Online Resourse 1

immediate

March 1, 2019

1 Derivation of Neighborhood Proximity

To derive the measure neighborhood proximity, let us consider two nodes u and v, with $N_{uv} \neq \varnothing$, between which the closeness is to be determined. We define our measure called *neighborhood proximity* of u and v, denoted as $\rho(u, v)$, in such a way that it depends on the following points:

(1) The total weight on the edges falling from the common neighbors $w \in N_{uv}$ to the nodes u and v, i.e., $\sum_{w \in N_{uv}} (a_{uw} + a_{vw})$.

(2) The total weight on the edges falling among the common neighbors, i.e., $\sum_{w,x \in N_{uv}} a_{wx}$.

(3) The similarity between the weights of the edges falling from each common neighbor $w \in N_{uv}$ to the nodes u and v.

The points (1) and (2) above can be combined into one factor $\sum_{w,x \in N_{uv}} (a_{uw} + a_{vw} + a_{wx})$ which is smaller than or equal to $\sum_{w \in N_{uv}} s_w$. Thus we have

$$0 < \frac{\sum_{w,x \in N_{uv}} a_{uw} + a_{vw} + a_{wx}}{\sum_{w \in N_{uv}} s_w} \leq 1 \tag{1}$$

The upper bound in the Eq. (1) is achieved when every neighbor of a common neighbor $w \in N_{uv}$ is either u, v or some $x \in N_{uv}$.

To understand how point (3) affects $\rho(u, v)$, let us assume that for some $w \in N_{uv}$, a_{uw} is much higher than a_{vw}. This means that the interaction between u and w is stronger while that between v and w is comparatively weak. So, in a sense w obstructs the interaction between u and v. If this situation occurs for each $w \in N_{uv}$, then the quality of interaction between u and v would worsen. So, to improve $\rho(u, v)$, the weights a_{uw} and a_{vw} must be as close as possible. For this reason, we consider the term $\min(a_{uw}, a_{vw})$. It has the highest value when both a_{uw} and a_{vw} are equal. Now taking the sum of these terms over all such w's we get the term

$$\sum_{w \in N_{uv}} \min(a_{uw}, a_{vw})$$

Now if u and v are neighbors of each other, then a_{uv} would also improve $\rho(u, v)$. Hence, finally we consider the term,

$$\sum_{w \in N_{uv}} \min(a_{uw}, a_{vw}) + a_{uv}$$

Note that

$$\sum_{w \in N_{uv}} \min(a_{uw}, a_{vw}) + a_{uv} \leq \sum_{w \in N_u} a_{uw} = s_u$$

Similarly,

$$\sum_{w \in N_{uv}} \min(a_{uw}, a_{vw}) + a_{uv} \leq s_v$$

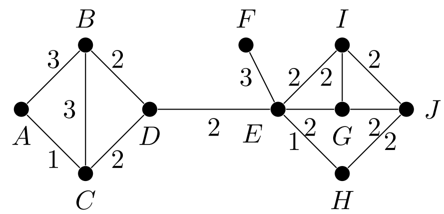

Fig. 1: A weighted network

Thus we have

$$0 < \frac{\sum_{w \in N_{uv}} \min(a_{uw}, a_{vw}) + a_{uv}}{\min(s_u, s_v)} \leq 1 \tag{2}$$

The upper bound in Eq. (2) is achieved when $a_{uv} = a_{vw}$, for all $w \in N_{uv}$ and $N_{uv} = N_u$ or $N_{uv} = N_v$. Now after understanding how the terms in Eq. (1) and (2) can affect the closeness of two nodes, defining $\rho(u, v)$ is mere a matter of taking a suitable combination of these two terms, for example the average or the product. Here, we take the product as it gives a stricter definition of closeness.(the product is high only when both the factors are high.) Therefore, the definition of $\rho(u, v)$ now takes the form,

$$\rho(u, v) = \left(\frac{\sum_{w,x \in N_{uv}} a_{uw} + a_{vw} + a_{wx}}{\sum_{w \in N_{uv}} s_w} \right) \left(\frac{\sum_{w \in N_{uv}} \min(a_{uw}, a_{vw}) + a_{uv}}{\min(s_u, s_v)} \right) \tag{3}$$

Here we have assumed that $N_{uv} \neq \varnothing$. For $N_{uv} = \varnothing$, we define $\rho(u, v)$ as follows:

$$\rho(u, v) = \frac{a_{uv}}{\min(s_u, s_v)} \tag{4}$$

If any of the vertices u and v is isolated, we simply write $\rho(u, v) = 0$. Now the measure is well defined for each vertex pair (u, v) and assumes all possible values between 0 and 1 including the bounds.

Consider the graph in Fig. 1 to compute the neighborhood proximity of a few vertex pairs. Using the definition of ρ, we get $\rho(A, D) = 3/4, \rho(D, E) = 1/3, \rho(D, F) = 1/3, \rho(E, F) = 1$ and $\rho(E, J) = 5/6$.