UNIVERSIDAD POLITÉCNICA DE YUCATÁN

SOCIAL NETWORK ANALYSIS

HOMEWORK 3

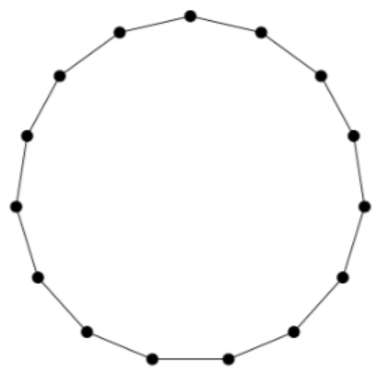
Professor:
Ing. Didier Gamboa

Centrality Measures and Network Models

Student:
Walter VIVES
DATA 8A

june 15, 2021

1 A network consists of n nodes in a ring, where n is odd. All the nodes have the same closeness centrality. What is it, as a function of n?



In a connected graph, the normalization closeness centrality of a node is the average length of the shortest path between the node and all other nodes in the graph.

$$c(x) = \frac{N}{\sum_y d(y,x)} \tag{1}$$

N is the number of nodes.
Where *d(y,x)* is the distance between vertices x and y.
Let us assume that there is a n nodes in a ring, where n is odd.

- n = 2m + 1
- m = 0, 1, 2, 3, 4, 5, 6, ...

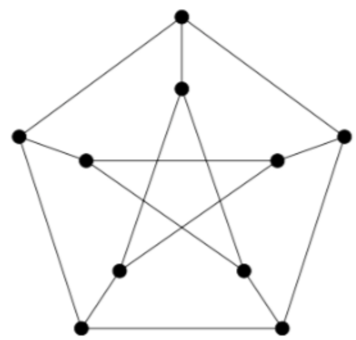
$$
\begin{aligned}
c(x) &= \frac{2m+1}{1+2+...+m+...} \\
c(x) &= \frac{2m+1}{1+2+...+(m-1)+m+m+(m-1)+...} \\
c(x) &= \frac{2m+1}{\frac{m(m+1)}{2}+\frac{m(m+1)}{2}} \\
c(x) &= \frac{2m+1}{\frac{2m(m+1)}{2}} \\
c(x) &= \frac{2m+1}{m(m+1)} \qquad (2)
\end{aligned}
$$

$$
\begin{aligned}
n &= 2m+1 \\
n-1 &= 2m \\
m &= \frac{n-1}{2} \\
m+1 = \frac{n-1}{2}+1 &= \frac{n-1+2}{2} \\
&= \frac{n+1}{2} \qquad (3)
\end{aligned}
$$

$$
\begin{aligned}
c(x) &= \frac{n}{(\frac{n-1}{2})\frac{n+1}{2}} \\
c(x) &= \frac{n}{\frac{n^2-1}{4}} \\
c(x) &= (\frac{n}{1})(\frac{4}{n^2-1}) \\
c(x) &= \frac{4n}{n^2-1} \qquad (4)
\end{aligned}
$$

Therefore, it's a function of n.

2 Calculate the closeness centrality of each node of the nodes in this network:



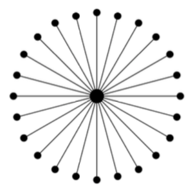
Given the formula:

$$\frac{N-1}{\sum d(j,i)} \tag{5}$$

Table 1: Closeness Centrality of each Node

-	1	2	3	4	5	6	7	8	9	10	$\sum d(j,i)$	*N-1*/$\sum d(j,i)$
1	0	1	2	2	1	1	2	2	2	2	15	9/15
2	1	0	1	2	2	2	1	2	2	2	15	9/15
3	2	1	0	1	2	2	2	1	2	2	15	9/15
4	2	2	1	0	1	2	2	2	1	2	15	9/15
5	1	2	2	1	0	2	2	2	2	1	15	9/15
6	1	2	2	2	2	0	2	1	1	2	15	9/15
7	2	1	2	2	2	2	0	2	1	1	15	9/15
8	2	2	1	2	2	1	2	0	2	1	15	9/15
9	2	2	2	1	2	1	1	2	0	2	15	9/15
10	2	2	2	2	1	2	1	1	2	0	15	9/15

3 A “star graph” consists of a single central node and n − 1 other nodes connected to it. What is the (unnormalized) betweenness centrality of the central node as a function of n?



