Hindawi Publishing Corporation
Journal of Applied Mathematics
Volume 2015, Article ID 580361, 9 pages
http://dx.doi.org/10.1155/2015/580361

Research Article

Computing Assortative Mixing by Degree with the *s*-Metric in Networks Using Linear Programming

Lourens J. Waldorp[1] and Verena D. Schmittmann[2]

[1] *University of Amsterdam, Weesperplein 4, 1018 XA Amsterdam, Netherlands*
[2] *Tilburg University, Warandelaan 2, 5037 AB Tilburg, Netherlands*

Correspondence should be addressed to Lourens J. Waldorp; waldorp@uva.nl

Received 13 May 2014; Accepted 8 February 2015

Academic Editor: Frank Werner

Calculation of assortative mixing by degree in networks indicates whether nodes with similar degree are connected to each other. In networks with scale-free distribution high values of assortative mixing by degree can be an indication of a hub-like core in networks. Degree correlation has generally been used to measure assortative mixing of a network. But it has been shown that degree correlation cannot always distinguish properly between different networks with nodes that have the same degrees. The so-called s-metric has been shown to be a better choice to calculate assortative mixing. The s-metric is normalized with respect to the class of networks without self-loops, multiple edges, and multiple components, while degree correlation is always normalized with respect to unrestricted networks, where self-loops, multiple edges, and multiple components are allowed. The challenge in computing the normalized s-metric is in obtaining the minimum and maximum value within a specific class of networks. We show that this can be solved by using linear programming. We use Lagrangian relaxation and the subgradient algorithm to obtain a solution to the s-metric problem. Several examples are given to illustrate the principles and some simulations indicate that the solutions are generally accurate.

1. Introduction

Assortative mixing by node degree (i.e., the number of connections of a node) is the tendency of nodes to be connected to other nodes of similar degree and an important concept in network analysis [1–3]. For example, assortative mixing in social networks could reflect the notion that well-connected people, who know many people, have a tendency to know mainly other well-connected people [4]. Or in a network of actors, where connections indicate that actors worked together on a film, actors with many connections are likely to have worked together with other well-connected actors [5]. As a final example, in symptom networks in psychopathology, where connections represent symptoms belonging to the same disorder, it appears that core symptoms, which are common to multiple disorders, have a high tendency to be mutually connected [6]. Assortative mixing is becoming more and more important, because it points to relevant network characteristics, such as self-similarity and other emergent properties, if it is detected in networks with a power-law degree distribution [1, 7]. Here we propose to compute assortative mixing in undirected networks using linear programming.

Several measures of assortative mixing for undirected graphs exist [5]. One of the most popular ones is the Pearson assortativity coefficient, or degree correlation r [1], which is Pearson's correlation applied to the degrees of each node in the network. Degree correlation is a normalized metric and obtains values between −1 and 1. However, Alderson and Li [8] have shown that the same value of degree correlation can occur from very different configurations of edges (topology) of a network and that the use of degree correlation may lead to incorrect conclusions. The reason for this is that degree correlation is normalized with respect to general networks with a specific degree sequence (number of connections for all nodes) that can have multiple edges, self-loops and can even be disconnected (i.e., consist of multiple disconnected components). However, in many situations the objective is to compare assortativity with networks that are similar, that is, that are connected and simple (no self-loops, connected, and no multiple edges). To remedy this objection to degree correlation, Li et al. [9] proposed the s-metric (first called L-metric and later s-metric), which is a linear transformation of the degree correlation. In the normalized version of the s-metric, normalization is calculated with the maximal and minimal values of s with respect to the class of networks that have the desired properties of connectedness and simplicity for the specified degree sequence. The minimal ($s_{\min}$) and maximal ($s_{\max}$) possible values of s in the class of simple and connected networks with the same degree sequence are compared to the obtained s-value for the network at hand. Then, in contrast to the degree correlation coefficient r which is normalized with respect to unrestricted networks of the same degree sequence, the normalized s-metric is obtained by comparing s of the network under consideration to similar networks that are within the same class of networks of the same degree sequence yet are maximal ($s_{\max}$) or minimal ($s_{\min}$) with respect to assortative mixing.

