{ACC = 0.6108}$$

Expected Clustering Coefficient :

$$E_{CC} = \frac{3(K-2)}{2(K-1)}$$

$$K = \frac{2E}{N} \qquad E = 14, N = 9$$

$$= \frac{2(14)}{9} = 3.1$$

$$E_{CC} = \frac{3(3.1-2)}{2(3.1-1)} = \frac{3(1.1)}{2(2.1)} = \frac{3.3}{4.2} = 0.785$$

2)

a. Erdos_renyi-graph

1. It is also often referred as the ER graph, is a random graph model introduced by mathematicians paul Erdos and Alfred Renyi.
2. It is used to generate random graphs for research and modelling purpose.
3. The Erdos-renyi graph model has two main variations

$G(n, p)$, $G(n, M)$

n: number of nodes

M: number of edges

p: probability of connecting any two nodes

b. Barabasi_albert_graph

1. BA graph is a type of random graph model used in network science and graph theory.
2. It was introduced by Reka Albert and Albert-Laszlo barabasi in 1999.
3. The BA model is designed to capture the "scale-free" property observed in many-real world complex networks.

C.

1. Karate club graph:

It Captures the interactions and friendships between members of the club.

2. Zachary's Karate club graph:

It is often used for Community detection and social network analysis.

3. Les Miserables graph:

It is a weighted graph where the edges represent interactions between characters, and edge weights may denote the frequency of interactions.

4)

a) specify with nodes and edges represent in network.

Nodes - Represent webpages
Edges - Represent hyperlinks or connections b/w web pages

b) specify node attributes and edge/link attributes.

Node :- URL, domain ; Edge:- hyperlink, timestamp

c) Link addition and removal

links are added to webpages when they are created, or updated

d) Hubs in network :

Hubs represent web pages that are have very high number of incoming links

e) Analytical applications :

- page rank analysis
- Community detection
- link prediction

6) Draw the plot of degree distribution and verify power law of any real world example (SNAP)

- Choose a dataset from stanford SNAP
- Load and create a network from the dataset

- Calculate the degree distribution
- plot it on log-log scale
- fit a powerlaw distribution and check the goodness of fit

powerlaw degree distribution should appear as a straight line on the log-log plot with a low p-value for goodness of fit test.