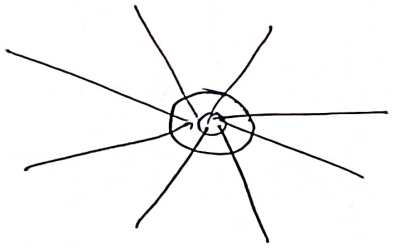
Week-10

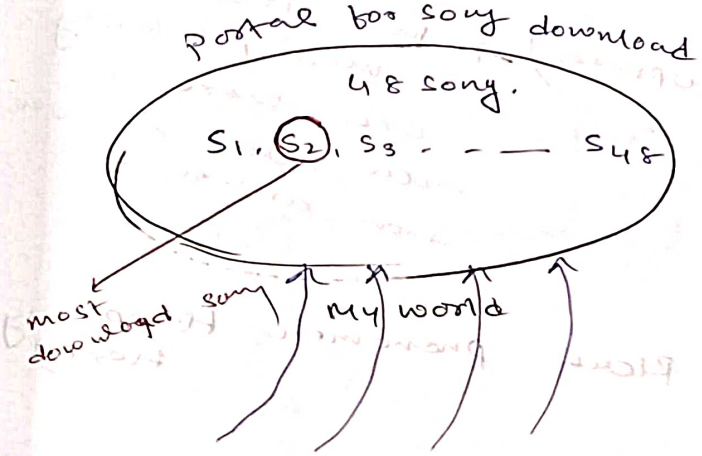
Rich Get Richer Phenomenon - 2

Rich Get Richer - A Possible Reason.

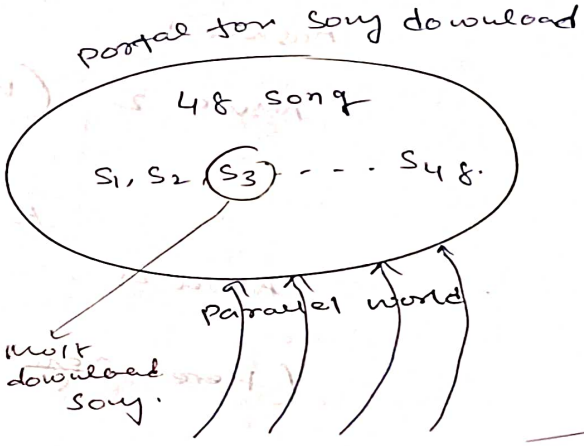


People having high degree can attract more.

→ Hypotheses validated on data for the Rich Get Richer phenomenon.



People come here to download song



people come here to download song

→ Here observed the distribution of 'song download were all different (means most song download in myworld is S_2 and in parallel world most song download is S_3)

→ Reason is "Rich get richer" phenomenon.

Download information.

- Highly downloaded songs biased. new people!

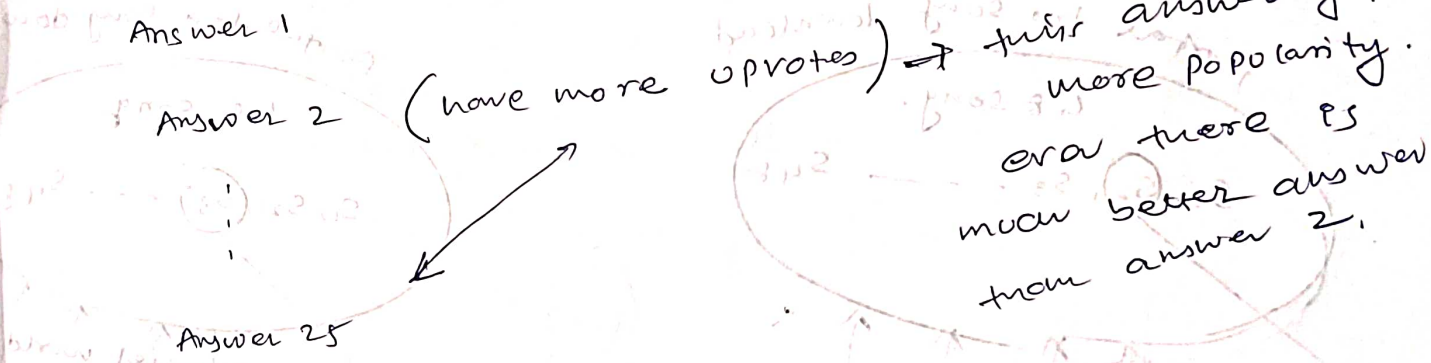
means,

Good gets better....

Example.

Quora.

Question.



(Here ~~Rega~~ Rich get Richer phenomen happening here)

(1) How can one detect the right talent by substracting out the popularity obtain by RGR phenomenon

(2) RGR

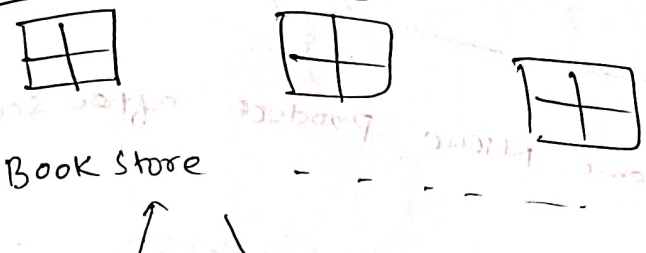
How does one judge the true value of an entity?

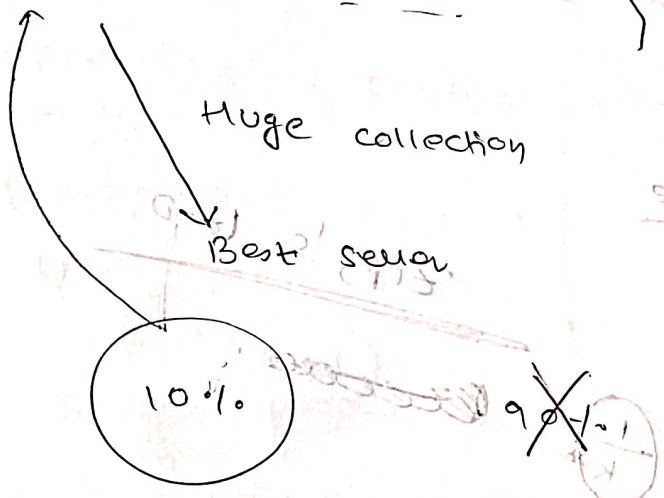
1) Rich get Richer is the reason for power law emergence.

(2) RGR adds noise to Popularity!
How to remove.