Obtaining the maximal and minimal network in the class of simple (undirected networks without self-loops and no multiple edges) and connected networks with the same degree sequence is not trivial [8]. The algorithm introduced in Alderson and Li [8] and described in van Mieghem et al. [2] ranks all edges according to the product of degrees of the pair of nodes and then connects nodes according to a nodes degree such that the graph is connected. Unfortunately, this algorithm does not always achieve the exact degree sequence as desired. Alternatively, van Mieghem et al. [2] proposed a rewiring approach increasing (or decreasing) the assortativity, while the degree sequence remains unchanged. The constraint of a connected graph here is sacrificed, however, resulting in possibly disconnected graphs. We propose to use a linear program (LP) to identify $s_{\max}$ and $s_{\min}$. Obtaining the minimum and maximum value of s within the class of graphs with a specific degree sequence that are simple and connected is formulated as a binary integer program (BIP). A binary integer program is a program to optimize a linear objective function given certain linear constraints for a binary (0/1) solution (e.g., [10, 11]). Here the objective function to be optimized is s and the constraints are that the optimal graph is simple and connected and has the specified degree sequence. The s-metric problem resembles the traveling salesman problem (e.g., [12]), except that the degree constraints can be different from value 2. As a consequence we cannot use the approach of Held and Karp [13], where the problem is reformulated to a 1-tree problem and was shown to lead under certain circumstances to a linear program with the same solution as the binary integer problem. We prove a weaker result that leads to an efficient solution of the problem by moving the degree constraints to a penalty term in the optimization (Lagrangian relaxation), which is solved by a subgradient algorithm.

2. Correlation and the s-Metric

Let $G = (V, E)$ be an undirected graph, where V is the set of nodes $\{1, 2, \ldots, n\}$ and E is the set of edges $\{(i, j) : i, j \in V\}$ with size $|E| = m$. The degree sequence d of graph G is the n vector $(d_1, d_2, \ldots, d_n)$ containing for each node the number of connections and the degrees d_i of the nodes in G. The degrees d_i are not necessarily ascending or descending. We are mostly interested in simple connected graphs, that is, no self-loops and no multiple edges with a single component. This class of graphs with degree sequence d is denoted by $\mathcal{S}(d)$. The unconstrained class of graphs with self-loops, multiple edges, and possibly multiple components having degree sequence d is denoted by $\mathcal{U}(d)$. It is immediate that $\mathcal{S}(d) \subset \mathcal{U}(d)$.

The degree correlation or assortative mixing of an undirected graph G is defined in terms of the degree sequence as [1, 8]

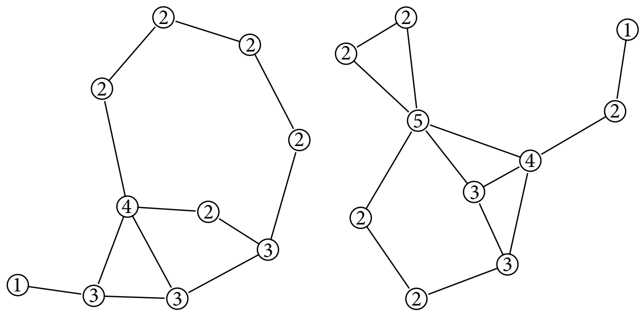
$$r = \frac{\sum_{(i,j)\in E} d_i d_j - m^{-1}\left((1/2)\sum_{i\in V} d_i^2\right)^2}{(1/2)\sum_{i\in V} d_i^3 - m^{-1}\left((1/2)\sum_{i\in V} d_i^2\right)^2}. \tag{1}$$

