Page :
Date : / /

Common Attributes:

1) User ID - Unique Identification for user
2) Session ID - " " " session
3) Device type -

CS techniques:

① Path Analysis:
To analyze the path the user used to reach that site. (so, previous sites/ references). The way the user takes on particular apps & websites. Helpful for business to find the most common path to reach the site.

- Drop off points:
The point where we left the product to buy. Like payment failure, specification pg. Optimise user flow.

Analyses the sequence of pgs visited by user to understand the most common path taken is called Path Analysis.

Steps:

1. Define goal - User engagement / optimisation
2. Gather CS Data - improve the sequence of pgs visited by user
3. Analysis the path - most common path
4. Optimise the user experience

Page :
Date : / /

② ~~Session Analysis~~: ~~(Unders~~

Steps:

1. Determine duration of session
2. Collect CS Data
3. Analysis the data (patterns)
4. Optimise

③ Conversion - funnel Analysis:

- applying filters to get target product
- which filter frequently used will be recommended for other.

CS Analysis Tools

① Google Analytics - Used for web Analytics tool (Tracking, analysis of behaviour, report generation)

② Adobe Analytics - Advanced CSA (Real time data analysis, Segmentation predictive)

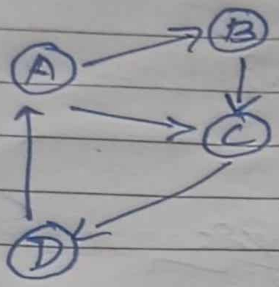
Challenges:

-> Data Volume
-> Data privacy & security
-> Integration with other systems

Page :
Date : / /

10/03/25

Prob: To calculate the page rank for the directed graph with nodes A, B, C, D & the given edges $A \to B$, $A \to C$, $B \to C$, $C \to D$, $D \to A$.



Sol:

① Represent the graph as a transition matrix i.e., Probability of transitions from one node to other node

$$M = \begin{array}{c|cccc} & A & B & C & D \\ \hline A & 0 & 0 & 0 & 1 \\ B & 1/2 & 0 & 0 & 0 \\ C & 1/2 & 1 & 0 & 0 \\ D & 0 & 0 & 1 & 0 \end{array}$$

$\frac{1}{2}$ the choice of is taken for probability, we can't have 1.

Max prob is 1, No of edges = 2, so, 1/2

Column wise, prob shld always be 1.

② Initialise the page rank vector.

$$P^0 = \begin{bmatrix} 1/4 \\ 1/4 \\ 1/4 \\ 1/4 \end{bmatrix}$$

Total nodes = 4
so, 1/4

$$\Rightarrow = \begin{bmatrix} 0.25 \\ 0.25 \\ 0.25 \\ 0.25 \end{bmatrix}$$

③ Apply the damping factor (d) accounts for the probability a user may randomly jump to any ~~a~~ page.

α or $d = 0.85$ (constant)

Page rank formula becomes

$$\underset{\text{(prob to jump to nxt pg)}}{P^{(k+1)}} = d \cdot H \cdot \underset{\text{(prob to jump to prev pg)}}{P^{(k)}} + (1-d) \cdot v$$

v is the vector that is uniform distribution.

$$v = \begin{bmatrix} 1/4 \\ 1/4 \\ 1/4 \\ 1/4 \end{bmatrix}$$

$$P^{(k+1)} = 0.85 \underset{4\times4}{\begin{bmatrix} 0 & 0 & 0 & 1 \\ 1/2 & 0 & 0 & 0 \\ 1/2 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}} \underset{4\times1}{\begin{bmatrix} 0.25 \\ 0.25 \\ 0.25 \\ 0.25 \end{bmatrix}} + (1-0.85) \begin{bmatrix} 0.25 \\ 0.25 \\ 0.25 \\ 0.25 \end{bmatrix}$$

