Page No.
Date: / /

Week 1

random.randrange(1,10) - (1,10) int
random.random() → [0,1] between
random.randint(a,b) → [a,b] int

d.items() - list of tuples having key, value)

- Total graphs for n nodes - $2^{{}^nC_2}$
- logn steps from one point to other, 6 degrees of separation

Week 2 - Zachary karate - 34 nodes, 78 edges

Synonymy Graph
Web graph
Link prediction - fb

N/w datasets		Formats
Frnship n/w	(undirected)	CSV ← ~~edge~~ list, Adjlist
Road	(")	GML
Email	(directed)	Pajek net
Citation	(")	GraphML
Collaboration	(undirected)	GEXF - xml + gephi

1. ~~Degg~~ Degree Distribution
   nx.degree(h) = dictionary of {1: 16, 2: 5 ...}

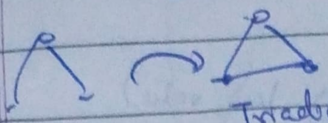
2. Density = $\frac{\text{edges present}}{\text{total possible edges}}$

3. Clustering coeff (node) = $\frac{\text{actual frnship b/w neighbour}}{\text{total possible b/w neighbr}}$

nlogn edges for ~~[illegible]~~ connected graph () $\frac{1}{n^2}$ [illegible]

Date: / /

__Week 3__ Granovetter's strenght of weak ties - job [illegible]



Triad

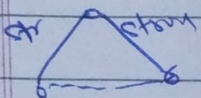
Triadic closure

Completely connected ⇒ clust coeff = 1

cc = 0 . If cc ↓ ⇒ suicide

Neighbourhood overlap = $\frac{\text{common frnds}}{\text{total frnds (dont count edge nodes)}}$

= local bridge - edge without triad, = weak tie



Strong Triadic closure prop (not weak tie)

Strong Tie ⇒ Talk more ⇒ not a local bridge ⇒ high neigbourh overlap

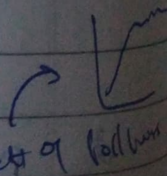
⇒ Embededness = no. of common frnds
(If ↑ emb ⇒ ↑ Trust)

High embdedness preffered
→ Yes ⇒ Trust
→ No bcz structural hole. monopoly (Brokerage

Closure & brokerage both are imp)

Facebook
- Maintained (passive engagement)
- oneway
- reciprocal

Twitter
Follow / Followed by ⇒ Friends (Strong Tie)



No. of DMs stay constant irrespect of followers

Communities
- more within & less across
- valid partitioning

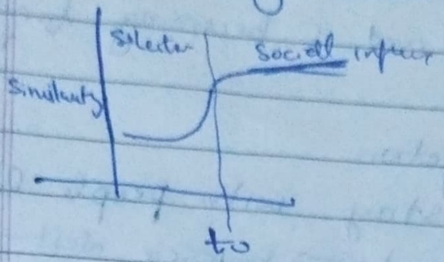
If you drop the highways ⇒ Valid Partitioning

Page No.
Date: / /

Week 4 — Selection vs Social influence

Wikipedia works? – Talk page in Background

Similarity measure = $\frac{\text{pgs edited by both}}{\text{Total by both}}$



If frndship across elder and child comms is less than half ⇒ Homophily (different grp have less frndship).

Homophily = 1 − $\frac{\text{Frndship across present}}{\text{expected no. of frnshp } \left(\frac{\text{Total edges}}{2}\right)}$

↓

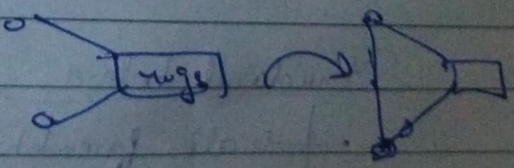
↑ses ⇒ more homophily

→ If −ve ⇒ Heterophily

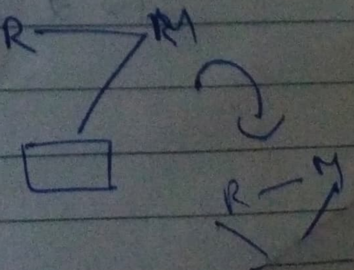
−ve ——— 0 ——— +ve

Diff people (actor, actress) | 0 ↓ half the times different | Similar people (teen, old)