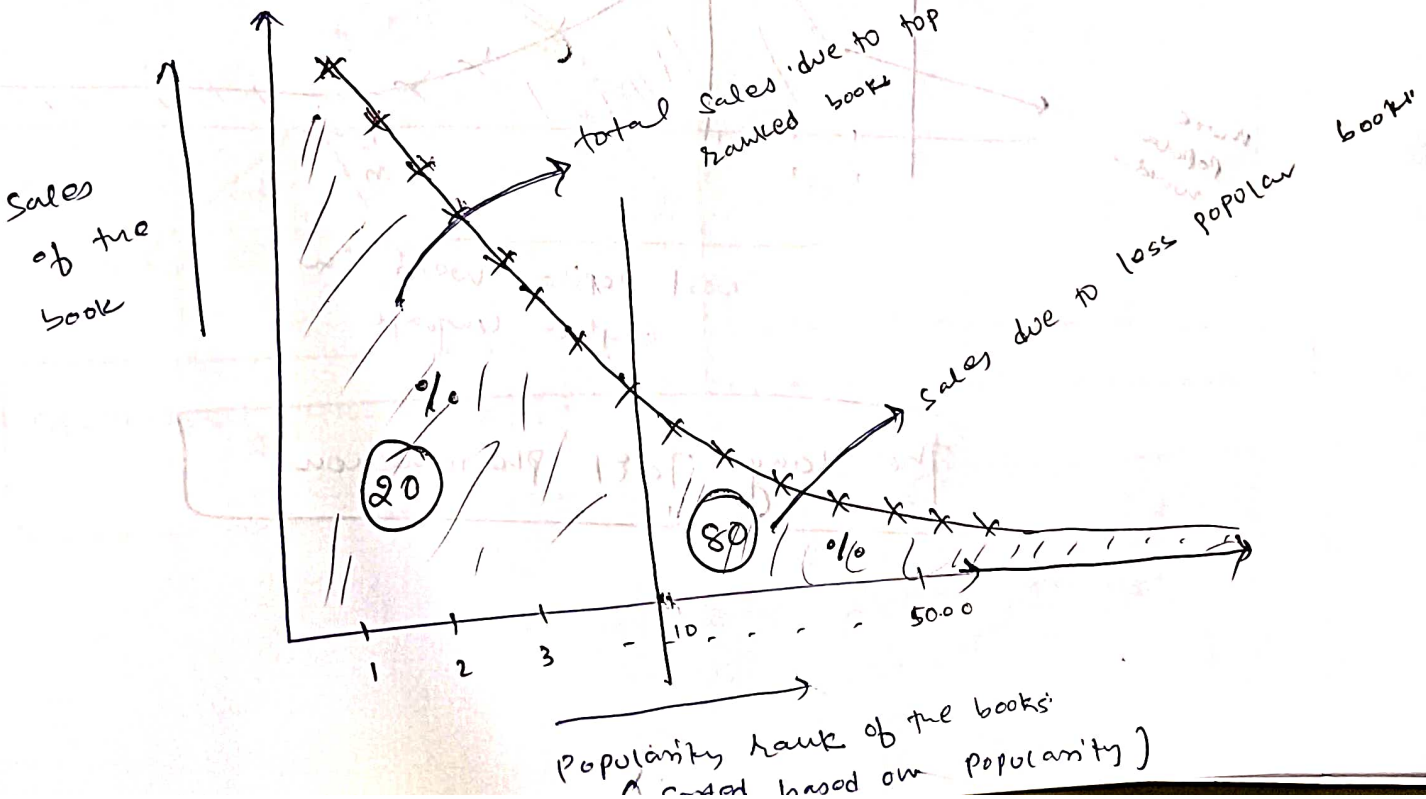
Rich Get Richer - The long Tail

Book store Example

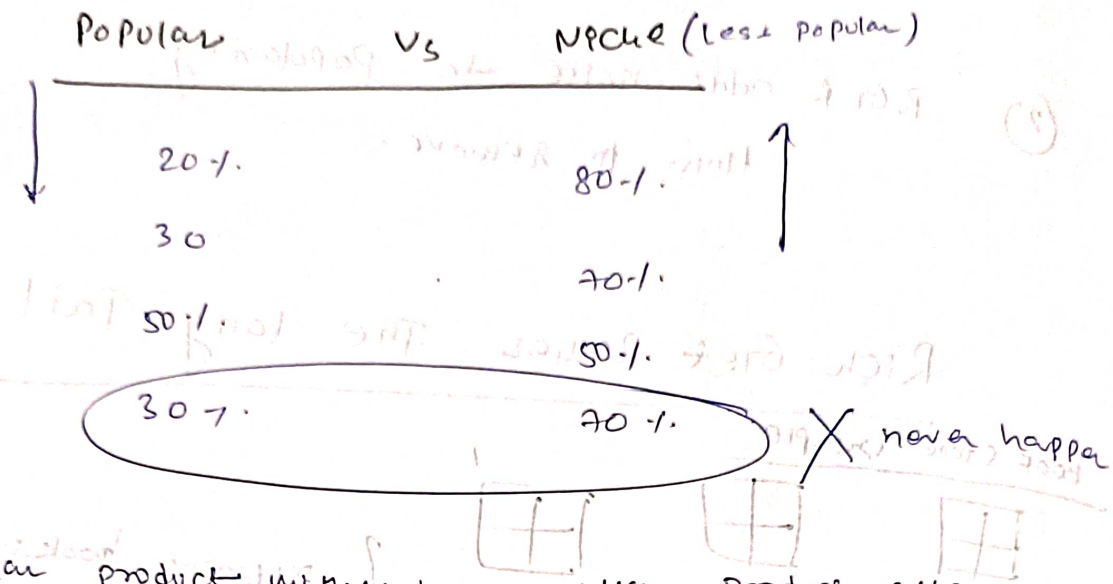




Why can't they just keep Best seller book?

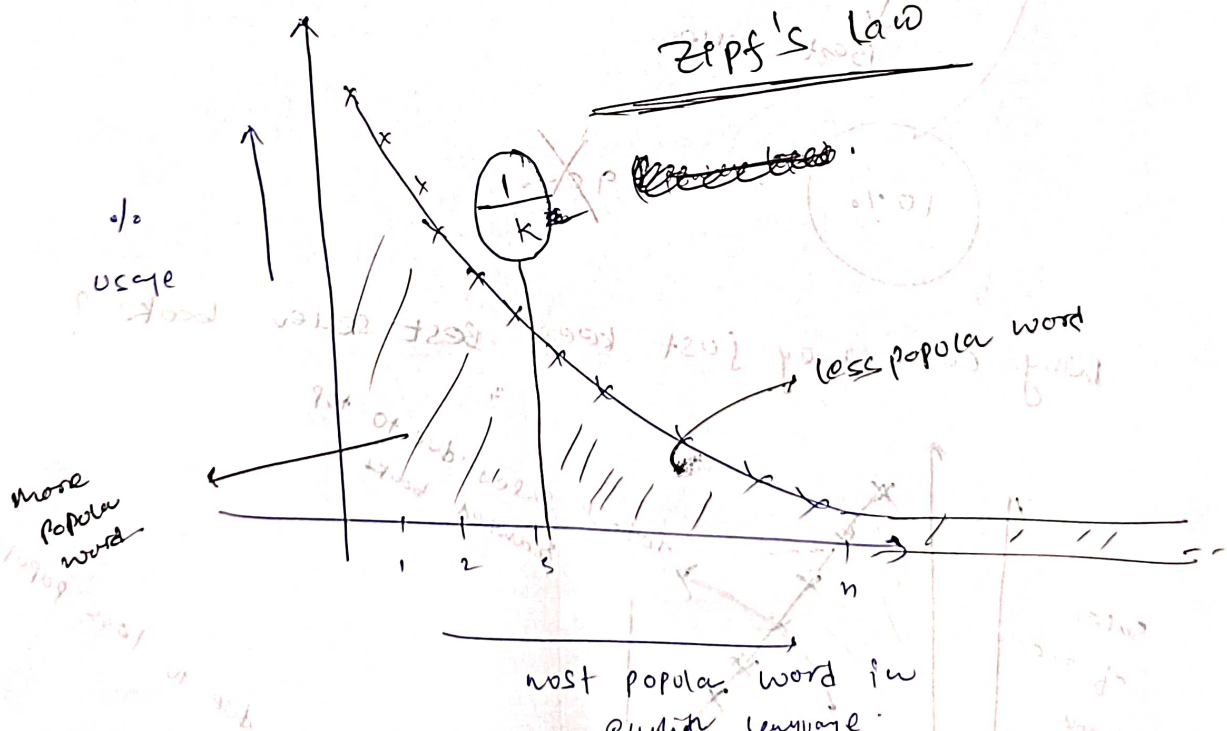


What is the total profit due to the top ranked books?



→ Popular product might become Niche product after some times.

English Language Example



The long Tail phenomenon

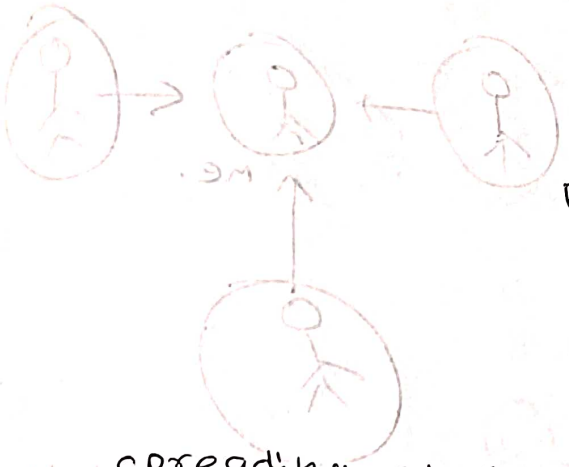
Epidemics -

social contagion: → Actual contagion

⇓

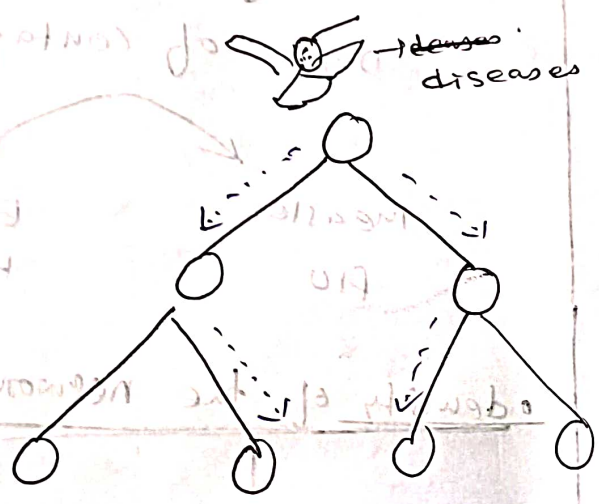
(Contagious Diseases.)

Ex - Ebola
Swine flu.
Black Death.



Spreading of Contagious Diseases.

Biologists &
Epidemiologists
Doctors
Social network scientists. ✓



spread of an idea ≠ spread of the diseases

⇓

they are two different process.

Reason: Spreading of ~~Disease~~ Idea

(i) ~~No~~ choice ✓

(bad song / good song.
So, you have choice)

ii) Invisible process.

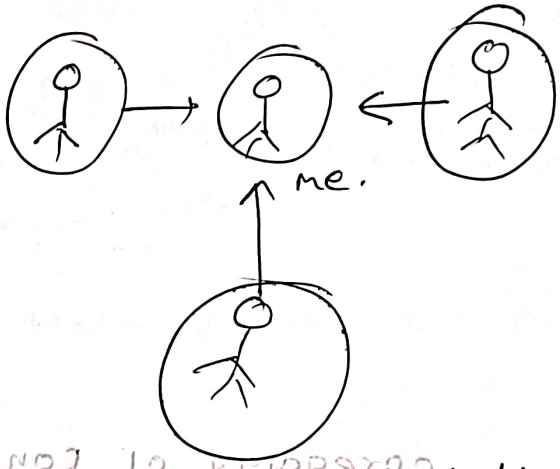
iii) Not Random.