Page :
Date : / /

$$p^{(k+1)} = 0.85 \begin{bmatrix} 0.25 \\ 0.25/2 \\ 0.25/2 + 0.25 \\ 0.25 \end{bmatrix} + 0.15 \begin{bmatrix} 0.25 \\ 0.25 \\ 0.25 \\ 0.25 \end{bmatrix}$$

$$= \begin{bmatrix} 0.2125 \\ 0.10625 \\ 0.31875 \\ 0.2125 \end{bmatrix} + \begin{bmatrix} 0.0375 \\ 0.0375 \\ 0.0375 \\ 0.0375 \end{bmatrix}$$

$$P(1) = \begin{bmatrix} 0.25 \\ 0.14375 \\ 0.35625 \\ 0.25 \end{bmatrix}$$

$$P(2) = 0.85 \begin{bmatrix} 0 & 0 & 0 & 1 \\ 1/2 & 0 & 0 & 0 \\ 1/2 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} 0.25 \\ 0.14 \\ 0.35 \\ 0.25 \end{bmatrix} + 0.15 \begin{bmatrix} 0.25 \\ 0.25 \\ 0.25 \\ 0.25 \end{bmatrix}$$

$$= 0.85 \begin{bmatrix} 0.25 \\ 0.25/2 \\ 0.25/2 + 0.14 \\ 0.35 \end{bmatrix} + \begin{bmatrix} 0.0375 \\ 0.0375 \\ 0.0375 \\ 0.0375 \end{bmatrix}$$

$$= \begin{bmatrix} 0.2125 \\ 0.10625 \\ 0.22525 \\ 0.2975 \end{bmatrix} + \begin{bmatrix} 0.0375 \\ 0.0375 \\ 0.0375 \\ 0.0375 \end{bmatrix} = \begin{bmatrix} 0.25 \\ 0.14 \\ 0.26 \\ 0.34 \end{bmatrix}$$

0.26
0.14
0.39
0.31

Page : ______
Date : / /

$$P(3) = 0.85 \begin{bmatrix} 0 & 0 & 0 & 1 \\ 1/2 & 0 & 0 & 0 \\ 1/2 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} 0.25 \\ 0.14 \\ 0.26 \\ 0.34 \end{bmatrix} + 0.15 \begin{bmatrix} 0.25 \\ 0.25 \\ 0.25 \\ 0.25 \end{bmatrix}$$

$$= 0.85 \begin{bmatrix} 0.34 \\ 0.25/2 \\ 0.25/2 + 0.\overset{14}{\cancel{26}} \\ 0.34 \end{bmatrix} + \begin{bmatrix} 0.0375 \\ 0.0375 \\ 0.0375 \\ 0.0375 \end{bmatrix}$$

$$= \begin{bmatrix} 0.289 \\ 0.10625 \\ 0.\overset{22525}{\cancel{32725}} \\ 0.289 \end{bmatrix} + \begin{bmatrix} 0.0375 \\ 0.0375 \\ 0.0375 \\ 0.0375 \end{bmatrix} = \begin{bmatrix} 0.3265 \\ 0.14 \\ 0.26 \\ 0.33 \end{bmatrix}$$

$$P(4) = 0.85 \begin{bmatrix} 0 & 0 & 0 & 1 \\ 1/2 & 0 & 0 & 0 \\ 1/2 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} 0.33 \\ 0.14 \\ 0.26 \\ 0.33 \end{bmatrix} + 0.15 \begin{bmatrix} 0.25 \\ 0.25 \\ 0.25 \\ 0.25 \end{bmatrix}$$

$$= 0.85 \begin{bmatrix} 0.33 \\ 0.33/2 \\ 0.33/2 + 0.14 \\ 0.26 \end{bmatrix} + \begin{bmatrix} 0.0375 \\ 0.0375 \\ 0.0375 \\ 0.0375 \end{bmatrix}$$

$$= \begin{bmatrix} 0.2805 \\ 0.14025 \\ 0.25925 \\ 0.221 \end{bmatrix} + \begin{bmatrix} 0.0375 \\ 0.0375 \\ 0.0375 \\ 0.0375 \end{bmatrix} = \begin{bmatrix} 0.318 \\ 0.177 \\ 0.296 \\ 0.258 \end{bmatrix} = \begin{bmatrix} 0.32 \\ 0.18 \\ 0.30 \\ 0.26 \end{bmatrix}$$