It is equivalent to the Pearson correlation coefficient and its value is between $-1 \leq r \leq 1$. The first term in the numerator calculates assortativity by degree for the graph G [7]. The first term in the denominator $(1/2)\sum_{i\in V} d_i^3$ can be interpreted as the maximal value for assortative mixing that can be obtained within the class of graphs that may have multiple connections and self-loops and need not be connected in $\mathcal{U}(d)$. The second term in the numerator and in the denominator $m^{-1}((1/2)\sum_{i\in V} d_i^2)^2$ can be seen as the central point in the class of graphs in $\mathcal{U}(d)$ with minimal assortativity [8]. Hence, the Pearson degree correlation or assortativity can be considered as the normalized assortativity within the general class of graphs that may contain self-loops and multiple edges and need not be connected that are in $\mathcal{U}(d)$ (see the discussion on p. 502 of Li et al. [7] for more details).

Another insightful way of considering the degree correlation is given by van Mieghem et al. [2]. Let A be the symmetric $n\times n$ adjacency matrix with a 1 if two nodes are connected and 0 otherwise, and let D be a diagonal matrix with the degrees on the diagonal and zero otherwise. The matrix $L = D - A$ is called the Laplacian matrix. Then it is shown that the degree correlation is

$$\begin{aligned} r &= 1 - \frac{d^{\mathsf{T}} L d}{d^{\mathsf{T}} D d - (2m)^{-1}\left(d^{\mathsf{T}} d\right)^2} \\ &= 1 - \frac{(1/2)\sum_{(i,j)\in E}\left(d_i - d_j\right)^2}{(1/2)\sum_{i\in V} d_i^3 - m^{-1}\left((1/2)\sum_{i\in V} d_i^2\right)^2}. \end{aligned} \tag{2}$$

From this version of degree correlation, it is easily seen that a regular graph has degree correlation $r = 1$, since all degrees are equal. Additionally, it is shown in van Mieghem et al. [2] that a connected Erdös-Renyi random graph has zero assortativity.



(a) $G_1 : r = -0.04, s^n = 0.60$ (b) $G_2 : r = -0.05, s^n = 0.83$

Figure 1: Two example graphs with $n = 10$ nodes and $m = 12$ edges (a) and $n = 10$ nodes and $m = 13$ edges (b) with the degrees inside the nodes.

For a graph $G = (V, E)$ the s-metric is defined by [7, 9]

$$s = \sum_{(i,j)\in E} d_i d_j. \tag{3}$$

It is clear from its definition that the s-metric can obtain values between 0 and $n(n-1)^2$ for the complete graph of size n, K_n. It is therefore convenient to normalize the s-metric such that it obtains values between 0 and 1. The normalized s-metric of graph $G = (V, E)$ is defined by [7, 9]

$$s^n = \frac{s - s_{\min}}{s_{\max} - s_{\min}}, \tag{4}$$

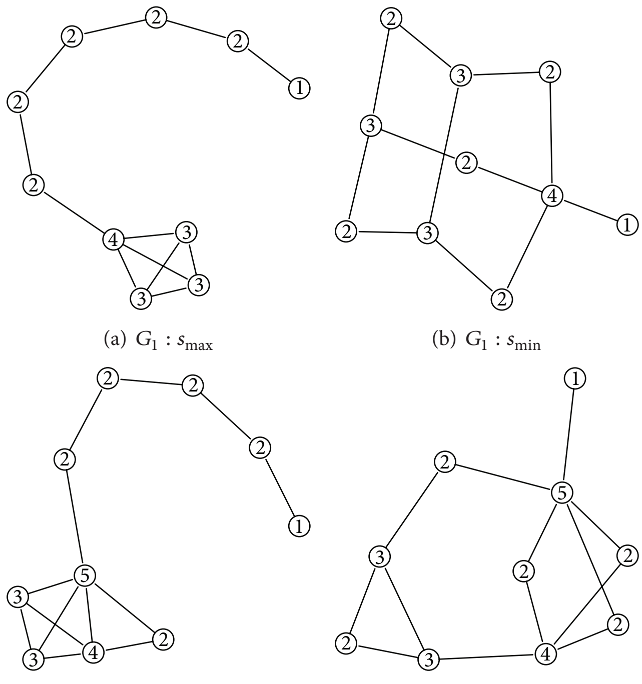
where $s_{\min}$ and $s_{\max}$ refer to the minimal or maximal value of s in a specific class, like $\mathcal{S}(d)$. The value of the normalized s-metric is between $0 \le s^n \le 1$. The first term in the nominator of s^n is $s = \sum_{(i,j)\in E} d_i d_j$, which is the same as that of degree correlation above. The difference between s^n and r is the normalization. The s-metric is normalized with respect to the maximum and minimum obtainable in a specific class, whereas degree correlation is always normalized by the central and maximal values of assortativity by degree within the general class of graphs having self-loops, multiple connections and being possibly disconnected in $\mathcal{U}(d)$.

