Degree distribution: $P(k)$

Path length: h

Clustering coefficient: C

Connected components: s

Avg. degree: $\bar{k} = \langle k \rangle = \frac{1}{N}\sum_{i=1}^{N} k_i = \frac{2E}{N}$

In directed networks we define an **in-degree** and **out-degree**. The (total) degree of a node is the sum of in- and out-degrees.

The **maximum number of edges** in an undirected graph on N nodes is

$$E_{\max} = \binom{N}{2} = \frac{N(N-1)}{2}$$

An **undirected** graph with the number of edges $E = E_{max}$ is called a **complete graph** and its **average degree is N-1**

Degree distribution $P(k)$: Probability that a randomly chosen node has degree k

N_k = # nodes with degree k

Normalized histogram:

$P(k) = N_k / N$ → plot

Distance (shortest path, geodesic) between a pair of nodes is defined as the **number of edges** along the **shortest path** connecting the nodes

- *If the two nodes are **not connected**, the distance is usually defined as **infinite (or zero)**

In **directed graphs**, paths need to follow the **direction of the arrows**

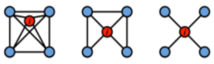
- Consequence: Distance is **not symmetric**: $h_{B,C} \neq h_{C,B}$

Diameter: The **maximum (shortest path) distance** between **any pair** of nodes in a graph

Clustering coefficient (for undirected graphs)

- How connected are i's neighbors to each other?
- Node i with degree k_i
- $C_i \in [0,1]$
- $C_i = \frac{2e_i}{k_i(k_i-1)}$ where e_i is the number of **edges between the neighbors of node i**

Note $k_i(k_i-1)$ ... max number of edges between k_i neighbors



$C_i = 1$ $C_i = 1/2$ $C_i = 0$

Clustering coefficient is undefined (or defined to be 0) for nodes with degree 0 or 1

Average clustering coefficient: $C = \frac{1}{N}\sum_{i}^{N} C_i$

Bridge edge: If we erase the **edge**, the graph becomes disconnected

Articulation node: If we erase the **node**, the graph becomes disconnected

G_{np}: undirected graph on n nodes where each edge (u,v) appears i.i.d. with probability p

Largest component = Giant component

21:06:05

Triadic closure = High clustering coefficient

How to find connected components:

- Start from random node and perform Breadth First Search (BFS)
- Label the nodes that BFS visits
- If all nodes are visited, the network is connected
- Otherwise find an unvisited node and repeat BFS

Bridge describes an edge's role where deleting this edge would **cause its corresponding vertices falling in different components**;

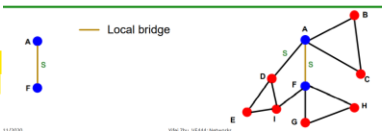
An edge is a **local bridge** if its **end points** have **no friends in common**

- (Alternatively, distance perspective?)
- **Span** of a local bridge: the **distance its endpoints** would be from each other if the edge were deleted

- If a node A has edges to nodes B and C, then the B-C edge is especially likely to form if A's edges to B and C are both strong ties.
- More formally, as Granovetter suggested:

A node A violates the **Strong Triadic Closure property** if it has strong ties to two other nodes B and C, and there is no edge at all(either a strong and weak tie) between B and C.

if a node A in a network satisfies the STC property and is involved in at least two strong ties, then any local bridge it is involved in must be a weak tie

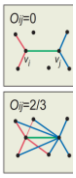


- **Edge overlap:**

$$O_{ij} = \frac{|(N(i) \cap N(j)) \setminus \{i,j\}|}{|(N(i) \cup N(j)) \setminus \{i,j\}|}$$

- $N(i)$ ... the set of neighbors of node i

- **Note: Overlap = 0** when an edge is a **local bridge**



Probabilistic relational classifier

- **Repeat** for each node i and label c

$$P(Y_i = c) = \frac{1}{|N\_i|}\sum_{(i,j)\in E} W(i,j)P(Y_j = c)$$

- $W(i,j)$ is the edge strength from i to j
- N_i is the number of neighbors of i

Structural balance property: For every set of three nodes, if we consider the three edges connecting them, **either all three of these edges are labeled +, or else exactly one of them is labeled +**.