Page :
Date : / /

$$P(5) = 0.85 \begin{bmatrix} 0 & 0 & 0 & 1 \\ 1/2 & 0 & 0 & 0 \\ 1/2 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} 0.32 \\ 0.18 \\ 0.30 \\ 0.26 \end{bmatrix} + 0.15 \begin{bmatrix} 0.25 \\ 0.25 \\ 0.25 \\ 0.25 \end{bmatrix}$$

$$= 0.85 \begin{bmatrix} 0.26 \\ 0.32/2 \\ 0.32/2 + 0.18 \\ 0.30 \end{bmatrix} + \begin{bmatrix} 0.0375 \\ 0.0375 \\ 0.0375 \\ 0.0375 \end{bmatrix}$$

$$= \begin{bmatrix} 0.26 \\ 0.17 \\ 0.33 \\ 0.29 \end{bmatrix}$$

$$P(6) = 0.85 \begin{bmatrix} 0 & 0 & 0 & 1 \\ 1/2 & 0 & 0 & 0 \\ 1/2 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \Bigg[$$

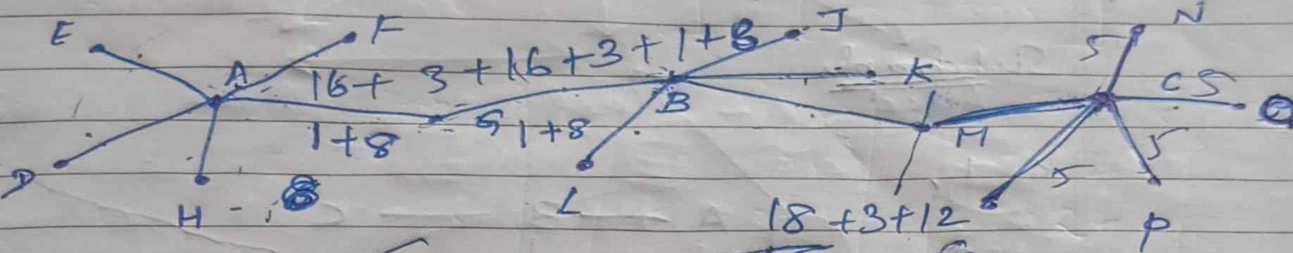
12/3/25

K-Core Graph

Given the following graph: Vertices: A, B, C, D, E, F
Edges: AB, BC, BE, BD, CD, CE, DE, DF

a) Compute k-core collapse sequence for k = 1, 2, 3, and 4 (10 marks)

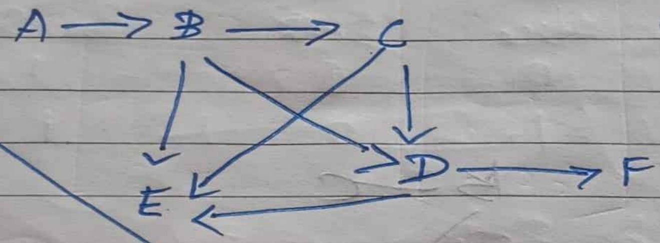
b) Calculate local & global centrality for the following graph (10 marks)



18+3+12
15
33

5+2+12 + 4 + 20

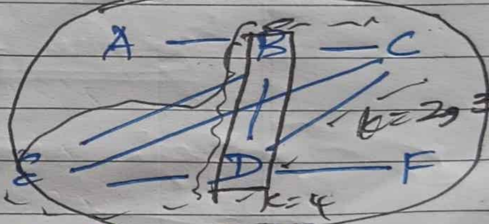
a) 5+ +12



24/12
36

38/5
42

k=1



k=1, ~~B, C, E, D~~ ~~A, F~~ A, F
~~B, C, E, D~~ ~~B, C, D, E~~

k=2 removed nodes A, F
So, remaining nodes > 2 ⇒ B, C, D, E

k=3 remove nodes less than degree 3.
remove A, F
B, C, D, E

Write nodes whose degree not 1, for k=1.