Example 1. We generated two topologically different graphs, shown in Figure 1. The first graph, shown in Figure 1(a), has $n = 10$ nodes, $m = 12$ edges, and ascending degree sequence $d^\top = (1, 2, 2, 2, 2, 2, 3, 3, 3, 4)$. The other graph in Figure 1(b) has $n = 10$ nodes, $m = 13$ edges, and ascending degree sequence $d^\top = (1, 2, 2, 2, 2, 2, 3, 3, 4, 5)$. The networks have similar degree sequence but the topology is different, as can be seen in Figure 1. However, both networks have approximately zero degree correlation (assortativity), $r_1 = -0.04$ and $r_2 = -0.05$, respectively. This suggests that the networks are similar in topology. The normalized s-metric does pick up the topological differences. The values of the normalized s-metric are $s^n = 0.60$ and $s^n = 0.83$, respectively, indicating that the networks can be correctly distinguished and may have a hub-like core. The values of s^n are higher because the normalization is made within the class of simple, connected graphs, whereas the correlation coefficient r is approximately zero with respect to normalization within the general class of graphs with self-loops and multiple connections that are possibly disconnected. This example shows that there are differences in degree sequence of a connected graph that are not picked up at all by degree correlation but are picked by the s-metric; and the reason for this is that the normalization for the s-metric is within $\mathcal{S}(d)$ whereas the normalization of the correlation is within the larger class $\mathcal{U}(d)$.

The main practical difference between the correlation coefficient r and normalized s-metric is that the correlation coefficient does not distinguish certain graphs that are configurationally (topologically) different whereas the normalized s-metric does distinguish between them. Alderson and Li [8] show that two networks can have a completely different topology but the correlations of both networks are approximately the same. Additionally, there is a large range of values of the coefficient of variation for which the correlation is approximately zero, whereas the s-metric shows clear changes across the same range of values. There is also a set of graphs for which both the degree corruption and the normalized s-metric are approximately zero (see Figure 1 in [8]). One solution to this is to compute the local assortativity [14]. Piraveenan et al. [3] show that assortative graphs can heave locally disassortative hubs, and vice versa, suggesting that local information can be relevant.

The main problem in calculating the normalized s-metric identified by Alderson and Li [8] is that the normalization requires calculation of $s_{\min}$ and $s_{\max}$. The algorithm proposed in Li et al. [7] orders all edges $(i, j) \in E$ according to the weights $d_i d_j$. Nodes are then connected corresponding to this ordering as long as the degree sequence remains unchanged and the graph is connected. Such a heuristic greedy search type algorithm seems to work reasonably well for the maximum within $\mathcal{S}(d)$ but does not always achieve the exact degree sequence d [7]. Additionally, only the maximum is obtained, not the minimum, with this algorithm [8]. The minimum is usually approximated by a lower bound, the minimum in the class $\mathcal{U}(d)$ of unrestricted graphs. The alternative algorithm introduced in van Mieghem et al. [2] randomly rewires two links simultaneously (leaving the degree sequence unchanged) such that nodes with the highest degree are connected and nodes with the lowest degrees are connected, increasing assortativity. Although with this algorithm the correct degree sequence is obtained, the resulting graph is not necessarily connected. In the following sections we present an integer program using a Lagrangian relaxation that obtains both $s_{\min}$ and $s_{\max}$ for simple, connected graphs that have exactly the specified degree sequence.