Focal closure



Membership closure



Date : / /

$(1-p)^k$ = B and C not becoming frnds on having k common frnds

$1-(1-p)^k$ = B and C becoming frnds with k common

Week5 Spatial Segregation
- We prefer to stay with people like you
- Unhappy → move to → happier dest
  ↳ threshold

Schelling Model
similar ≥ 25% ⇒ $\frac{2}{08}$ = threshold
if 8 neighbrs

⇒ Spatial seg doesnot always induce cluster (community)

---

3 friendship = stable
1 friendship = stable → all 3 +ve
2 frnd = unstable → one +ve and 2-ve
3 hatred = convert → 1 +ve, 2-ve (stable)

Structural balance
- → all friendsh +ve
- → 2 clusters with + within (intra)
  − across (inter)

→ ~~bl~~ btwness of edge of max it connects 2 communities
→ Girvan Newman Algo - find btwness and drop the one with max

Page No.
Date: / /

Gephi - ↑ resolution ⇒ less communities
Week 4 & 5 later.

Week 6

Web Graph - Larry Page & Sergey Brin
2 ques (1) How to collect? → random walking
(2) How it solves prob. of who is imp? (Page Rank)
a) Gold coin Distribution (equal sharing)
b) Random walking " → one coin as you move
equal share ⇒ random dropping
2 aspect (1) Gain rich
(2) He not have more people to distribute

* Degree and PageRank are not correlated
independent * Google crawl webpages and do coin dropping

Week 7

3 Types of Follow → smoking (action)
→ product
→ info of interest

why do we follow → explicit benefit (getting free pass)
→ informational - (farmer change seeds)

→ Social reinforcement - many people
→ Diffusion

2 factors (1) Payoff
A B
pa → payoff (1-p)b
fract

(2) Choose large people

for A → $p \geq \frac{b}{b+a}$
A fraction

Soln
Diffusion Problem? People scared for new
(1) ↑se Payoff
(2) Key People focus

→ ~~gt~~ b/wness of edge ≥ more it connects & communities
→ Girvan Newman Algo - find b/wness and drop the one with max

Gephi - ↑ resolution ⇒ less communities

Week 4 & 5 later.

Week 6

Web Graph - Lary Page & Sergey Brin

2 ques (1) How to collect? → random walking

(2) How it solves prob. of who is IMP? (Page Rank)

a) Gold coin Distribution (equal sharing)

b) Random Walking " → one coin as you move

equal sharing = random dropping

2 aspect (1) Cousin rich

(2) He not have more people to distribute

* Degree and PageRank are not correlated

* teleport ~ Google crawl webpages and do coin dropping

Week 7

3 Types of Follow → smoking (action)

→ product

→ info of interest

why do we follow → explicit ~~benefit~~ (getting free pass)

→ informational - (farmer)

→ Social reinforcement - many people change decision

→ Diffusion

2 factors, (1) Payoff (2) Choose key people

A B

pa → payoff $(1-p)b$

frac

for A → $p \geq \frac{b}{b+a}$

A friend fract

Diffusion Problem? People scared for new

(1) ↑se Payoff

(2) Key People focus

Date : / /

If strong density n/w, difficult to diffuse ...

cluster density — D = if in a cluster atleast D fraction of friends are in cluster itself

10 frnds, 3 in cluster $\Rightarrow$ cluster density = 0.3

q = threshold for every node to adopt

Claim 1 — Idea cant cascade if any cluster have density $> 1-q$

Claim 2 — If idea couldnt cascade, $\Rightarrow$ there exist a cluster with density $> 1-q$

Incomplete cascade $\Rightarrow$ cluster of density $> 1-q$

Collective Action (for Revolting)

If threshold of people attained $\Rightarrow$ revolt

Week 8 — Better-rated lists = Better rated places

Hub $\longrightarrow$ Authority

Repeated Improvement - recursive

you point to good

good point to you

PageRank eg $\rightarrow$ converge

conservation of energy

Big vector + small vector = close to big vect

$(30, 70) + (5, 3) = (35, 73)$

eigen vector - linearly independent

$\frac{2^{n+1}}{2}$

Page No.