In order to calculate Between Centrality:

$$g(v) = \sum_{s \neq t \neq v} \frac{\sigma_{st}(v)}{\sigma_{st}} \tag{6}$$

$\sigma_{st}(v) =$ Total number of shortest path through "v".
$\sigma_{st} =$ Total number of shortest path between node u and w.

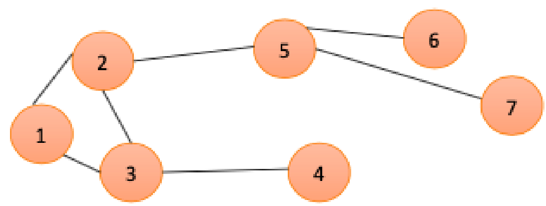
$$g(centralNode) = 23 + 22 + 21 + 20 + 19 + 18 + 17 + 16 + 15 + 14 + 13 + 12 + 11 + 10 + 9 + 8 + 7 + 6 + 5 + 4 + 3 + 2 + 1$$

$$g(centralNode) = 276$$

It is the same using the formula:

$$\frac{(n-1)(n-2)}{2}$$
$$\frac{(25-1)(25-2)}{2} = 276$$
Where n is the number of nodes. (7)

4 Consider the following network. For each question, in case of a tie, answer with all the tied top nodes.



a **Which node has the highest degree centrality?**

The nodes 2,3, 4 have the highest degree centrality given that they have 3 edges.

b **Which node has the highest betweenness centrality?**

The node 2 has the highest degree centrality since if we removed the network will be disconnected.

c **Which node has the highest closeness centrality?**

Closeness centrality indicates how close a node is to all other nodes in the network. It is calculated as the average of the shortest path length from the node to every other node in the network. [1]

Therefore, the node 2 has the highest degree closeness centrality.

5 Consider a social network where a connection represents a sexual relationship.

Read the report by Liljeros et al. (2001) about a study of such a network based on a sample of 4781 Swedes. (If you do not have access to the journal through your institution, you can download a preprint of the paper at https://arxiv.org/abs/cond-mat/0106507

What is the maximum degree in this network? What does it mean? If you consider the subnetworks with nodes corresponding to males and females, respectively, do they have the same degree distribution? Why or why not?

The maximum degree is the degree of the vertex with the greatest number of edges incident to it. In this case, X is the maximum number of sexual partner that a person has had during the twelve months prior to the survey. It can be reduced the propagation of sexually-transmitted diseases if it noticed a behavior in the network.

No, they will have different degree distribution since in the study, males have more connections than females therefore, it could possible that exist male nodes that have a a higher degree than others.

6 Consider the Random Graph Model. Show that the binomial distribution of the degree given by

$$P(k) = \binom{n-1}{k} p^k (1-p)^{n-1-k} \tag{8}$$

follows a Poisson Distribution given by,

$$P_k = e^{-<k>} \frac{<k>^k}{k!} \tag{9}$$

under the condition k $\ll$ N.

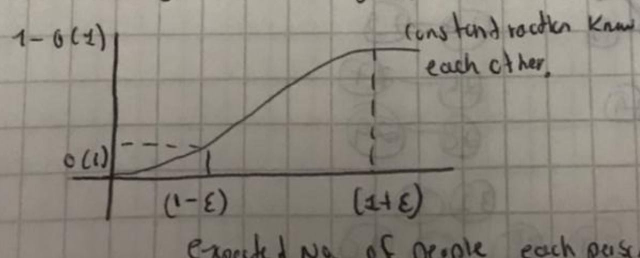
In the random graph model

Large graphs appear in many context such as the world wide web; the internet, social network journal citations and other place perhaps the most basic such model is the G(n,p) model of a random graph. In the study of the G(n,p)

These n is the no of vertex of the graph & p is the edge p.d. For each pair of distribution v & w, p is the probability that the edge (v, w) present of each edge is statistically independent

For small p with $p = d/n$, $d < 1$, each connected component in the graph is small, for $d < 1$, there is giant component consisting of.