Example 2. We continue with the two graphs from the previous example with $n = 10$ nodes shown in Figure 1. The graphs that maximize the s-metric within the class of simple, connected graphs for G_1 and G_2 are shown in Figures 2(a) and 2(b). It can be seen that node degree is similar to that of neighbors, as the optimization requires. In contrast, the graphs with minimal value s in Figures 2(c) and 2(d) have



(a) $G_1 : s_{\max}$ (b) $G_1 : s_{\min}$ (c) $G_2 : s_{\max}$ (d) $G_2 : s_{\min}$

Figure 2: Graphs that have maximal value s ((a) and (b)) and minimal values ((c) and (d)) for the graphs G_1 and G_2 corresponding to the ones in Figure 1. Degrees are inside the nodes.

low degree nodes connected to high degree nodes. For G_1 the value for the graph is $s = 85$ and the minimal and maximal values within the class $\mathcal{S}(d)$ are $s_{\min} = 79$ and $s_{\max} = 89$, so that

$$s^n = \frac{85 - 79}{89 - 79} = 0.600. \tag{5}$$

The minimal and maximal values within $\mathcal{U}(d)$ are 79 and 93, respectively. For G_2 the value for the graph is $s = 122$ and the minimal and maximal values within the class $\mathcal{S}(d)$ are $s_{\min} = 108$ and $s_{\max} = 125$, so that

$$s^n = \frac{122 - 108}{125 - 108} = 0.833. \tag{6}$$

The minimal and maximal values within $\mathcal{U}(d)$ are 107 and 142, respectively. As expected the minimal values within $\mathcal{S}(d)$, $s_{\min}$, were similar to the minimal values within $\mathcal{U}(d)$. However, the maximal values are quite different for $\mathcal{S}(d)$ and $\mathcal{U}(d)$, indicating the importance that the maximal value needs to be determined specifically for $\mathcal{S}(d)$.

3. Linear and Integer Programming to Obtain s^n

In this section we first give the formulation of the s-metric problem for binary integer programming. We use the cutset formulation to ensure that the resulting graph is connected. We then show that in general the polytope for a linear program is not integral. However, we can relax the constraints and show that while constraining the solution to be integral, we need not constrain it to be binary, which is what we require.

The objective is to find a graph such that the minimal and maximal values of g within the class of simple and connected graphs with degree sequence d, $\mathcal{S}(d)$, can be obtained to compute s, the normalized s-metric. We will show that obtaining the minimum (or maximum) can be reformulated into a binary integer program (BIP) with a linear relaxation such that a solution can be obtained within polynomial time.

3.1. Problem Formulation. The objective function

$$s = \sum_{(i,j)\in E} d_i d_j \tag{7}$$

is to be optimized subject to two constraints to obtain a graph in $\mathcal{S}(d)$. Since maximization is equivalent to minimizing its negative, we limit our discussion to minimization. The objective function can be rewritten in terms of an incidence vector and the weights $c_{ij} = d_i d_j$, the product of degrees for nodes i and j. Let $E(V) = \{(i,j) : i, j \in V\}$ denote the set of all possible edges for the nodes in V (Kronecker set). Then for an edge e in the set $E(V)$, $x_e = 1$ when an edge is present and $x_e = 0$ otherwise. We can then rewrite the objective function s for preferential attachment as

$$s = \sum_{e\in E(V)} c_e x_e. \tag{8}$$

To obtain a graph in the class of simple and connected graphs with degree sequence d in $\mathcal{S}(d)$, we require several types of constraint. The first is that the degree sequence of the graph should be exactly d. Define, for any $S \subset V$, the cut by the partition $S \cup \overline{S} = V$, where $\overline{S} = V \setminus S$ is the complement of S. Denote the edges in the cut by $\delta(S) = \{(i,j) \in E : i \in S, j \in \overline{S}\}$. Then $\delta(\{i\})$ is the incidence set of node i, and the degree constraint can be defined as $\sum_{e\in\delta(\{i\})} x_e = d_i$. The second type of constraint is to ensure that the graph is connected, for which we use the cutset formulation [11, 15, 16]. The idea is that for any nonempty subset $S \subset V$ of nodes at least one edge has to connect both sets of vertices in S and $\overline{S}$. This implies that for any nonempty $S \subset V$ the cutset $\delta(S)$ has to contain at least one edge. Finally, we require that the optimal solution is binary; that is, x should be in $\mathbb{B}^p = \{0,1\}^p$. We now have the binary integer program (BIP) as follows:

$$z = \min_x \sum_{e\in E(V)} c_e x_e, \tag{9}$$

$$\sum_{e\in\delta(\{i\})} x_e = d_i \quad \forall i \in V, \tag{10}$$

$$\sum_{e\in\delta(S)} x_e \geq 1 \quad \forall S \subset V,\ 2 \leq |S| \leq \left\lfloor \frac{n}{2} \right\rfloor, \tag{11}$$

$$x \in \mathbb{B}^p \quad p = \frac{n(n-1)}{2}, \tag{12}$$

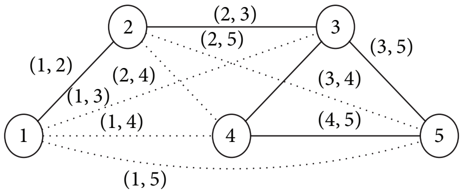


Figure 3: Graph of Example 3. G has $n = 5$ nodes and $m = 5$ edges (solid lines); all other possible edges including labels in $E(V)$ are shown as dotted lines.

where $\lfloor\cdot\rfloor$ means the integer part. The degree constraints in (10) combined with the cutset constraints (11) imply nonnegativity since $d_i \geq 1$ for all $i \in V$. The cutset constraints (11) run to size $\lfloor n/2 \rfloor$ since we do not need to check subsets twice. For any cut S we have $|S| + |\overline{S}| = |V|$ and so for a cut $|S| \geq \lfloor n/2 \rfloor + 1$, we have that this S is the same as $\overline{S}$ when $|S| \leq \lfloor n/2 \rfloor$.

Example 3. Consider a graph G with $n = 5$ nodes and $m = 5$ edges shown in Figure 3 (solid lines). The edge set $E = \{(1,2), (2,3), (3,4), (3,5), (4,5)\}$ can be considered as an instance of the incidence vector x of dimension 10. The incidence vector in this case is

$$\begin{aligned} x &= ((1,2), (1,3), (1,4), (1,5), (2,3), (2,4), \\ &\quad (2,5), (3,4), (3,5), (4,5)) \\ &= (1,0,0,0,1,0,0,1,1,1). \end{aligned} \tag{13}$$

For this G we have degree sequence $d = (1,2,3,2,2)'$ and so

$$\begin{gathered} c^{\mathsf{T}} x = \sum_{e \in E} c_e x_e = \sum_{(i,j) \in E} d_i d_j x_{ij}, \\ 1 \cdot 2 \cdot 1 + 1 \cdot 3 \cdot 0 + 1 \cdot 2 \cdot 0 + \cdots + 3 \cdot 2 \cdot 1 + 2 \cdot 2 \cdot 1 = 24. \end{gathered} \tag{14}$$

With 5 nodes there are 5 constraints for the degrees to equal d_i, giving the degree constraints (10). For connectedness there are sets $S \subset V$ of size $2 \leq |S| \leq \lfloor 5/2 \rfloor = 2$, and so the number of cutsets to be checked is $\binom{5}{2} = 10$. For this x the 10 cutsets are shown in Table 1. For example, take the first partition $S = \{1,2\}$ and $\overline{S} = \{3,4,5\}$. All elements in $E(V)$ for this partition are in the first row. Since nodes 1 and 2 are in the same set S, no edge is required here for connectedness; nodes 1 and 3 are in S and $\overline{S}$, respectively, and so an edge is required for connectedness. This is repeated for all different partitions S and $\overline{S}$. In total there are 15 constraints from the degrees and the cutset constraints.

