Network Science Cheatsheet

Made by
Remy Cazabet

Assortativity

Assortativity - Homophily - Mixing Patterns

A network is said to be **assortative** or to demonstrate **homophily** if its nodes tend to connect more with other nodes that are **similar** than to nodes that are different.
Similarity in this case must be understood in term of nodes properties. Some typical examples can be age, gender, language, political beliefs, etc.
Homophily is considered a common feature of many networks, in particular social networks[a], as reflected in the aphorism *Birds of a feather flock together*.
Typical examples would be *age, gender, ethnicity* or *politicla opinions* in social networks networks such as Twitter[b]

[a]McPherson, Smith-Lovin, and Cook 2001.
[b]McPherson, Smith-Lovin, and Cook 2001.

Disassortativity - Heterophily

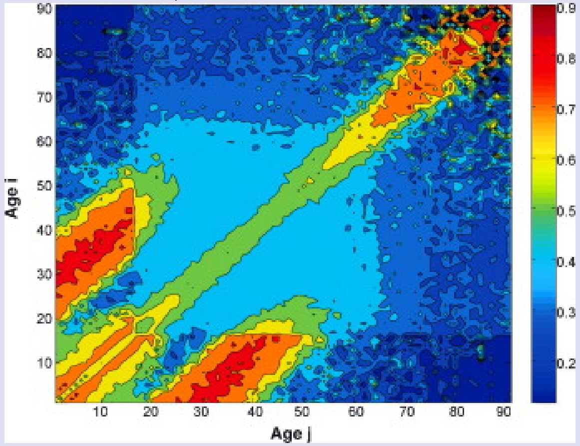
Some networks can also demonstrate **heterophily**, or **disassortativity**, i.e., a greater number of connections with nodes that are different (for instance, in a sentimental relationship network, women tend to connect more with men than with other women, and reciprocally).

Mixing Patterns

The notion of nodes connecting to each other with preferences based on their attributes can be generalized to the concept of **Mixing Patterns**. Beyond homophily/heterophily, nodes with property p_1 can be preferentially connected to nodes with property p_2 (and not p_3 or p_4) while nodes having property p_3 can have a preference for nodes having the same property, for instance.

Mixing Patterns - example

Example of mixing patterns of age in a network of interaction between individuals, reproduced from[a].



We can see that there is some level of assortativity (high values on the diagonal), but that there are also some more complex mixing patterns, for instance between age 10 and 40, approximately, here interpreted as child-parents relationships.

[a]Del Valle et al. 2007.

Note on interpreting homophily

Homophily can be a link creation mechanism (nodes have a preference to connect with similar ones, so the network end up to be assortative), or a consequence of influence phenomenons (because nodes are connected, they tend to influence each other and thus become more similar).
Without access to the dynamic of the network and its properties, it is not possible to differentiate those effects.

Categorical or Numerical homophily

Attributes of nodes can be either categorical (no natural order between values, discrete number of possible values), or numerical (natural order, discrete or continuous). Although the general idea remains the same, the way to compute homophily differs according to type of attributes we are interested in.

Assortativity Index - Definition

When the property for which we study homophily is **categorical**, homophily can be defined[a] by comparing the fraction of edges that connect nodes of the same category, and the expected value of such edges if the network was random. More formally, it is expressed as:

$$r = \frac{\sum_i e_{ii} - \sum_i a_i^2}{1 - \sum_i a_i^2}$$

where $e_i i$ is the fraction of edges connecting two nodes of category i, and a_i the fraction of all edges connected to a node of category i (sum of degrees divided by number of edges).

[a]Newman 2003.

Assortativity index - Example

Let's see a fictional example of how to compute the assortativity index. Nodes are individuals, edges represent for instance some social interaction. Columns/Rows correspond to blood types, and numbers are expressed in fraction of the total number of edges.

Blood Types	A	AB	B	O	a_i
A	0.30	0.05	0.1	0.05	*0.5*
AB	0.05	0.05	0	0	*0.1*
B	0.1	0	0.2	0	*0.3*
O	0.05	0	0	0.05	*0.1*
a_i	*0.5*	*0.1*	*0.3*	*0.1*	**1**

$$r = \frac{(0.3+0.05+0.2+0.05)-(0.5^2+0.1^2+0.3^2+0.1^2)}{1-(0.5^2+0.1^2+0.3^2+0.1^2)} = \frac{0.6+0.36}{1-0.36} = 0.375$$

Asortativity index - Properties

An assortativity index of $r = 0$ means that the network has no assortative mixing, $r = 1$ corresponds to a perfectly assortative network (edges exist only between nodes of the same category), and $r = -1$ to a perfectly disassortative network (no edge between nodes of the same category).

Assortativity and Modularity

Assortativity is related to the Modularity, a measure of the quality of *communities*, by the following relation:

$$r = \frac{Q}{Q_{max}}$$

Indeed, $\sum_i e_{ii} - \sum_i a_i^2$ corresponds to the definition of the Modularity, while $1 - \sum_i a_i^2$ corresponds to the maximal value that the Modularity could reach if all nodes were in the same communities.

Homophily for numeric variables

When the property for which we study homophily is **numeric**, homophily r can be defined as the **Pearson Correlation Coefficient** between values at both end of each edge. For details, see Newman 2003.

Numeric Assortativity index - Properties

Homophily $r = 0$ means that the network has no assortative mixing, $r > 0$ corresponds to an assortative network (nodes with high values tend to connect to high values), and $r < 0$ to a disassortative network (nodes with high values are preferably connected to low values).

Degree assortativity

Degree assortativity[a], sometimes simply called *assortativity*, is a particular case of homophily measured in term of node degrees, i.e., the numerical value associated to each node is its degree. The existence of a degree assortativity can be interpreted in term of a *rich club phenomenon*: hubs prefer to connect to other hubs. ER, Configuration and BA random graph models have a degree assortativity equals to 0, while many real networks have positive values, and some negative ones.

[a]Newman 2003.

Limits of Assortativity

A limit of assortativity coefficients as we have defined them is that they summarize the whole network as a single value. However, different parts of the network might have different types of assortativity.

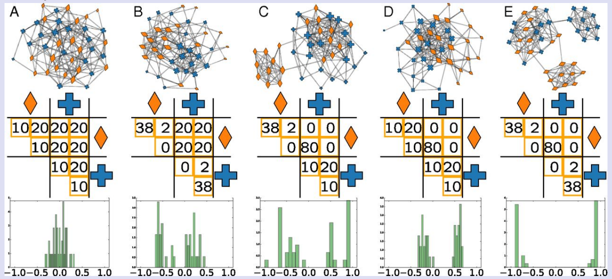


Illustration of different local assortativity behaviors leading to the same global assortativity value (bottom: distribution of local assortatvity). Figure from[a], in which the authors propose a measure of **multiscale assortativity**.

[a]Peel, Delvenne, and Lambiotte 2018.

References

[1] Sara Y Del Valle et al. "Mixing patterns between age groups in social networks". In: *Social Networks* 29.4 (2007), pp. 539–554.

[2] Miller McPherson, Lynn Smith-Lovin, and James M Cook. "Birds of a feather: Homophily in social networks". In: *Annual review of sociology* 27.1 (2001), pp. 415–444.

[3] Mark EJ Newman. "Mixing patterns in networks". In: *Physical review E* 67.2 (2003), p. 026126.

[4] Leto Peel, Jean-Charles Delvenne, and Renaud Lambiotte. "Multiscale mixing patterns in networks". In: *Proceedings of the National Academy of Sciences* 115.16 (2018), pp. 4057–4062.