Date : / /

$A^k(v) = \underbrace{\lambda_1^k v_1}_{\text{Big vector}} + \underbrace{\lambda_2^k v_2}_{\text{small vector}} \quad (\lambda_1 > \lambda_2)$

$\Rightarrow$ ans is a direction of big vector

Page Rank by Matrix Mult

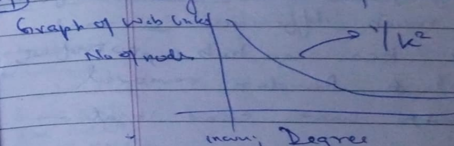
Markov mat = every col sum = 1

highest eigen val = 1

Week 9

Degree Dist^n of Real world N/w

(1)



Central limit theorem

sum of few random variables = Bell curve

(2) Song downloads

(3) telephone com (no of calls vs time

Powerlaw $\Rightarrow f(k) = \frac{1}{k^{2}} \longrightarrow$ b/w 2 to 3

Detect $\rightarrow f(k) = \frac{1}{k^a} \overset{\log\log}{\Longrightarrow} y = -\alpha \cdot x$

Date: / /

Reason of Powerlaw — Preferential Attachment
a new prsn attracted to someone having many ...

Node with high deg attract more — Rich get richer

Rich get Richer = Mathew effect = Preferential Attachment

Attack Survivability

→ Forced removal (selective) — removing high degree nodes

→ Random removal — removing nodes at random

Real world n/w

graph became disconnected fast

Random n/w

* But not much difference in random network bcz no hubs, no preference — edges are randomly added & no pref for addition of nodes

Selective removal & Random removal both have same situation

Barabasi Albert Model = Rich-getting richer phenomenon

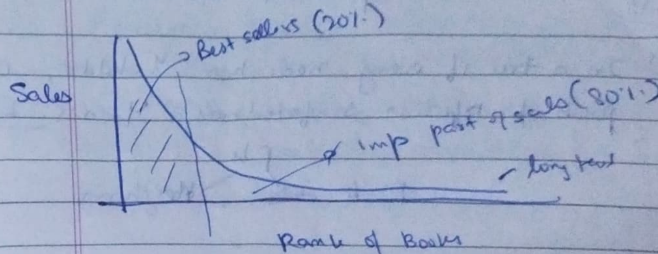
Erdos Renyi — Random Graph implementation

Page No.
Date: / /

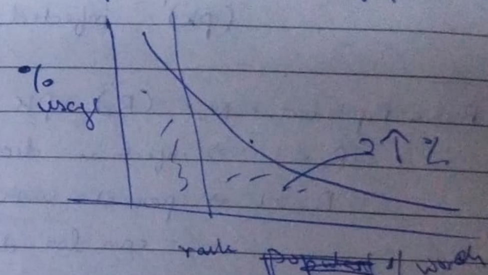
RGR adds more SE!

WEEK 10

when history is rewind, 1st rank (richest) is not the same



Eng → Zipf's law for popular words



long tail phenomena → popular words have less no. and not popular have more

Social Contagion

Spread of Idea vs Disease
choice · no choice
visible · Invisible

Factors → ① Pathogen
② Network

Page No.
Date : / /

Pathogen also derives the N/W

If ~~Fat~~ flue $\Rightarrow$ Dense N/W

branching model — If HIV $\Rightarrow$ Sparse N/W

At level 1 — In a tree if every node has 4 children, 0.5 prob of infection $\Rightarrow$ infected nodes $= 4 \times 0.5 = 2$

$= pk$

prob ~~dist~~ neighbours

At many levels — Branching model:

on ith level $\rightarrow$ k^i = total people

$(pk)^i$ = infected people

Basic Reproductive No. $(R_0) = pk$

$R_0 < 1 \Rightarrow$ infection dies ultimately with $p=1$

$R_0 > 1 \Rightarrow$ persists with +ve prob

- can also die but with low prob

knife edge property

R_0 0.9 — R_0 1.2 epidemic

SIR Model (Measles)

SIS (common cold)