Homophily test: If the fraction of **cross-attributes edges** is **significantly less than 2pq**, then there is evidenced for homophily.

Balance theorem: If a labeled **complete graph** is balanced, then **either all pairs of nodes are friends**, or else the nodes can be **divided into two groups**, X and Y , such that **every pair of nodes in X like** each other, **every pair of nodes in Y like each other**, and **everyone in X is the enemy of everyone in Y** .

	Stag	Hare	q-mix
Stag	4,4	0,3	4q, 4q+3(1-q)
Hare	3,0	3,3	3q+3(1-q), 3(1-q)
p-mix	4p+3(1-p), 4p	3(1-p), 3p+3(1-p)	计算q用player1的等式

计算p用player2的等式

A choice of strategies – one by each player – is Pareto-optimal if there is no other choice of strategies in which **all players receive payoffs at least as high**, and **at least one player receives a strictly higher** payoff.

Pareto/social optimal 和是否为equilibrium/dominant无关

Social optimality: A choice of strategies – one by each player – is a social welfare maximizer (or socially optimal) if it **maximizes the sum of the player's payoffs**

Suppose n numbers are drawn independently from the uniform distribution on the interval $[0,1]$ and then sorted from smallest to largest. The expected value of the number in the k^{th} position on this list is $\frac{k}{n+1}$.

Now, if the seller runs a second-price auction, and the bidders follow their dominant strategies and bid truthfully, the seller's expected revenue will be the expectation of the second-highest value. Since this will be the value in position $n-1$ in the sorted order of the n random values from smallest to largest, the expected value is $(n-1)/(n+1)$, by the formula just described. On the other hand, if the seller runs a first-price auction, then in equilibrium we expect the winning bidder to submit a bid that is $(n-1)/n$ times her true value. Her true value has an expectation of $n/(n+1)$ (since it is the largest of n numbers drawn independently from the unit interval), and so the seller's expected revenue is

$$\left(\frac{n-1}{n}\right)\left(\frac{n}{n+1}\right) = \frac{n-1}{n+1}.$$

The two auctions provide exactly the same expected revenue to the seller!

Price of Anarchy(POA): the ratio between the system performance with **strategic players** and the **best-possible** system performance

2020/12/4 21:56:35

Perfect Matching: When there are an equal number of nodes on each side of a bipartite graph, a perfect matching is an assignment of nodes on the left to nodes on the right, in such a way that

- each node is connected by an edge to the node it is assigned to, and
- no two nodes on the left are assigned to the same node on the right.

1:18:03

Revenue equivalence theorem: given certain conditions, any mechanism that results in the same outcomes (i.e. allocates items to the same bidders) also has the same expected revenue.

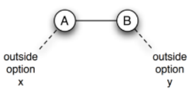
Stationary strategy: the same split in every period in which they are scheduled to propose

- The split (a1,b1) that A will offer whenever he is scheduled to propose a split;
- the split (a2,b2) that B will offer whenever she is scheduled to propose a split; and
- reservation amounts a' and b', constituting the minimum offers that A and B respectively will accept from the other.

Nash Bargaining Solution: When A and B negotiate over splitting a dollar, with an outside option of x for A and an outside option of y for B (and x + y ≤ 1), the Nash bargaining outcome is

- $x + \frac{1}{2}s = \frac{x+1-y}{2}$ to A, and
- $y + \frac{1}{2}s = \frac{y+1-x}{2}$ to B

Surplus: s = 1-x-y



A will offer B the least he can get B to accept his offer

- b1=b'

B will offer A the least he can get A to accept his offer:

- a2=a'

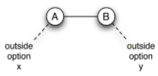
B's reservation amount is the indifference amount between accepting A's offer and rejecting A's offer