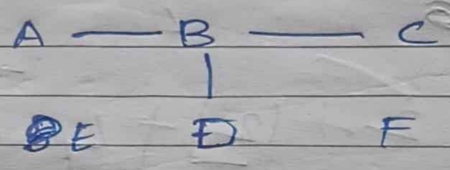
k=4

$\boxed{k < 1 \text{ remove}}$

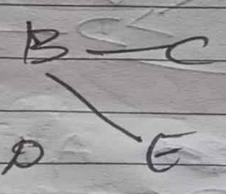
	k		Removed Nodes	Remaing nodes
a)	k = 1	A, F	–	A, B, C, D, E, F
	k = 2		A, F	B, C, D, E
	k = 3	C, E	–	B, C, D, E
	k = 4	B, D	C, E	B, D,

Subgraphs

k = 1,



k = 2



Page :
Date : / /

b) Local ⇒ particular node, how many edges

A, B, C ⇒ 5
G, H ⇒ 2
J, K, L ⇒ 1
All other pts ⇒ 1

Global Centrality

Global centrality for node A ⇒ $5+2+12+\cancel{8}+4+20 = 43$

Node C ⇒ $5+2+12+4+20 = 43$

Node B ⇒ $5+2+12+2+(4 \times 3) = 33$

Node G ⇒ $1+8+1+8+3+16 = 37$

37 Node H ⇒

JKL
48

others 57

Types of Network Analysis

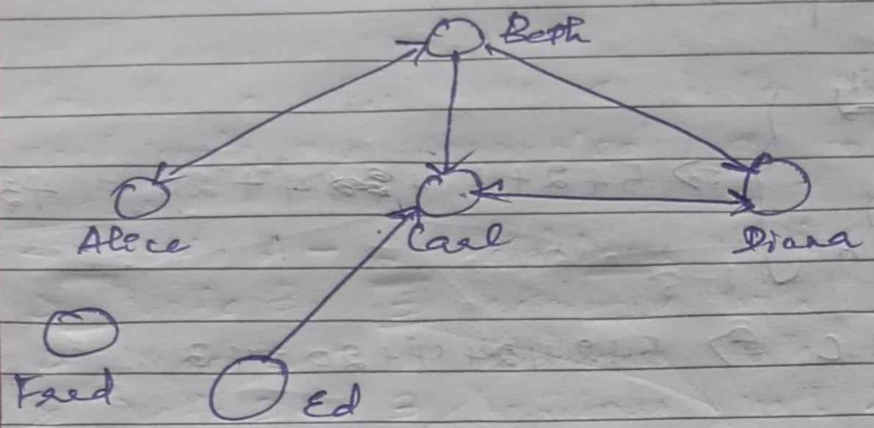
Page :
Date : / /

14/3/25

(i) Ego Network Analysis

Alters - Direct connections to a single node
- Follows star topology

Problem: Extract ego n/w and analyze as independent (Directed graph)



Row Sum = 0
⇓
Independent person

Ego n/w for alice

(Alice)	Alice	Beth	Row Sum
Alice	0	1	1
Beth	1	0	1

Ego n/w for Beth

Beth	Beth	Alice	Carl	Diana	Row sum
$\boxed{\text{Beth}}$	0	1	1	1	3
Alice	1	0	0	0	1
Carl	0	0	0	1	1
Diana	0	0	1	0	1

More influenced person

Page :
Date : / /

Ego N/w for Carl

Carl	Carl	Diana	Row Sum
Carl	0	1	1
Diana	1	0	1

Ego N/w for Diana

Diana	Diana	Carl	Row Sum
Diana	0	1	1
Carl	1	0	1

Ego N/w for Ed

Ed	Ed	Carl	Row Sum
Ed	0	1	1
Carl	0	0	0

1st ⇒ Fred has no connections - so independent.
2nd ⇒ ~~Ed is not independent.~~

(2) Whole N/w Analysis (Three Big centrality)
- Degree (local)
- Betweeness (global)
- Clustering Coeffients (local)

Page :
Date : / /

Clustering Coefficient

$$C = \frac{\text{No of connections among neighbours}}{\text{Max. possible connections among neighbours}}$$