Note that the number of constraints for this formulation is exponential in the size of the graph n, because all nonempty subsets $S \subset V$ are required for connectedness, with $2 \leq |S| \leq \lfloor n/2 \rfloor$ (see, e.g., [11], p. 471). Specifically, the number of constraints to check whether all subsets $S \subset V$ of size $2 \leq |S| \leq \lfloor n/2 \rfloor$ connect to at least one other node is $r(n) = 2^{h(n)} - h(n) - 1$, where $h(n) = \lfloor n/2 \rfloor$. One possibility is to change formulation to obtain a connected graph, for example, to a subtour elimination formulation, as used in minimal spanning trees (e.g., [11, 12, 15]). The subtour elimination formulation for the constraints in optimizing s is

$$\sum_{e \in E(S)} x_e \leq |S| - 1 \quad \forall S \subset V,\ 2 \leq |S| \leq \left\lfloor \frac{n}{2} \right\rfloor. \tag{15}$$

However, the subtour elimination formulation is in this case inappropriate since the total number of edges can in general exceed $|V| - 1$, which means we are not dealing with a tree.

3.2. Linear and Integer Program. We are interested in designing a linear program that will result in incidence (binary) vectors such that the constraints in (10) and (11) are satisfied. We therefore need to bound the solution we obtain in $\mathbb{R}_+^p$ to be at most 1; that is,

$$x_e \leq 1 \quad \forall e \in E(V). \tag{16}$$

Let $P = \{x \in \mathbb{R}_+^p : x \text{ satisfies (10), (11)}\}$ be the polyhedron associated with the s-metric problem. Let $Q = \{x \in \mathbb{R}_+^p : x \text{ satisfies (11)}\}$ be the polyhedron of the cutset constraints. The polyhedron P resembles the subtour elimination polytope (SEP), associated with the symmetric traveling salesman problem [12]. The difference is that for the s-metric problem the degrees are not fixed to be 2 but are equal to the original graph.

Example 4. Consider a graph with three nodes and two edges, $G = (V, E)$ with $V = \{1,2,3\}$ and $E = \{(1,2), (1,3)\}$; see Figure 4(a). Here, $x = (1,1,0)$ is the incidence vector. The cutset constraints (11) and the degree constraints (10) are in this example:

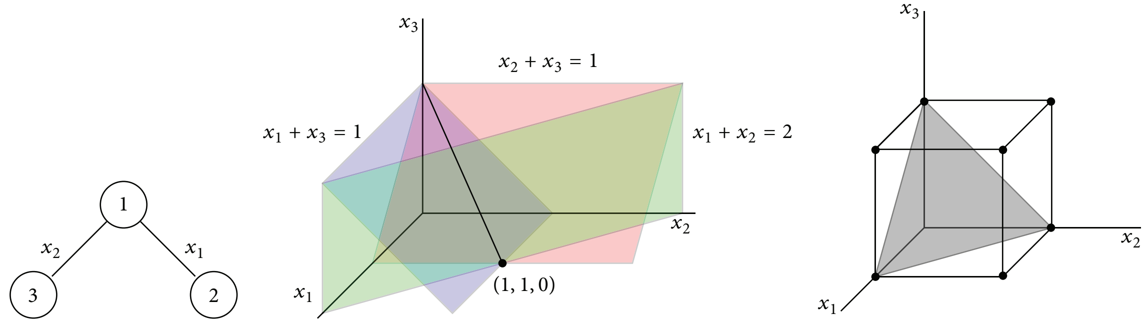
$$\begin{gathered} \text{cutset} \begin{cases} x_1 + x_2 & \geq 1 \\ x_1 \quad + x_3 & \geq 1 \\ \quad x_2 + x_3 & \geq 1, \end{cases} \\ \text{degree} \begin{cases} x_1 + x_2 & \geq 1 \\ x_1 \quad + x_3 & \geq 1 \\ \quad x_2 + x_3 & \geq 1, \end{cases} \end{gathered} \tag{17}$$

This is a description of P in three dimensions. If we consider the degree constraints only, then we see that the only extreme point is $x^* = (1,1,0)$. This is because the face $F = \{(1,1,0)\}$ has dimension 0, since the three degree constraints are supported and its associated matrix has rank 3 (see Figure 4(b)). Note that this is not true for other points like $(1,0,1)$. It is also seen that the polytope P is not integral, since the point $x^* = (1/2, 1/2, 1/2)$ is an extreme point. This point is in the center of the triangle in Figure 4(c).

