Applied Mathematics and Computation 311 (2017) 262–271

Contents lists available at ScienceDirect

Applied Mathematics and Computation

journal homepage: www.elsevier.com/locate/amc

Two-walks degree assortativity in graphs and networks

Alfonso Allen-Perkins[a], Juan Manuel Pastor[a,b], Ernesto Estrada[c,*]

[a] Complex System Group, Universidad Politécnica de Madrid, 28040 Madrid, Spain
[b] E.T.S.I.A.A.B, Universidad Politécnica de Madrid, Avd. Puerta de Hierro 4, 28040 Madrid, Spain
[c] Department of Mathematics and Statistics, University of Strathclyde, Glasgow G1 1XQ, United Kingdom

ARTICLE INFO

Keywords:
Graphs
Real-world networks
Walks
Assortativity

ABSTRACT

Degree assortativity is the tendency for nodes of high degree (resp. low degree) in a graph to be connected to high degree nodes (resp. to low degree ones). It is usually quantified by the Pearson correlation coefficient of the degree–degree correlation. Here we extend this concept to account for the effect of second neighbours to a given node in a graph. That is, we consider the two-walks degree of a node as the sum of all the degrees of its adjacent nodes. The two-walks degree assortativity of a graph is then the Pearson correlation coefficient of the two-walks degree–degree correlation. We found here analytical expression for this two-walks degree assortativity index as a function of contributing subgraphs. We then study all the 261,000 connected graphs with 9 nodes and observe the existence of assortative–assortative and disassortative–disassortative graphs according to degree and two-walks degree, respectively. More surprisingly, we observe a class of graphs which are degree disassortative and two-walks degree assortative. We explain the existence of some of these graphs due to the presence of certain topological features, such as a node of low-degree connected to high-degree ones. More importantly, we study a series of 49 real-world networks, where we observe the existence of the disassortative–assortative class in several of them. In particular, all biological networks studied here were in this class. We also conclude that no graphs/networks are possible with assortative–disassortative structure.

1. Introduction

Networks represent the topological skeleton of a wide range of systems in nature and society [1–4]. The characterization of their structure is crucial since it shapes the evolutionary, functional, and dynamical processes that take place in those systems [4–6].

It is well known that links generally do not connect nodes regardless of their characteristics. In social networks, for instance, evidence suggests that individuals prefer to associate with others of similar age, religion, education or occupation as themselves [7]. Assortativity or assortative mixing is a graph metric that refers to the tendency for nodes in networks to be connected to other nodes that are similar (or different) to themselves in some way [8]. Typically, it is determined for the *degree* (i.e. the number of direct neighbours, k) of the nodes in the network [9–12]. The tendency for high-degree nodes to associate preferentially with other high-degree nodes plays a major role in many important processes, such as epidemic

* Corresponding author.
E-mail addresses: alfonso.allen@hotmail.com (A. Allen-Perkins), juanmanuel.pastor@upm.es (J.M. Pastor), ernesto.estrada@strath.ac.uk (E. Estrada).

http://dx.doi.org/10.1016/j.amc.2017.05.025

spreading, synchronization or network robustness, among others [9,13–16]. However, assortativity may be applied to any characteristics of a node, including non-topological vertex properties, such as language or race [8]. Most of the research done in this area has been summarized in the review of Noldus et al. [17]. Other extensions to account for interactions beyond the nearest-neighbours have also been proposed in the recent literature [18].

The aim of this work is to define an assortativity index that captures the influence of first and second neighbours of a node. We then express this two-walks assortativity in terms of the subgraphs contributing to it.

The paper is organized as follows. In Section 2, the preliminaries are presented. In Section 3, the concept of two-walks degree assortativity is introduced and analysed. Main result is demonstrated in Section 4. Numerical results are presented in Section 5. Conclusions are summarized in Section 6.