The phase transition at the threshold $d = 1$ from every small $O(n)$ size size component to a ~~ga~~ $\Omega(n)$ sized component is illustrated.
If the expected number of friends each person has is more than one, than a giant component will be present a fraction of all the people.
On the other hand, if in expectation each [redacted] has less than one friend, the largest component is a small fraction of the whole happens. arbitrary b/w d slightly less than



One of the simple quantities to observe in a real graph is the no. of vertices of given degree, called the vertex degree distribution.

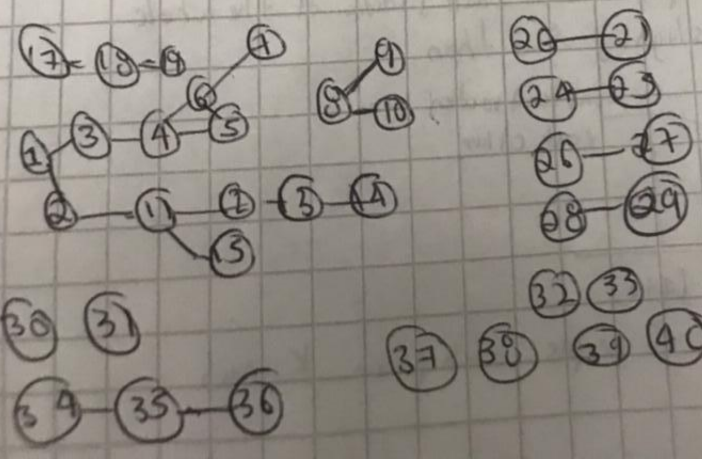
Such p is the probability of an edge being present, the expected degree of a vertex is $d = pn$.

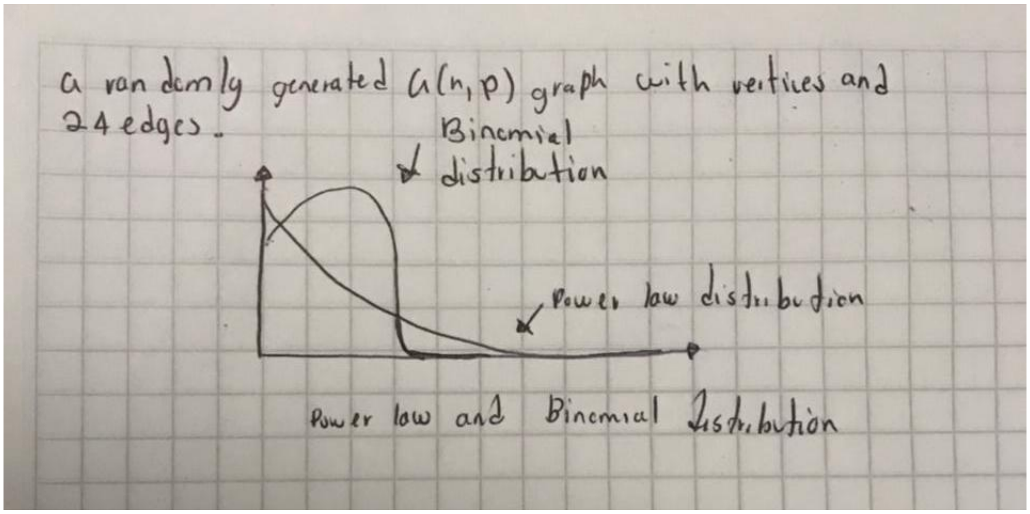
Prob (vertex has degree k)

$$= \binom{n-1}{k} p^k (1-p)^{n-k-1} = \binom{n}{k} p^k (1-p)^{n-k}$$

$\binom{n-1}{k}$ is the no. of ways of choosing k edge, out of the possible $(n-1)$ edge and $p^k (1-p)^{n-k-1}$ is the probability that the k selected edge are present and remaining $(n-k-1)$ are not.

Although the $h(n,p)$ model is important mathematically more complex moder are needed





7 Sociologist estimate that a typical person knows about 1,000 individuals on a first name basis. Considering that the society can be modeled with a Random Graph Network of N = 7 $\times$ 10^9 of people, obtain the maximum and minimum degree (expression and numerical value).