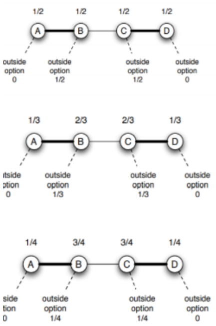
- Accept: b1
- Reject: py+(1-p)b2
- Indifference: b1=py+(1-p)b2 (1)

Similarly, a2=px+(1-p)a1 (2)

Solving (1), (2),

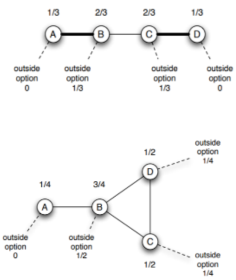
a1 = $\frac{(1-p)x+1-y}{2-p}$, b1= 1-a1= $\frac{y+(1-p)(1-x)}{2-p}$





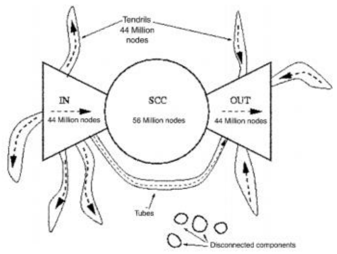
In any network with a stable outcome, there is a balanced outcome
Balance can be used as a way to refine stable outcomes to align with experimental results.
Nash bargaining solution: captures the weak power advantages, and some fairness

Computational issue



Nash bargaining solution: captures the weak power advantages, and subtle differences because of the network structure

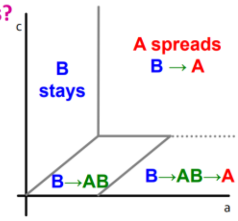
Cooperative game theory: how a collection of players will divide up the value arising from activity

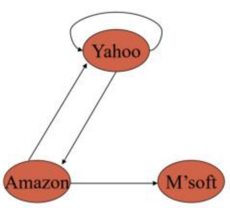


In: new pages
Out: cooperate pages that links to private pages.
Tendrils: (1)nodes that reachable from IN but cannot reach the SCC
(2)nodes that reach OUT but cannot be reached from SCC
Tubes: Both (1) (2)

***B* is the default throughout the network until new/better *A* comes along. What happens?**

- **Infiltration:** If **B** is too compatible then people will take on both and then drop the worse one **(B)**
- **Direct conquest:** If **A** makes itself not compatible – people on the border must choose. They pick the better one **(A)**
- **Buffer zone:** If you choose an optimal level then you keep a static "buffer" between **A** and **B**





$$0.8\begin{bmatrix} 1/2 & 1/2 & 0 \\ 1/2 & 0 & 0 \\ 0 & 1/2 & 0 \end{bmatrix} + 0.2\begin{bmatrix} 1/3 & 1/3 & 1/3 \\ 1/3 & 1/3 & 1/3 \\ 1/3 & 1/3 & 1/3 \end{bmatrix}$$

$$\begin{array}{c} y \\ a \\ m \end{array}\begin{bmatrix} 7/15 & 7/15 & 1/15 \\ 7/15 & 1/15 & 1/15 \\ 1/15 & 7/15 & 1/15 \end{bmatrix}$$

which cannot lead to convergence

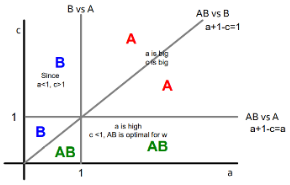
Non-stochastic!

$$\begin{array}{c} y \\ a \\ m \end{array} = \begin{array}{cc} 1/3 & 1/3 \\ 1/3 & 0.2 \\ 1/3 & 0.2 \end{array} \quad \cdots \quad \begin{array}{c} 0 \\ 0 \\ 0 \end{array}$$

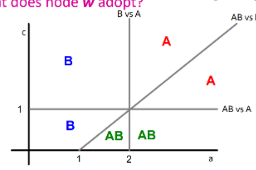
instead you have to 0.8[ ]+[ ]

fixing teleport

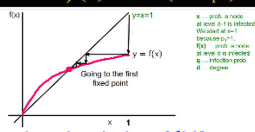
Infinite path, start with Bs
Payoffs for *w*: A:a, B:1, AB:a+1-c
What does node ***w*** adopt?



