Small World Network

Neeraj Nagi(03CS1005) and Harish Daiya(03CS1004)

Lecture by Prof. Niloy Ganguly

1 Small World Network

1.1 Small world phenomenon

According to small world phenomenon everyone might not be directly connected to each other, but can be reached through a few steps of social acquaintences. This concept was popularized by the famous 1967 small world experiment by social psychologist Stanely Milgram, according to which any two randomly selected U.S. citizens can be connected by 6 acquaintences.Which popularized the famous phrase “Six degree of separation”. The networks which show this kind of behaviour are known as small world networks.

Six degrees of separation refers to the idea that, if a person is one "step" away from each person he or she knows and two "steps" away from each person who is known by one of the people he or she knows, then everyone is no more than six "steps" away from each person on Earth. Several studies, such as Milgram's small world experiment have been conducted to empirically measure this connectedness. While the exact number of links between people differs depending on the population measured, it is generally found to be relatively small. Hence, six degrees of separation is somewhat synonymous with the idea of the "small world" phenomenon.

1.2 Properties of small world network

Small world networks are characterized by abundance of clique or subgraphs which are few edges shy of being cliques. Which requires high cluster coefficient so that any two nodes within a subgraph are connected in most cases.

Also any two nodes in the network should be connected by atleast one short path, so the mean shortest path distance should be small. Examples include road maps, food chains, electric power grids, metabolite processing networks, neural networks, telephone call graphs and social influence networks.

1.3 Testing a network for small world effect

For an undirected network of n nodes the mean geodesic path between vertices can be expressed as

- Considering self loops
  L= $\frac{1}{\frac{1}{2}n(n+1)}\sum_{i\geq j}$ d_{ij}
- Assuming whole graph is connected
  L= $\frac{1}{\frac{1}{2}n(n-1)}\sum_{i>j}$ d_{ij}

where L is the mean geodesic path and d_{ij} is the geodesic distance from vertex i to vertex j. Most of the networks observed in real life(e.g. film actors, company directors, emails, internet, electronic circuits) have L $\leq$6.

The shortest distance between vertices can be found using standard BFS for unweighted graphs. Since the time required for performing BFS from a single node is O(m), the total time required to find L would be O(mn). For networks, with more than one component, we can find L for each connected component, and take the harmonic mean of these L's to get the overall mean geodesic path length.

2 Clustering

2.1 Clustering Coefficient (p)

The clustering coefficient of a particular node in a network is defined as the relation between the total number of connections among its neighbours to the total number of connections possible between its neighbours. Its the measure of whether or not the node is connected to a dense neighbourhood. High clustering coefficient indicate that the network follows small world phenomenon.

For a vertex i, the clustering index C_i is given by,
$C_i = \frac{|\{e_{jk}\}|}{K_i(k_i-1)}$
where k_i is the degree of vertex i.
e_{jk} is number of edges among the neighbours of vertex i.

- It can also be calculated as the ratio of no. of triangles including vertex i to no. of triplets including vertex i.

For nework of infinite nodes and constant degree clustering coefficient tends to zero

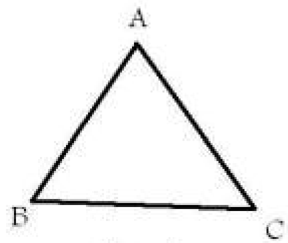

Fig - 1

- Clustering Coeff. $C^{(1)}$=3*no. of traingles/no. of triplets(fig 1.)
  so, $C^{(1)}=\frac{ABE*3}{EAB+ABE+BEA}=1$

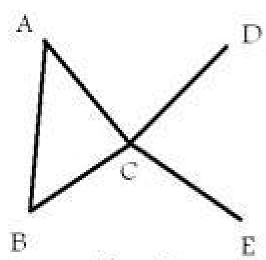

Fig - 2

- In fig. 2, $C^{(1)}= \frac{3*ABE}{3+ABC+ABD+EBC+EBD+CBD}$
  $= 3/8$

Global Clustering Coefficient ($C^{(2)}$) Clustering coefficient of a graph is taken as the average of all the clustering coefficient of all its nodes.

- $C^{(2)}= \frac{1}{n}\sum C_i$
- For previous example
  $C^{(2)}=\frac{1}{5}(1+1+\frac{2}{8}+0+0)= 9/20$

3 Assortativity

Out of several different possible explanation of the formation of communities in the social network, assortativity is most prominent. It states "like goes with like".

3.1 Example

Consider the following men-women marriage relation depending on their race. Here(i,j) element of the matrix gives the fraction of men of i^{th} race marry women of j^{th} race.

	black	hispanic	white	other
black	0.256	.016	.035	.013
hispanic	.012	.157	.058	.019
white	.013	.023	.306	.035
others	.005	.007	.024	.016

Diagonal elements represent the fraction of couples in partnership with members of their own group and off-diagonal those in partnership with members of other group. Inspection of the matrix shows that matrix has consideably more weight along the its diagonal than off it indicating that assortative mixing does take place. The amount of assortative mixing in a network can be quantified by measuring how much of the weight in the mixing matrix falls on the diagonal and how much off it. Let us define $e_{ij}$to be the fraction of all edges in a network that joins the vertex of type i with type j. According to the matrix, we can say that index i represents man and j represents female. Which makes e_{ij}asymmetric. The matrix should satisfy the sum $\sum_{ij} e_{ij}=1$, $\sum_j e_{ij}= a_i$ and $\sum_i e_{ij}= b_i$ where a_i and b_i are the fraction of each type of end that is attached to vertices of type i. Now we define a quantitative measure r of the level of assortative mixing in the network,

$$r = \frac{\sum_i e_{ii} - \sum_i a_i b_i}{1 - \sum_i a_i b_i}$$

It takes the value 1 for the perfectly assortative network, since in that case, the entire weight on the matrix lies along its diagonal. Conversely, if there is no assortative mixing at all, then r becomes quite low. Networks can also be disassimilative: vertices may associate preferentially with others of different types - the opposite attracts phenomenon. In that case, r becomes negative.

References

[1] ”Assortative mixing in networks”, M. E. J Newmann, Phys. Rev. Lett. 89, 20871 (2002)

[2] Milgram, S. (1967). The small-world problem. Psychology Today 1, 61-67.

[3] Travers, J., & Milgram, S. (1969) . An experimental study of the small world problem. Sociometry 32, 425-443.

[4] Watts, D.J. and Strogatz, S.H. “Collective dynamics of “small worlds networks”.