SIR	SIS
Comes to an end	Forever
	but in one case if all become susceptible - it dies forever

Page No.
Date : / /

15.8

In SIR Model, theory of reproduction is not prevalent

— x —

Percolation Model - water pipe

Pg 193 — If $R_0 < 1$ $q^* = 0 = q_n$

(prob disease present at nth level)

$R_0 > 1 \Rightarrow q^* > 0$ (+ve prob)

$q_n = 1 - (1 - pq_{n-1})^k$

Week 11

2 people connected by at avg 6 hops

Milgrams Expt proved it

Watts - Strogatz Model

email - small world

Reason $\rightarrow$ Homophily + Weak ties are reason for SW

$\rightarrow$ Rewiring prob - for making outside/far off connections

$\rightarrow$ Decentralised Search - local ~~nw~~ search

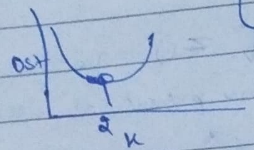
$\rightarrow$ Prob of putting a edge to far away node $= \dfrac{1}{d(u,v)^k}$ $k = 0,1,2 \ldots$

$\uparrow$ dist $\Rightarrow$ $\downarrow$ prob of edge b/w them

~~Mapping~~ Designing road /metro in city
Use small world n/w with p = rewiring prob

Ideal k = ?

$\boxed{k=2}$ ideally

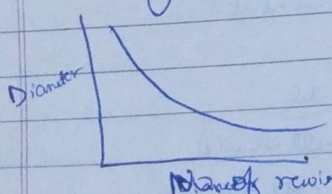


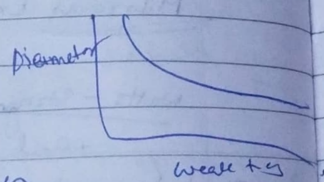
For 2D > k = 2
1D, k = 1

dist b/w random nodes in (n' nodes) small world
$\propto \log N$

Week 2

We can add random edges rather than rewiring to reduce the diameter





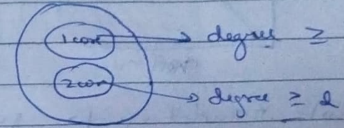
Myopic = Decentralised search

Node knows only dist of neighbour to target

As nodes ↑se, myopic search time is a log plot - logarithmic all n/ws

3 factors to get viral
1. Quality of meme
2. Structure of n/w
3. Key ~~pe~~ nodes

Right key nodes? ⟶ core nodes - densely connected



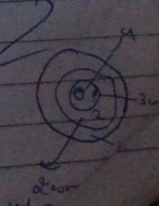
k - shell decomposition
- making core buckets

1-core = B1 ∪ B2 ∪ B3
2-core = B2 ∪ B3
3-core = B3

Coding cascade model
2 set of people - check Influential Power

Independent cascade → whether infection passed or not

Telling core nodes?
1. Calculate influential power of all taking one as seed - time consuming



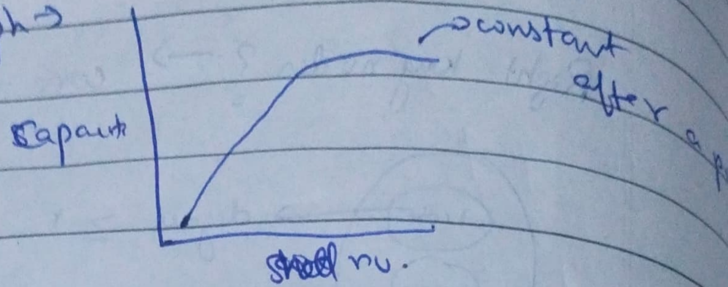
2. Try by calculate CORE no.
CORENESS == Degree

Date : / /

Pseudo-core

- Take a graph, calculate cascade capacity of each shell (influential power)
- Plot x = shell no., y = capacity

Graph →



→ It is not mandatory to go extreme inner core - rather can influence outer core too. (more in no.)

⇒ lot more people from outer core which can help = pseudo core

---

Granovetter - Weak Ties

Girvan - B/wness for bridge

Milgram - small world search algo

Watts Strogatz ⇒ Homo + weak tie

email = small world

collaboration local netwo