Spreading of diseases

No choice ✗

- Eye flue Example

Invisible process.



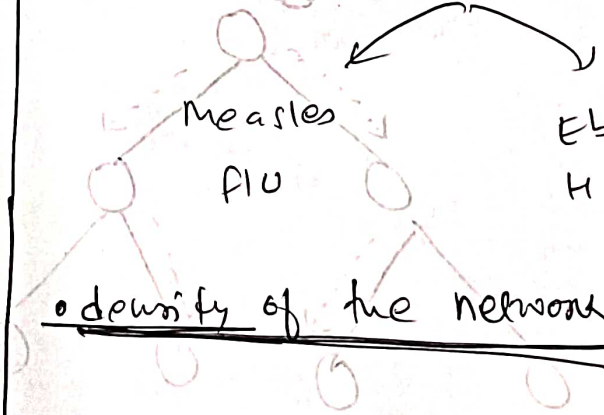
Who will infect ? don't know.

- Random process

pathogen

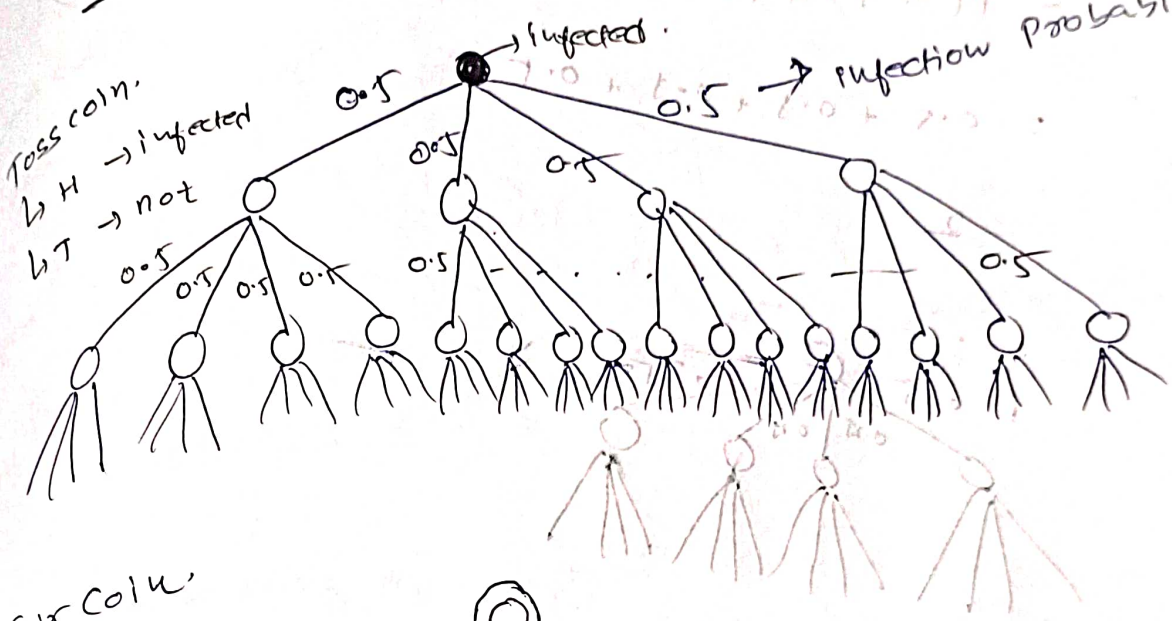
↓

Degree of contagiousness

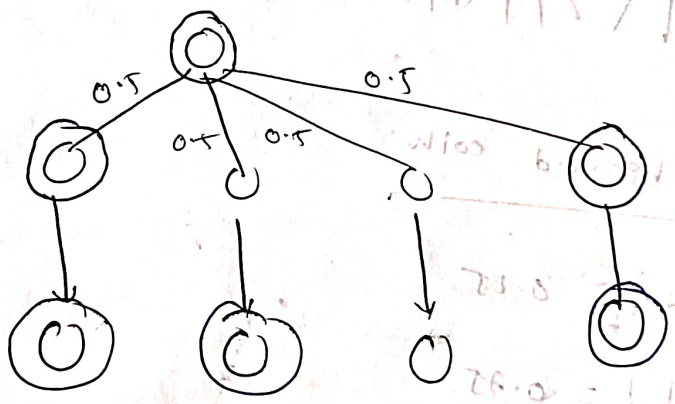


- density of the network also affect

Simple Branching process for Modeling Epidemics



Fair coin.



Experiment → E_1 E_2 E_3 E_4

$E_1 \downarrow x_1$ $E_2 \downarrow x_2$ $E_3 \downarrow x_3$ $E_4 \downarrow x_4$

Number of heads

Random variable X = number of heads.

$x_i = 1$ if head appears

$= 0$ else

$$X = x_1 + x_2 + x_3 + x_4$$

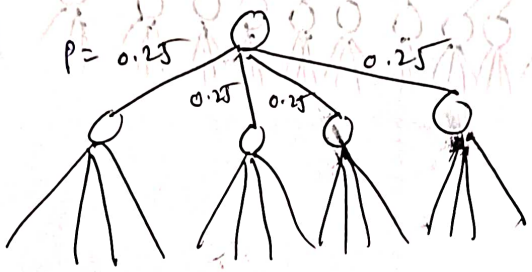
↑ expected value

$$E[X] = E[x_1] + E[x_2] + E[x_3] + E[x_4]$$

$$= P(x_1=1) + P(x_2=1) + P(x_3=1) + P(x_4=1)$$

$$= 0.5 + 0.5 + 0.5 + 0.5$$

$$= 2$$



Toss a biased coin:

P (head) = 0.25

P (tail) = 0.75

Expected number of heads =

$$E[X] = E(x_1) + E(x_2) + E(x_3) + E(x_4)$$

$$= 0.25 \times 4$$

$$= 1$$

Generalise

K-children (every node having k children)

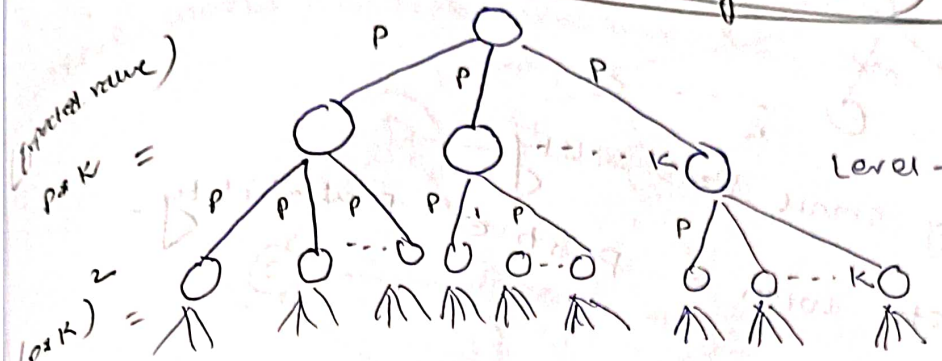
probability of infection = P

$$E[X] = P + P + \cdots \; K$$

$$= P * K$$

Looking at many levels.

(Branching Model)



(Imagine value)

$P * K =$

$(P*K)^2 =$

$(P*K)^3 =$

Level-1 = K

level-2 $K + K + \cdots K = K^2$ (K Times)

level-3 $= K^3$

at Level i = K^i people

P = probability of infection.

at level-1 $= \underbrace{P+P+P\cdots+P}_{K \text{ times}} = PK$ no. of person would (expected) be affected.

> at ith level $(PK)^i$ no. of people infected.

$R_0 = PK$

↓

Basic Reproductive number R_0.

$R_0 < 1$ (can not affect a single person)

let $P = 0.25$, $K = 2$

$R_0 = 0.5$

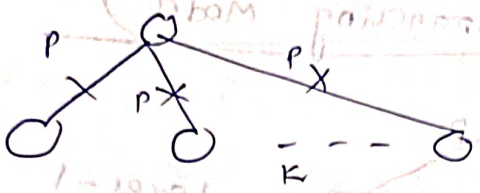
Infection ultimately dies away

$R_0 > 1$ $P = 0.25$ $K = 40$

$R_0 = 10$ (infect 10 people.)

diseases become Epidemics (but Not 100%.)