Qn: Consider a n/w with 4 nodes. The neighbours of each node are connected as follows:

* Node A – 2 neighbours with 1 connection b/w them
* Node B – 3 neighbours with 3 connection b/w them
* Node C – 1 neighbours with no connection b/w them
* Node D – 3 neighbours with 2 connection b/w them

Ans: Total possible connections $= \frac{n(n-1)}{2}$ n – no of neighbours

Max possible connections

$$A = \frac{2(2-1)}{2} = 1$$

$$B = \frac{3(3-1)}{2} = 3$$

$$C = \frac{1(1-1)}{2} = 0$$

$$D = \frac{3(3-1)}{2} = 3$$

Neighbours (Given)
A – 2
B – 3
C – 1
D – 3

Connections Among Neighbours (Given)
A – 1
B – 3
C – 0
D – 2

Page :
Date : / /

CC(A) = 1/1 = 1
CC(B) = 3/3 = 1
CC(C) = 0/0 = Not Defined (No connections)
CC(D) = 2/3 = 0.67

$$\text{Avg clustering Coefficients} = \frac{1+1+0.67}{4} = \frac{2.67}{4}$$

$$\text{Avg clustering Coefficients} = 0.667$$

(3) Dynamic N/w analysis

- N/w evolved over time
- focus only on changes in relationship, behaviour, ~~monitoring ...~~
- monitoring how communications pattern shift in an organisation during some crises.
- Track metrics like node churn, edge persistence and network growth rate.
- Often used for event-driven analysis.
- Understand temporal patterns such as formation, dissolution.

(4) Multiplex Network Analysis

- Examine how different types of connections interact & influence behaviour
- Each type of relationship is represented as separate layer
- Analyze interdependencies b/w layers

- Common metrics - layer overlap, multiplex degree

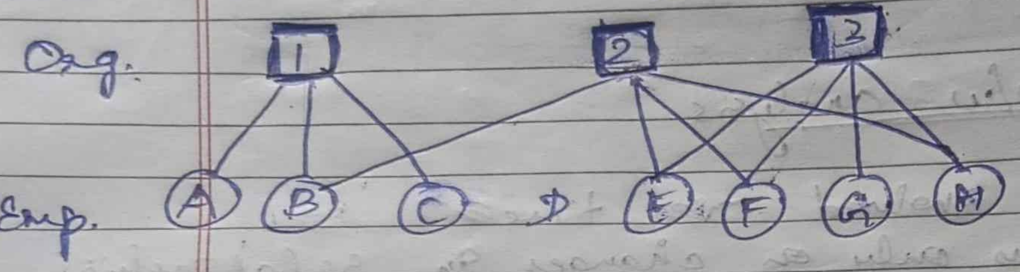
17/3/25

(5) Two Node N/w Analysis:

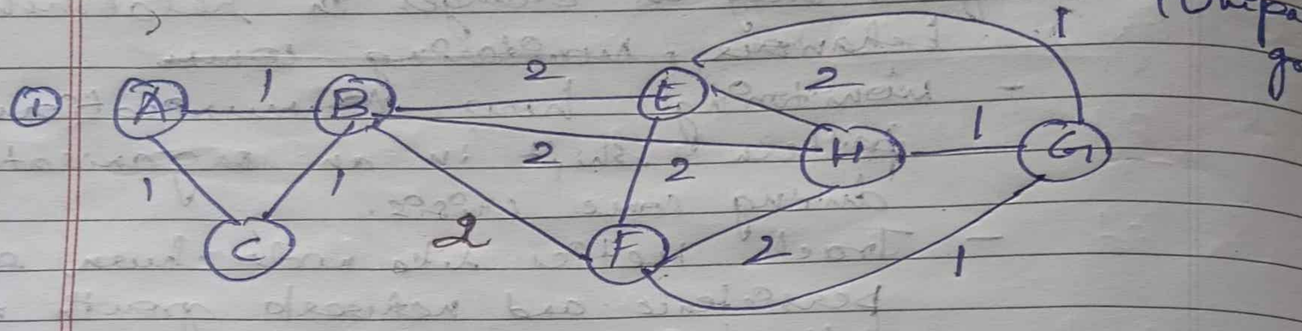
- Relationship b/w 2 different nodes

Ex: Employee & Org, Members & clubs.