Same reward structure as before but now payoffs for *w* change: A:a, B:1+1, AB:a+1-c
Notice: Now also AB spreads
What does node ***w*** adopt?



Fixed Point: $f(x) = 1 - (1 - qx)^d$

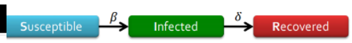


What do we know about the shape of $f(x)$?

- $f(0) = 0$
- $f(1) = 1 - (1 - q)^d < 1$
- $f'(x) = q \cdot d(1 - qx)^{d-1}$
- $f'(0) = q \cdot d$

f'(x) is monotone: If g'(y)>0 for all y then g(y) is monotone. In our case, 0≤x,q≤1, d>1 so f'(x)>0, so f(x) is monotone. f'(x) non-increasing since term $(1-qx)^{d-1}$ in f'(x) is decreasing as x decreases.

$f'(x)$ is monotone non-increasing on [0,1]!



- Each time unit, any infected individual contact k other individuals
- The probability of infection on contact: π
- Per unit of time, contact kS/N susceptible individuals
- Infect kπS/N individuals
- **Transmission rate: β= πk**

$$R_0 = \frac{\beta}{\delta} = \frac{k\pi}{\delta}$$

R_0: Expected number of secondary cases caused by a single primary infected individual

- **Reduce the infectious period $\frac{1}{\delta}$** -> therapeutics
- **Reduce the β**
  - **Reduce π:** sewage systems, hand-washing, air filters, and so on)
    - mitigating the number of contagious particles that are exchanged among individuals.
  - **Reduce k:** quarantines, social distancing and travel, restrictions
    - Reduce the contact rate

Estimate of R_0:

- **Estimating *q*:** Given an infected node count the proportion of its neighbors subsequently infected and average
- Then: $R_0 = q * d * \frac{avg(d_i^2)}{(avg\ d_i)^2}$

d ... avg degree
d_i ...degree of
Correction factor due to skew degree distribution of the net

Empirical R_0:

$$\mathcal{L} = \sum_{u \in V} \sum_{v \in N_R(u)} -\log(P(v|\mathbf{z}_u))$$

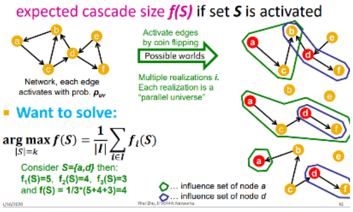
Intuition: Optimize embeddings to maximize likelihood of random walk co-occurrences

Parameterize $P(v|\mathbf{z}_u)$ using softmax:

$$P(v|\mathbf{z}_u) = \frac{\exp(\mathbf{z}_u^\top \mathbf{z}_v)}{\sum_{n \in V} \exp(\mathbf{z}_u^\top \mathbf{z}_n)}$$

Why softmax?
We want node *v* to most similar to nod (out of all nodes *n*).
Intuition: $\sum_i \exp(x_i)$ ... $\max \exp(x_i)$

- **Find most influential set *S* of size *k*:** largest expected cascade size ***f(S)*** if set ***S*** is activated



- **Want to solve:**

$$\arg\max_{|S|=k} f(S) = \frac{1}{|I|}\sum_{i \in I} f_i(S)$$

Consider $S=\{a,d\}$ then: $f_1(S)=5$, $f_2(S)=4$, $f_3(S)=3$ and $f(S) = 1/3*(5+4+3)=4$

... influence set of node a
... influence set of node d

- **Model dynamics are:**

$$\frac{dS}{dt} = -\frac{\beta SI}{N} \qquad \frac{dR}{dt} = \delta I$$

$$\frac{dI}{dt} = \beta SI - \delta I$$