⇒ there is a still chance that Epidemic die away for $R_0 > 1$



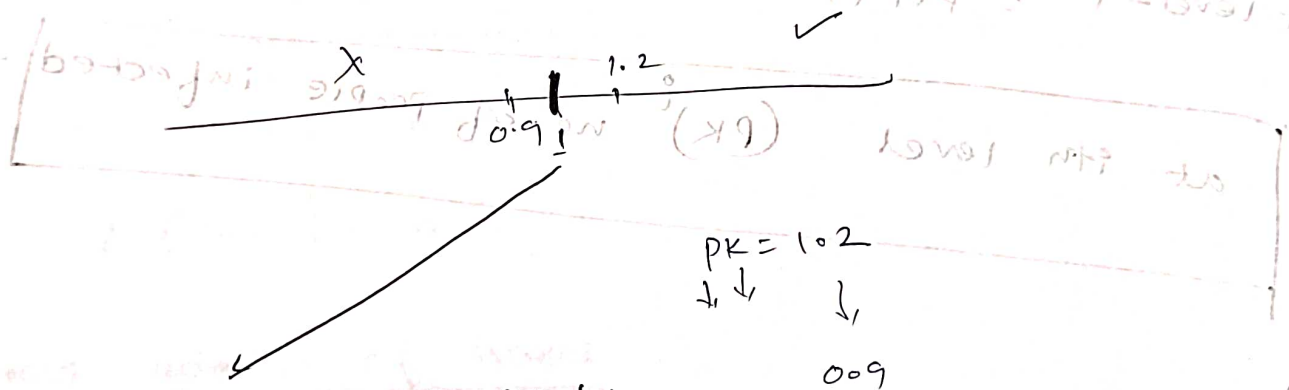
very small Probability

Persists with a positive probability.

Note:-

$R_0 < 1$: Dies away with probability = 1

$R_0 > 1$: Persist in the n/w with +ve probability.



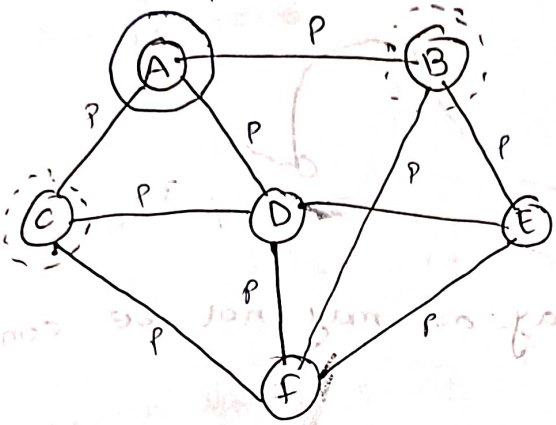
Knife-edge property

Little bit you come left side by decreasing p or K value the disease will vanish but you little bit you goes right the disease will become Epidemics.

Modeling edidemecs on Complex Networks.

real world network does not look like tree.

initially:



Network.

SIR Model. (SIR Epidemics model)

for any type of network.

SIR

- Susceptible (not infected)
- Infected &
- recovered / removed.

T_I = Time for which you are going to be infected.

See after week 12

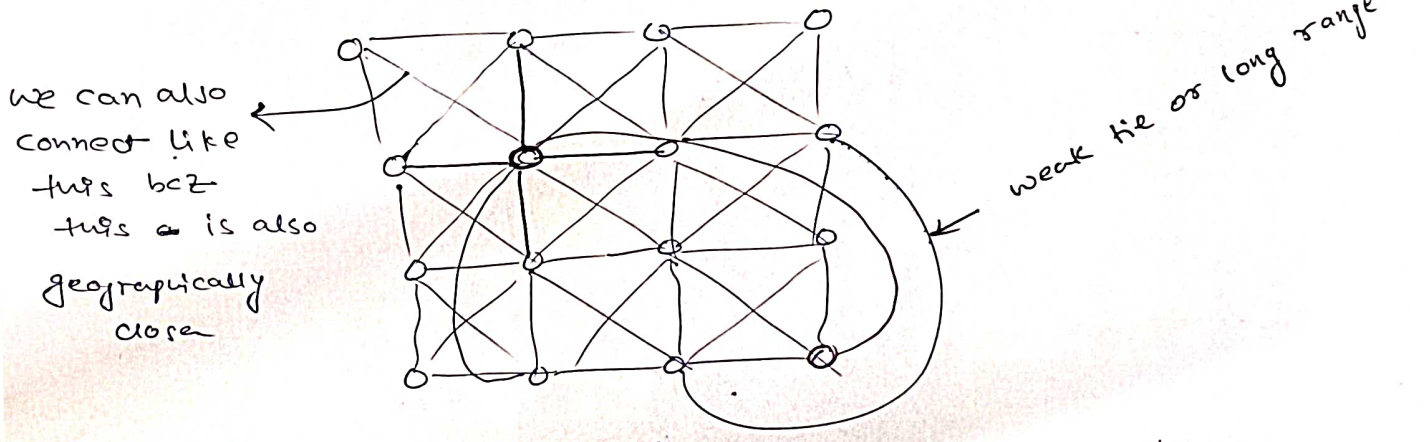
Week - 12

How to go viral on Web

Programming illustration - Small world Networks.

For making a small world network:

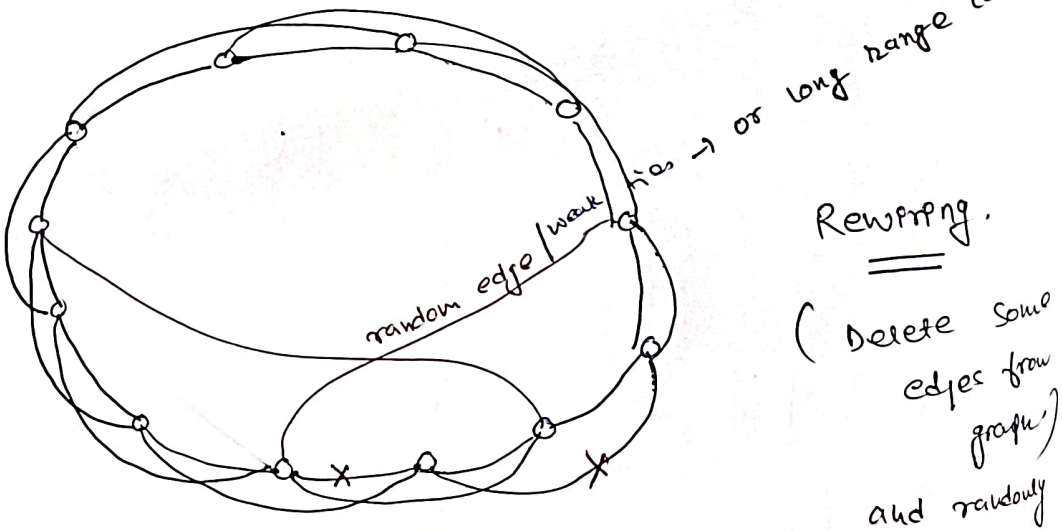
2D Homophily :- geographically closer to the Network



All the link are made because of homophily.

Weak Tie and Homophily makes the world small.

1-D (Ring)



Rewiring.

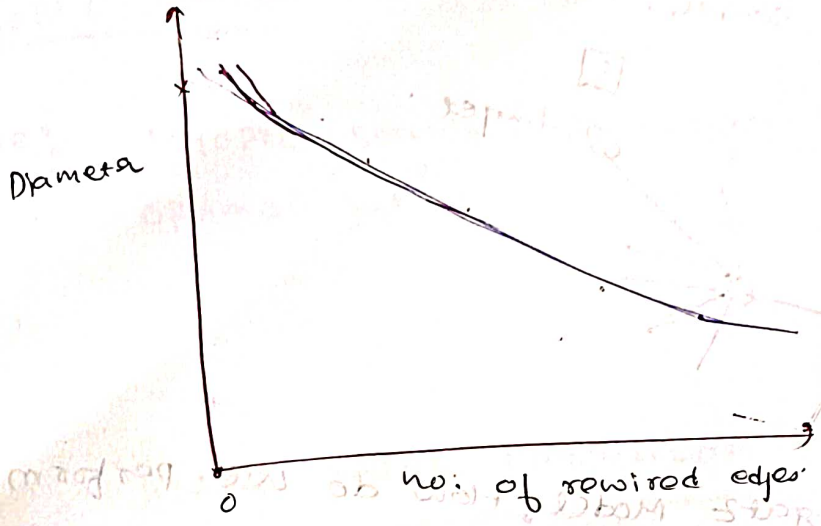
(Delete some edges from graph.)

and randomly and edge

Rewiring

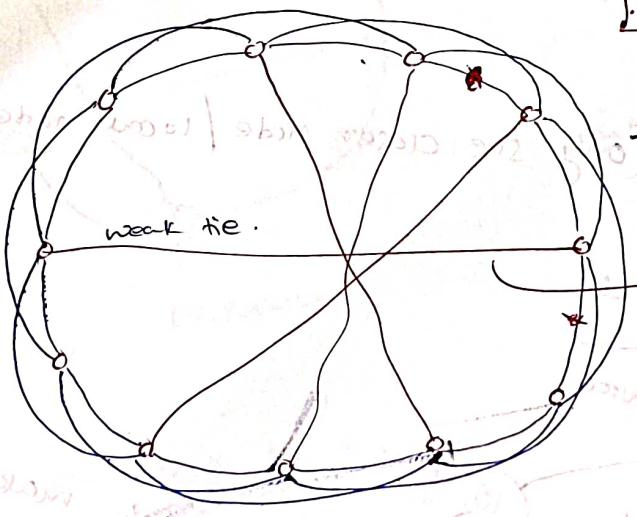
step-① Delete a random edge

step-② add a random edge.