Graph (1) How many members involved in more than 1 club?

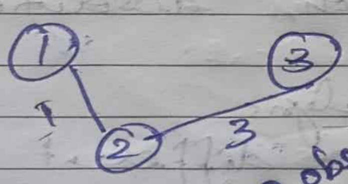


(1)



(Unipath graph)

(2) How many employees are common to each organisation



46% of edges occur in the n/w

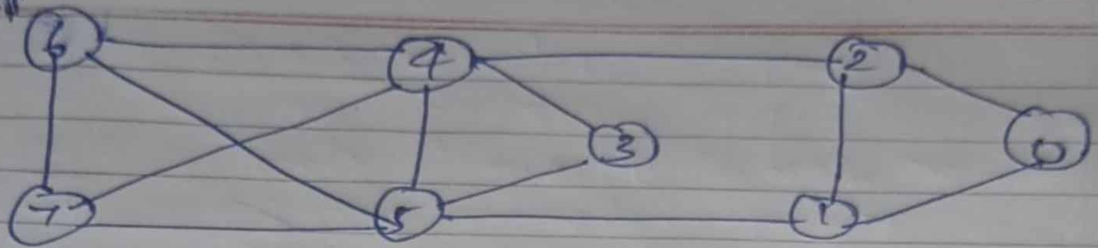
⇑

density for 2 mode graph $= \frac{L}{M \times N} = \frac{\text{No of obs edges}}{\text{Max possible edges}} = \frac{11}{24} = \frac{11}{8 \times 3} = 0.458$

Adj. Matrix

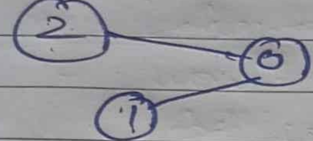
	1	2	3
A	1	0	0
B	1	1	0
C	1	0	0
D	0	0	0
E	0	1	1
F	0	1	1
G	0	0	1
H	0	1	1

Problem

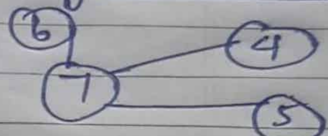


Q) Find the ego centric n/w for node 0, 1, 2, 3, 4, 5, 6, 7 (Undirected)

Sol: Ego n/w for 0

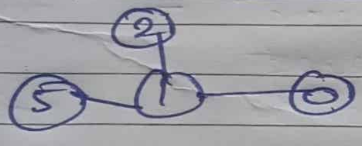


Ego n/w for 7

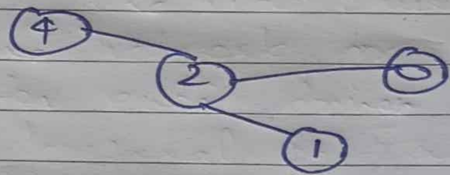


∴ Node 4 & 5 are the most influenced person, as they have star topology

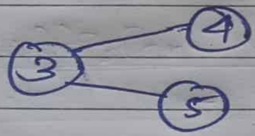
Ego n/w for 1



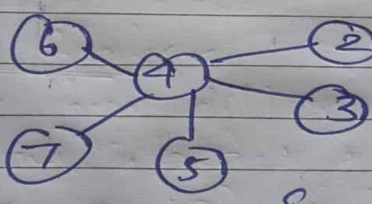
Ego n/w for 2



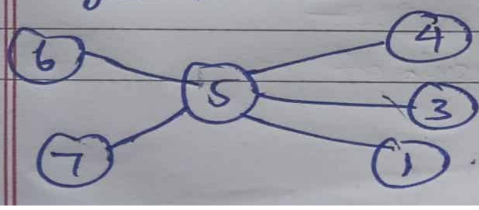
Ego n/w for 3



Ego n/w for 4



Ego n/w for 5



Ego n/w for 6