For a random network with N $\approx$ 7 x 10^9 nodes, A random society is anticipated to have the most linked individual (the biggest degree node): kmax = 1,185 acquaintances.

The degree of the least connected individual is kmin = 816. The dispersion of a random network is k = $\langle k \rangle$ 1/2 , which for $\langle k \rangle$ = 1,000 is k = 31.62

Random network dispersion is k = <$\langle k \rangle$ 1/2, which is k = 31.62 for $\langle k \rangle$= 1.000. This indicates that a typical person's number of friends is between $\langle k \rangle$ $\pm$k, a relatively between 968 and 1032.

All people should have a comparable number of friends together in a random society. Therefore, we have no outliers if people are connected by chance: Nobody is really popular and just a few buddies remain behind. This surprising finding is the result of an essential random network characteristic. This forecast clashes flagrantly with reality.

8 Show that in a Scale Free Network, the maximum degree $k_m axand$

$$kmax = k_{min} N^{\frac{1}{\gamma-1}} \tag{10}$$

The maximum degree k_{max} and k_{min} follows the relation given by

$$K_{max} = K_{min} N^{\frac{1}{\gamma - 1}}$$

Where N = The no. of nodes in the network

γ = The degree exponent

- Networks whose degree distribution follows a power law are said to be scale-free networks
- The power law degree distribution can be defined in both discrete and continuous formalisms.

→ Discrete Formalism

$$P_K = C K^{-\gamma}$$

[C is a constant
It's value is $1/\xi(\gamma)$]

$$\sum_{K=1}^{\infty} P_K = 1$$

[P_K diverges for $K=0$ (isolated nodes) for such nodes. we need to specify]

$$C \sum_{K=1}^{\infty} K^{-\gamma} = 1$$

[The scale-free property is independent of the formalism used to describe the degree distribution.

$$\Rightarrow C = \frac{1}{\sum_{K=1}^{\infty} K^{-\gamma}} = \frac{1}{\xi(\gamma)}$$

Therefore, $\boxed{P_K = \frac{K^{-\gamma}}{\xi(\gamma)}}$

$$\xi(\gamma) \sum_{K=1}^{\infty} K^{-\gamma}$$

Where $\xi(\gamma)$ = The riemann-zeta function

Now,

$$P(K) = C K^{-\gamma}$$

$$\int_{K_{min}}^{\infty} P(K) \, dK = 1$$

$$\Rightarrow \quad c = \frac{1}{\int_{k_{min}}^{\infty} k^{-\gamma} dk} = (\gamma-1) k_{min}^{\gamma-1}$$

$\therefore$

$$\boxed{P(k) = (\gamma-1) k_{min}^{\gamma-1} k^{-\gamma}}$$

here,

k_{min} is the smallest degree for which the power law for the discrete formalism holds.

~~Note~~ P_k encountered in the probability that a randomly chosen nodes has degree k. In contrast, only the integral of $p(k)$ encountered in the continuous formalism has a physical interpretation

$$\int_{k_1}^{k_2} P(k)\, dk$$

Probability that a randomly chosen nodes has degree between k_1 and k_2

$$\int_{k_{max}}^{\infty} P_k\, dk = \frac{1}{N} \Rightarrow \int_{k_{max}}^{\infty} c k^{-\gamma} dk = \frac{1}{N}$$

$$c \left\{ \frac{k^{-\gamma+1}}{-\gamma+1} \right\}_{k_{max}}^{\infty} = \frac{1}{N} \Rightarrow \frac{c}{\gamma-1} \left\{ \frac{1}{k_{max}^{\gamma-1}} \right\} = \frac{1}{N}$$

$$k_{max}^{\gamma-1} = \frac{Nc}{\gamma-1} = N k_{min}^{\gamma-1}$$

$$\Rightarrow \boxed{k_{max} = k_{min} N^{\frac{1}{\gamma-1}}} \quad \therefore \text{ hence proved } \checkmark$$

References

[1] Golbeck, J. (2013). Analyzing the social web. Newnes.