In one D. we are only adding random edge (not deleting random edge)

Diameter :- longest shortest distance



→ Every node is connected to two node towards left and two node towards right.

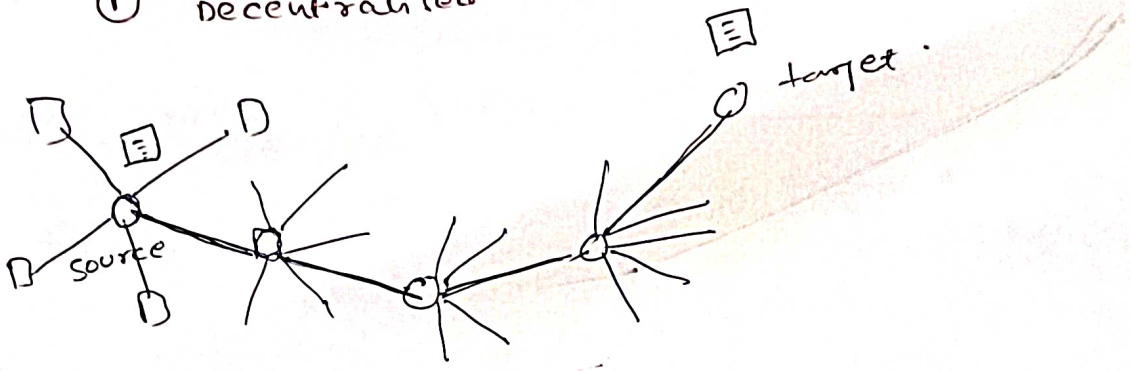
→ adding random edge or only weak tie.

Note :- the number of rewired edges or weak ties increases, then the diameter of the network decreases.

Myopic Search :

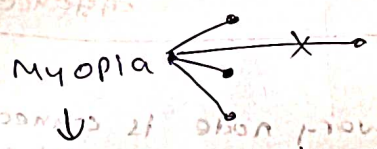
Search on the Small World Network

① Decentralised search.



Given a watts-Strogatz Model: How do we perform Decentralised search

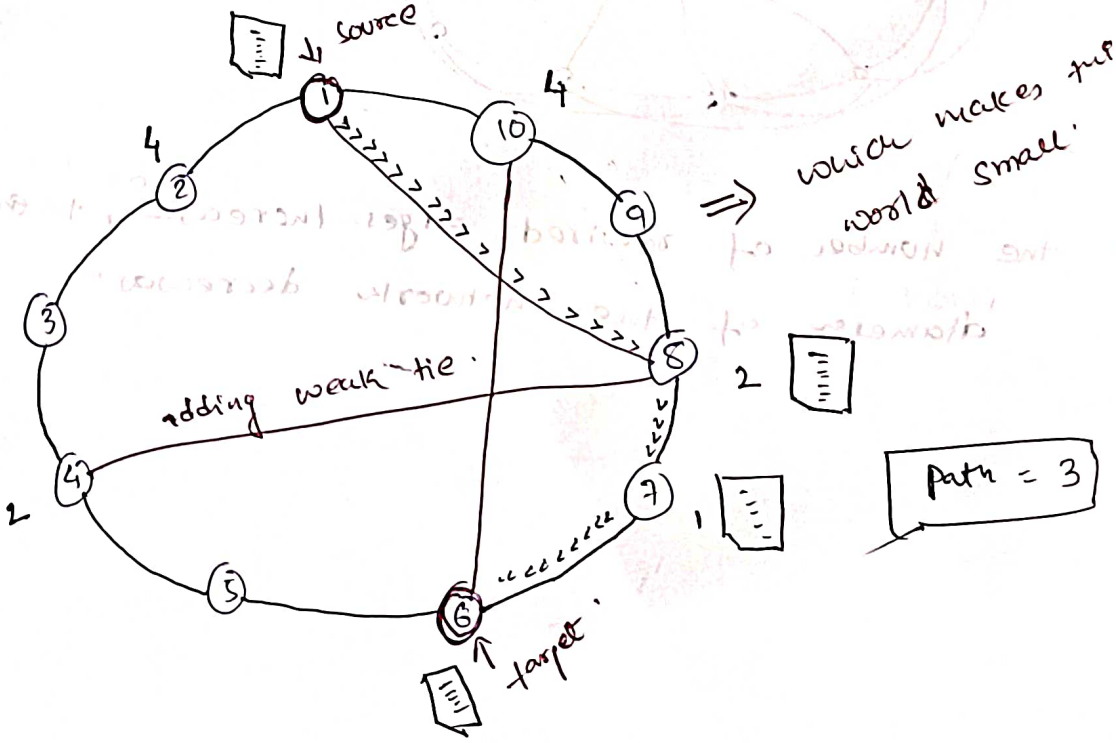
⇒ This kind of search is known as Myopic search.



↓

Short sightedness (only see closer node / local node)

1D watts-Strogatz Model



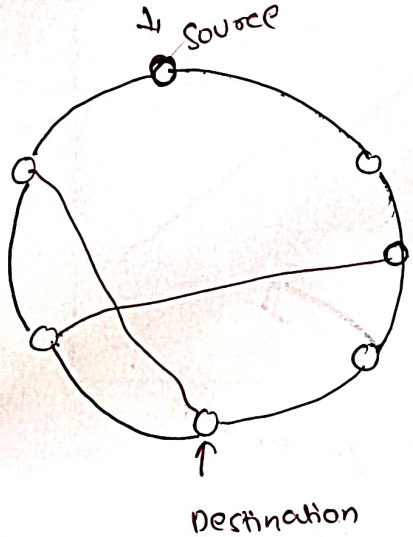
⇒ which makes this world small.

→ Note :- Since node 10 does not know the long range contact of node 1 / weak tie.

Myopic search does not give optimal path:

Note :- Myopic search / Decentralised search does not give optimal path.

Myopic search comparision to optimal search:



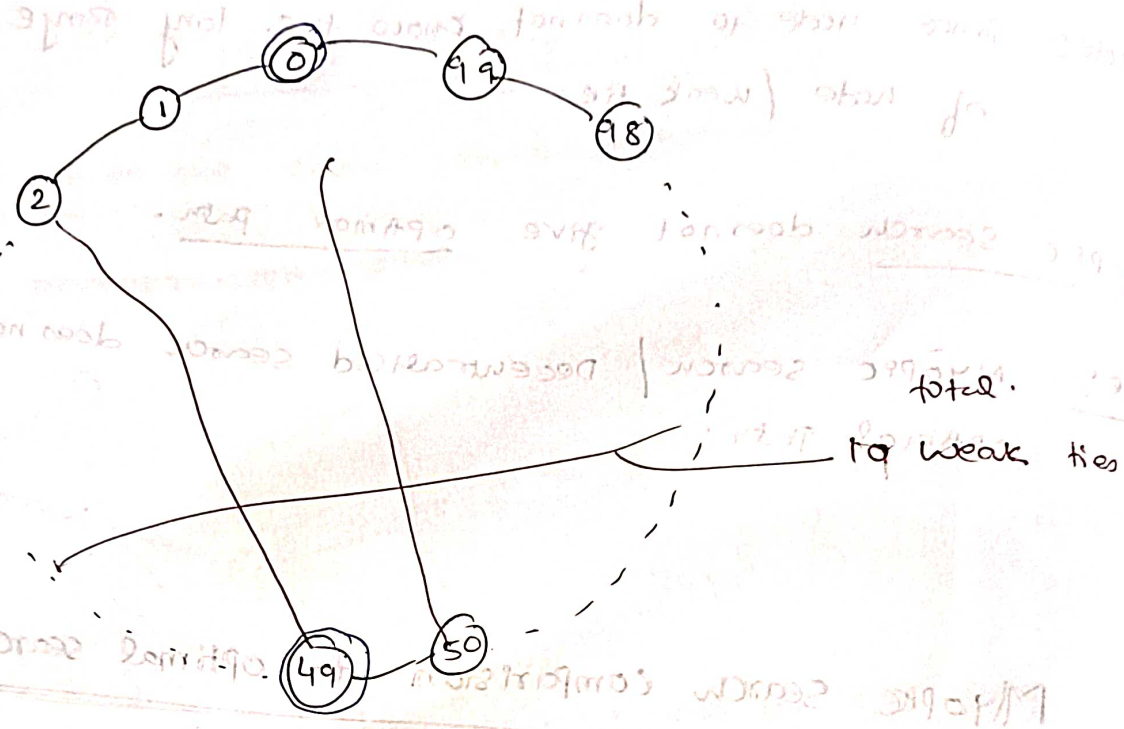
1) Myopic search
2) optimal search.

← In this example Myopic search and optimal search both gives the same result.

Aim :- our aim is to compare Myopic search with the optimal search for the various different pair.

So, which pair do we choose ?

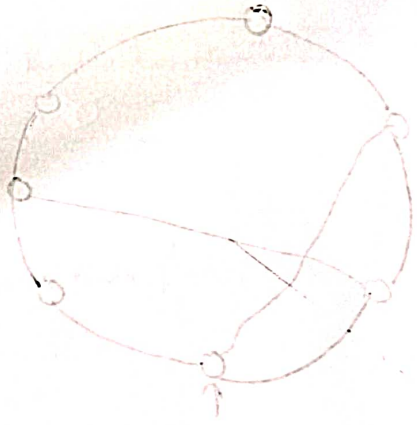
Ans: Best pair for delay this, are the pair which are diametrically opposite.