2. Preliminaries

Here we consider simple, undirected graphs $G = (V, E)$, i.e., graphs without multiple edges, self-loops, directions or weights in their edges. The notation used is standard and the reader can check for instance [19]. Let us define some of the measures used in this work in order to make it self-contained. First, we define the degree assortativity index [8]. Mathematically, it is written as:

$$r_k = \frac{\frac{1}{m}\sum_{(i,j)\in E} k_i k_j - \left\{\frac{1}{m}\sum_{(i,j)\in E}\frac{1}{2}\left[k_i + k_j\right]\right\}^2}{\frac{1}{m}\sum_{(i,j)\in E}\frac{1}{2}\left[k_i^2 + k_j^2\right] - \left\{\frac{1}{m}\sum_{(i,j)\in E}\frac{1}{2}\left[k_i + k_j\right]\right\}^2} \tag{2.1}$$

where k_i and k_j are the degrees at both ends of link $(i, j) \in E$ and m is the number of links. A positive assortativity index $r_k > 0$ indicates the tendency of higher degree nodes in the graph to be connected to other higher degree nodes. On the other hand, $r_k < 0$ indicates the tendency of higher degree nodes to be connected to lower degree nodes. It was previously proved the following result [11].

Lemma 1. *Let $G = (V, E)$ be a simple graph and let k_i be the degree of the vertex i. Let $|P_1|$, $|P_2|$ and $|P_3|$ be the number of edges, of paths of length two and of paths of length three, respectively. Finally, let $|C_3|$ be the number of triangles in G. Then, the assortativity coefficient can be written combinatorially as:*

$$r_k = \frac{|P_3| + 3|C_3| - \frac{|P_2|^2}{|P_1|}}{|P_2| + 3|S_{1,3}| - \frac{|P_2|^2}{|P_1|}} \tag{2.2}$$

Let $|P_{r/s}|$ be the ratio $|P_r|/|P_s|$, $|S_{1,\,3}|$ the number of star graphs of four nodes, and $C = 3|C_3|/|P_2|$. Then G is:

(1) *assortative ($r > 0$): if and only if $|P_{3/2}| + C > |P_{2/1}|$,*
(2) *neutral ($r = 0$): if and only if $|P_{3/2}| + C = |P_{2/1}|$, and $3|S_{1,3}| - |P_2|(|P_{2/1}| - 1) \neq 0$, and*
(3) *disassortative ($r < 0$): if and only if $|P_{3/2}| + C < |P_{2/1}|$*

It is worth mentioning that the denominator of Eq. (2.2) is non-negative. Consequently, the sign of r_k depends only upon the sign of the numerator, which is determined by the following structural factors: the global clustering coefficient (i.e. $C = 3|C_3|/|P_2|$), the intermodular connectivity (i.e. $|P_{3/2}| = |P_3|/|P_2|$) and the branching (i.e. $|P_{2/1}| = |P_2|/|P_1|$) [11].

The number of subgraphs contributing to the degree assortativity can be obtained using the following results [20].

Lemma 2. *Let $G = (V, E)$ be a simple graph with n nodes. Let k_i be the degree of the vertex i. Let $|C_3|$ be the number of triangles in G. Then, the number of edges $|P_1|$, path of length two $|P_2|$ and three $|P_3|$ are given, respectively by*

$$|P_1| = \frac{1}{2}\sum_{i=1}^{n} k_i,$$

$$|P_2| = \frac{1}{2}\sum_{i=1}^{n} k_i(k_i - 1),$$

$$|P_3| = \sum_{(i,j)\in E}(k_i - 1)\big(k_j - 1\big) - 3|C_3|.$$

Lemma 3. *Let $G = (V, E)$ be a simple graph. Let k_i be the degree of the vertex i in G. Let A be the adjacency matrix of G. Let $|P_1|$ and $|P_2|$ be respectively the number of edges and the number of paths of length two in G. Let $|S_{T1D}|$ be the number of subgraphs S_{T1D} in G (see Table 1). Let $|C_i|$ be the number of cycles of i nodes in G. Then, $|C_3|$, $|C_4|$ and $|C_5|$ are given, respectively by*

$$|C_3| = \frac{1}{6}tr\big(A^3\big),$$

$$|C_4| = \frac{1}{8}tr\big(A^4\big) - 2|P_1| - 4|P_2|,$$

$$|C_5| = \frac{1}{10}tr\big(A^5\big) - 3|C_3| - |S_{T1S}|.$$