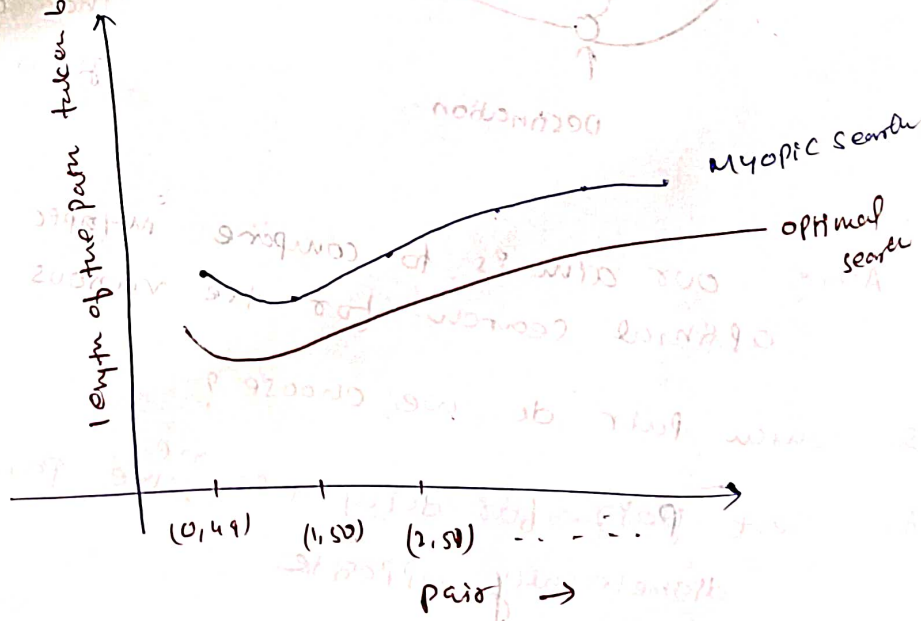


Pair:

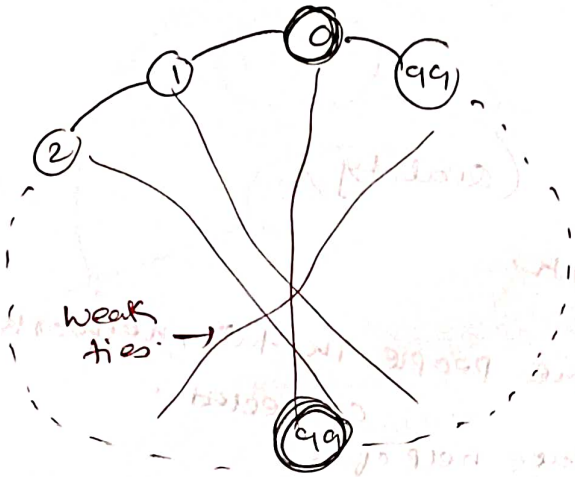
(0, 49)
(1, 50)
⋮
(50, 99)

51 Pairs.



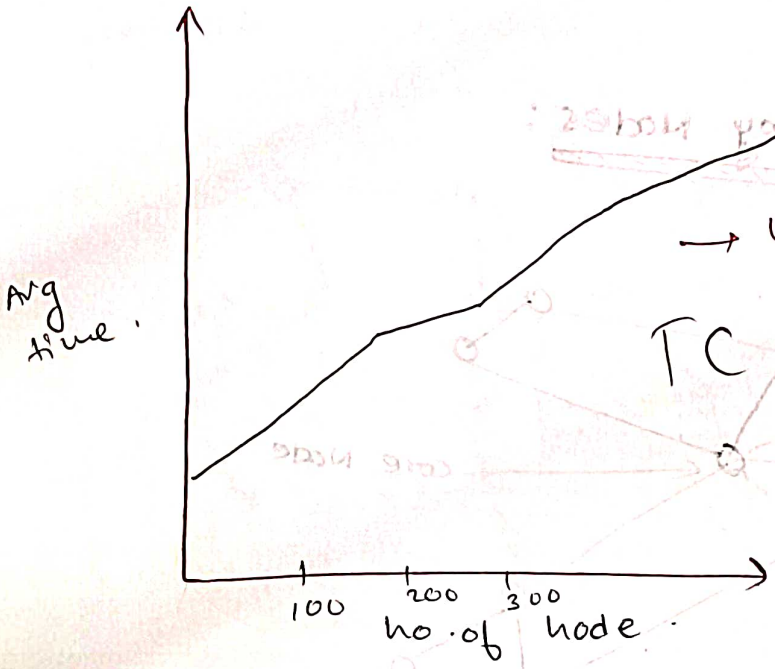


Time Taken by Myopic Search :-



Pair		time taken
(0, 49)	→	t_1
1, 50	→	t_2
⋮		⋮
(50, 99)	→	t_n

avg



→ Linear time.

TC : O(log n)

Myopic search.

How to be viral

How a meme goes viral.

(1) Novelty of meme (quality)

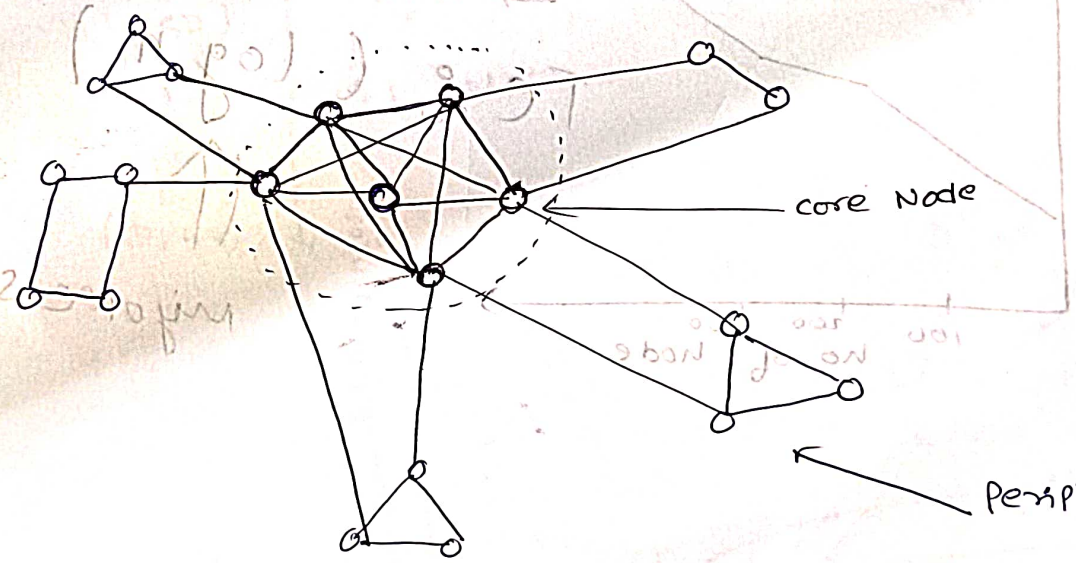
(2) Structure of Network

↓

Means how the people in this network is connected.

(3) Key Nodes (popular person take help of)

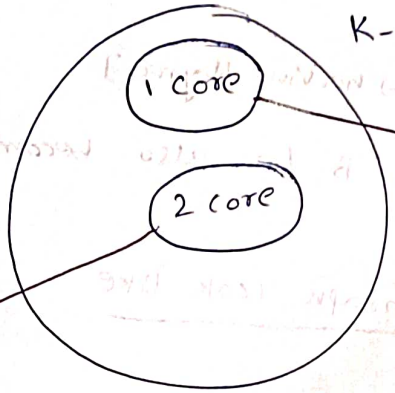
Finding the Right Key Nodes:



→ Onces the inner component is infected, your meme will go, a
reach to the periphori node as well.

→ core node is infected the periphori node.

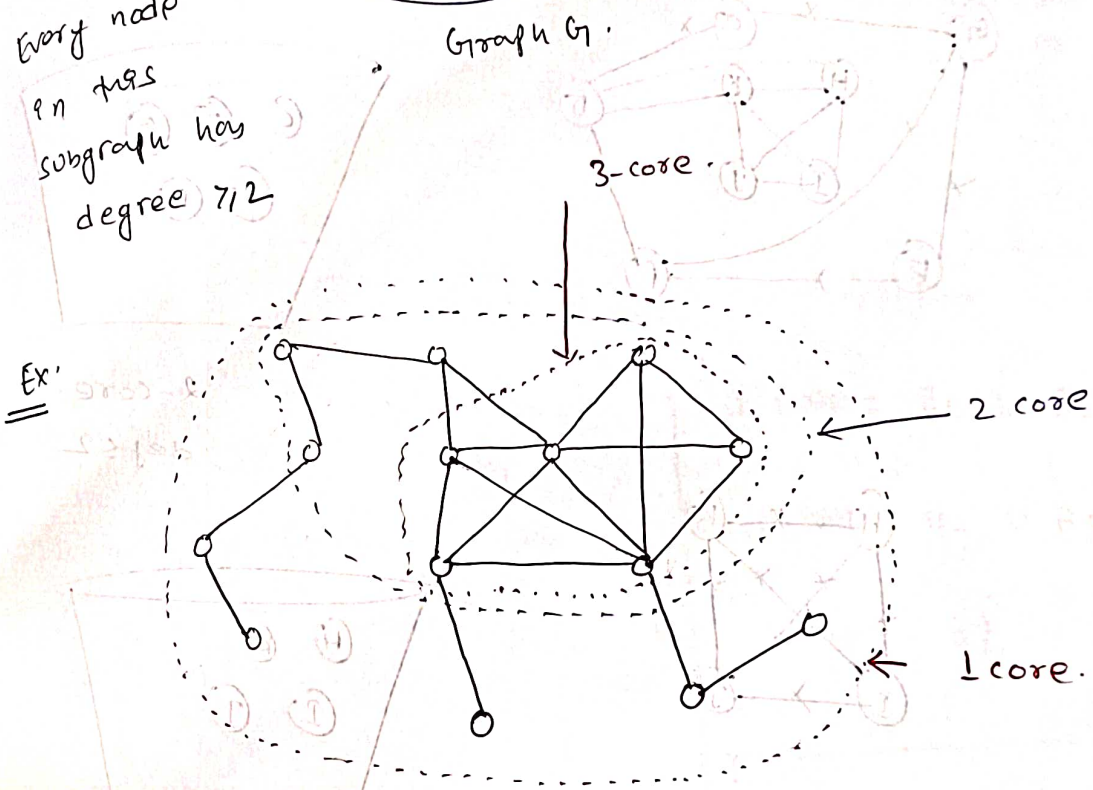
How to Identify core Node?



Every node in this subgraph has degree ≥ 1

Every node in this subgraph has degree ≥ 2

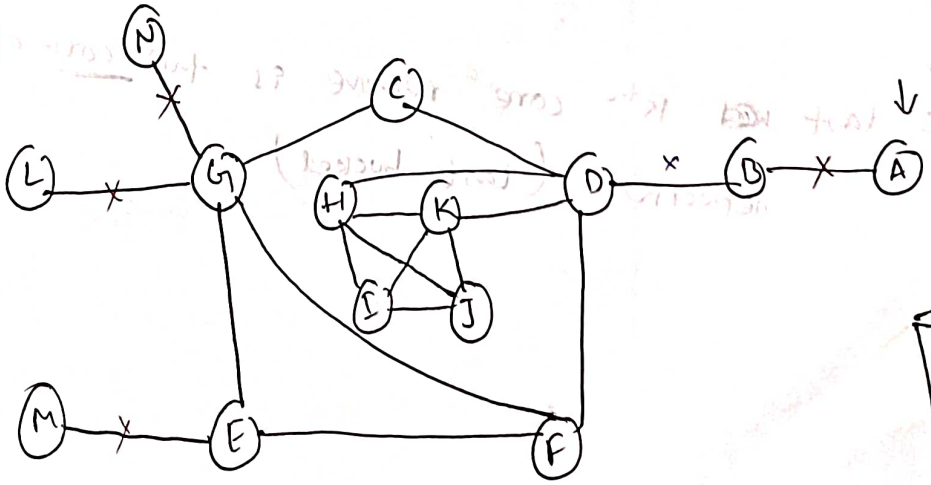
Graph G.



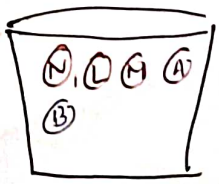
Ex.

K- Shell decomposition Algorithm

K=1



↓ bucket.



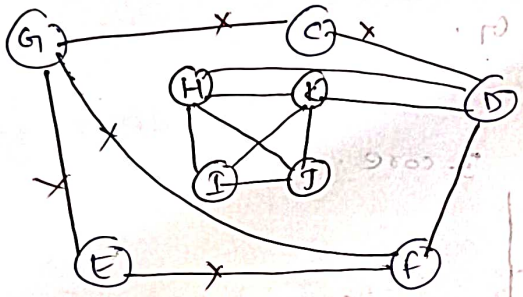
Core-1 (deg = 1)

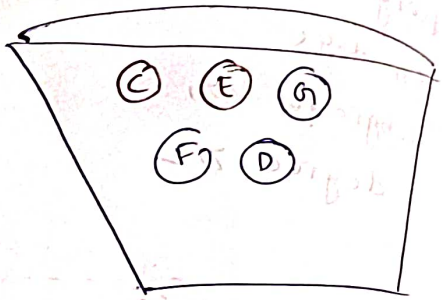
if you keep remove with degree 1. new balls keep emerging in network which has degree 1.

Ex - if we remove (A) → having degree 1
then the degree of B has also become 1.

Now, the New Graph look like

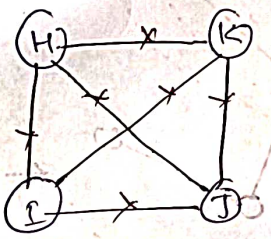
$k=2$



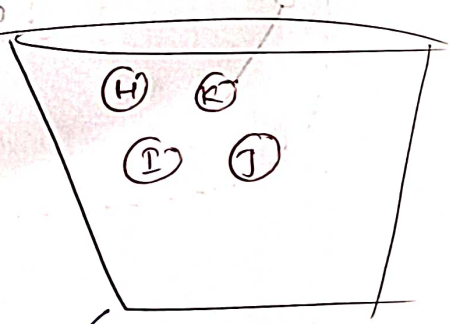


2-core.
deg ≤ 2

$k=3$



Network become empty

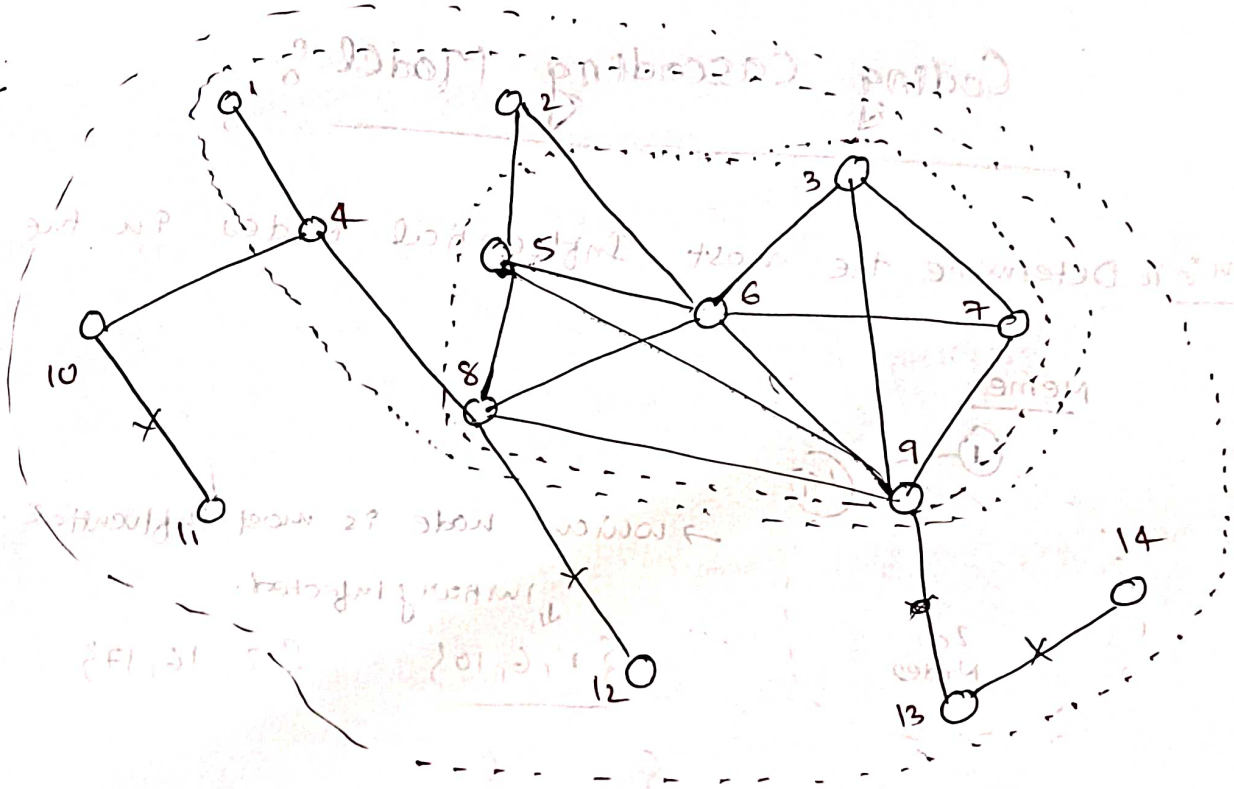


3-core.

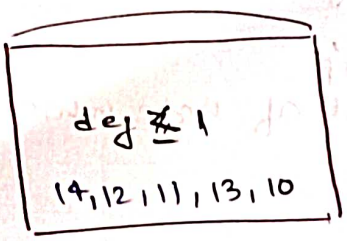
core of the Network

Note:- the last k^{th} core remove is the core of the network. (last bucket)

Example.

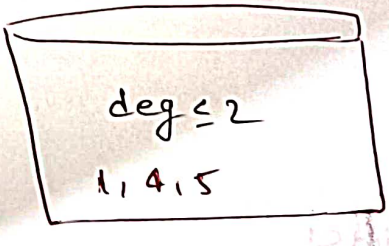


1 core



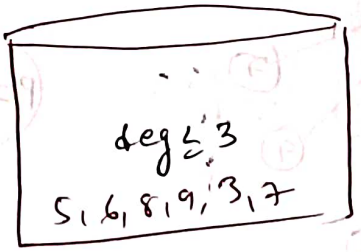
$= B_1$

2-core



$= B_2$

3-core



$= B_3$

1 core $= B_1 \cup B_2 \cup B_3$

2-core $= B_2 \cup B_3$

3-core $= B_3$

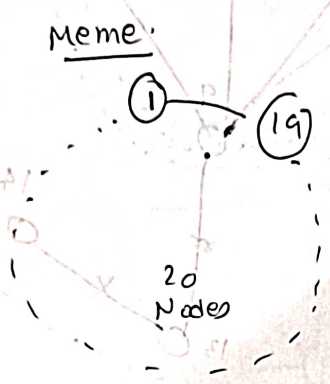
K core $= B_K \cup (B_{K+1}) \cup (B_{K+2}$

- - - -

$= \bigcup_{j \geq K} B(j)$

Coding Cascading Model's

Aim's To Determine the most Influential Nodes in the Net



→ which node is most influential.

↓ Initially infected.

{1, 6, 10} {2, 16, 17}

Infection | cascading model.

{1, 2, 3, 10, 6, 12} → those are the set of node which are i when the process stop

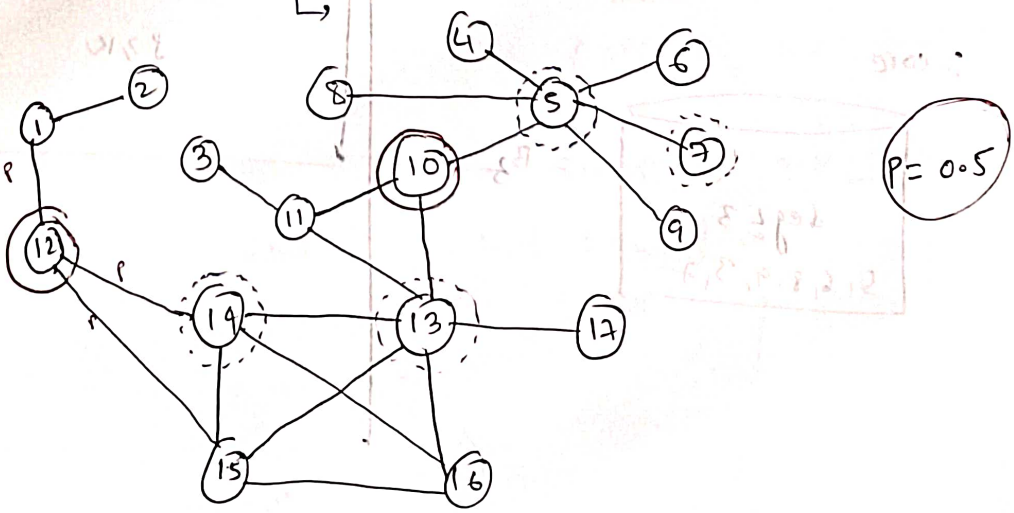
6 → node.

influential power of these set is 6.

Infection | cascades Model

Independent Cascade Model.

similar to SIR model



at t=0

let, say {12, 10} starting infection node. or seed node

at t=1

14, 5 get infected.

Now, 12, 10 will not infect further.

at t=2

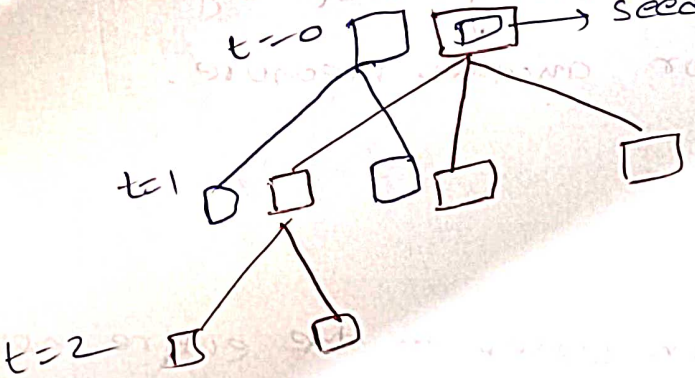
13, 7 get infected by 14, 5 respectively.

at t=3

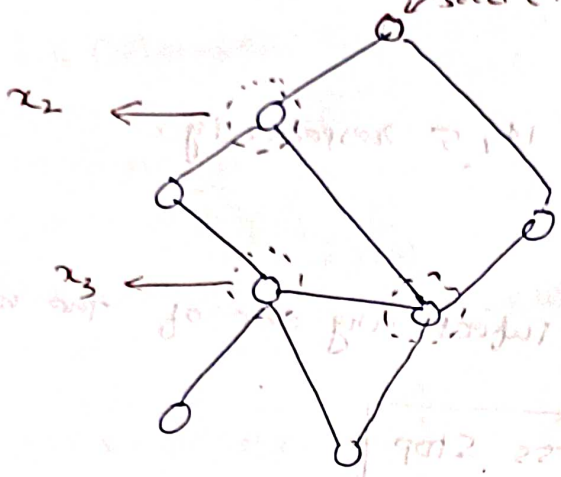
let say 13, 7 does not infect any one of the node.

and hence the process stop

So, the influential power = 6 = {12, 10, 14, 5, 13, 7}



Core Node in Cascading / Cascadeness / Influ



⇒ at the end which over x value is maximum is the mos influential in the network

⇒ This is very time consuming method.

⇒ can i say which over node having most degree is th most influential. or, another measure.

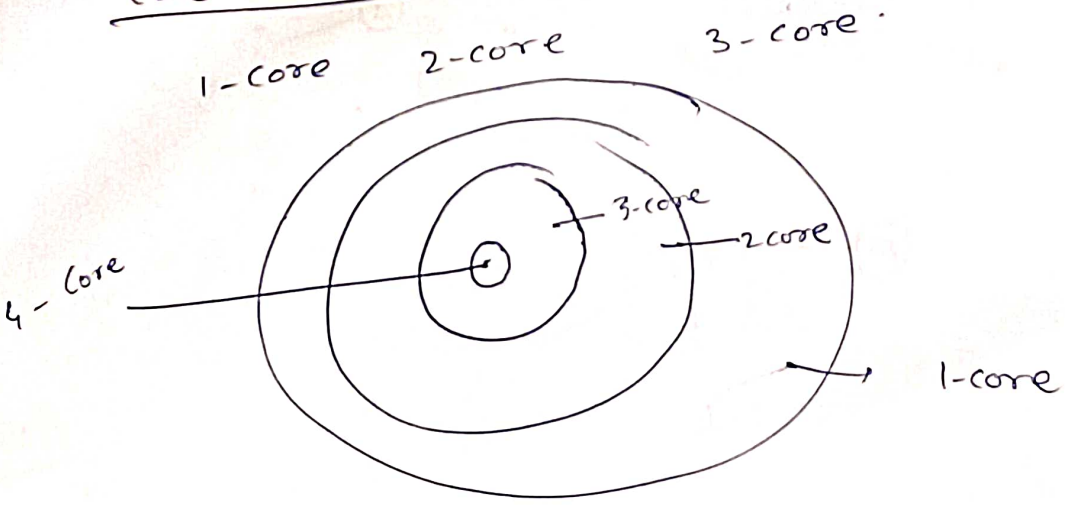
i) Degree

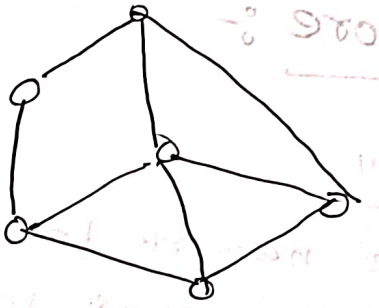
ii) Betweenness

iii) Closeness (how closer to the entire netwo

iv) Coreness (maximum)

core-number





descending order

Degree (desc)

A
B
C
.
.
Z

Betweenness (desc)

B
E
A
.
.
Z

closeness (desc)

coreness (desc)

highest value of coreness which best define which are the most influential node.

Influential - Power

B
C
Z
.
.
A

Pseudo core :-

Cascade capacity of a cell

→ If you start infecting the network from the shell o the seed node from the shell. what is the final amount of network that you have infected.

To make your meme go viral

→ It is not always necessary for you to go to the inner mos core and convince these people.

→ look at the some of the shell outer from core shell, they have actually the same cascade capacity than the core node. Instead of the core node. we found a lot more people here which are actually calle Pseudo core whom you can actually infect. so the your meme go viral.

— o —