From this example it is clear that in general we need to consider points in $P \cap \mathbb{Z}^p$ such that we only obtain integral solutions. Our problem is different from the traveling salesman problem approached by Held and Karp [13] as a

Table 1: The incidence vectors x for different cutsets $\delta(S)$ showing edges between S and its complement $\overline{S}$.

S	$\overline{S}$	x_e									
		$(1,2)$	$(1,3)$	$(1,4)$	$(1,5)$	$(2,3)$	$(2,4)$	$(2,5)$	$(3,4)$	$(3,5)$	$(4,5)$
$\{1,2\}$	$\{3,4,5\}$	0	1	1	1	1	1	1	0	0	0
$\{1,3\}$	$\{2,4,5\}$	1	0	1	1	1	0	0	1	1	0
$\{1,4\}$	$\{2,3,5\}$	1	1	0	1	0	1	0	1	0	1
$\{1,5\}$	$\{2,3,4\}$	1	1	1	0	0	0	1	0	1	1
$\{2,3\}$	$\{1,4,5\}$	1	1	0	0	0	1	1	1	1	0
$\{2,4\}$	$\{1,3,5\}$	1	0	1	0	1	0	1	1	0	1
$\{2,5\}$	$\{1,3,4\}$	1	0	0	1	1	1	0	0	1	1
$\{3,4\}$	$\{1,2,5\}$	0	1	1	0	1	1	0	0	1	1
$\{3,5\}$	$\{1,2,4\}$	0	1	0	1	1	0	1	1	0	1
$\{4,5\}$	$\{1,2,3\}$	0	0	1	1	0	1	1	1	1	0



(a) (b) (c)

Figure 4: The graph G from Example 4 in (a) and the polytope in (b) obtained from the degree constraints, and the polyhedron Q in (c) from only the cutset including the bound 1 constraints. The degree constraints in (b) show that their intersection at $(1,1,0)$ is the only possible solution here. The line in (b) indicates where the faces induced by the constraints $x_1 + x_3 = 1$ and $x_2 + x_3 = 1$ intersect.

1-tree problem, because the degree sequence we seek for the solution is in our case fixed to the one of the original graph. It follows that any simple graph with the number of edges $\geq n$ will not lead to a 1-tree. A general solution to the s-metric problem cannot therefore be linked to a 1-tree. We can establish a weaker result than integrality of P, though, which allows us to use a linear program effectively but is not guaranteed to lead to an integral result. When considering only the cutset constraints in Q, we find that the constraints of unity for each x_e imply that considering the convex hull of the integral solutions will result in binary solutions.

Theorem 5. *Let* $\operatorname{conv}(Q \cap \mathbb{Z}^p)$ *be the convex hull over the integer points in* $Q = \{x \in \mathbb{R}_+^p : x \text{ satisfies } (11), (16)\}$. *An LP can be used to find the optimal solution such that*

$$\min\left\{c^{\top}x : x \in \operatorname{conv}\left(Q \cap \mathbb{Z}^p\right)\right\} = \min\left\{c^{\top}x : x \in Q \cap \mathbb{B}^p\right\}. \tag{18}$$

Proof. We can rewrite the polyhedron Q such that it is associated with a set-covering problem. For such a set-covering problem we use Proposition 6.3 of Nemhauser and Wolsey [11] which, together with the constraint from cutset (11), then proves the claim.

In a set-covering problem the edges cover all nodes and can be described by a minimization over a polyhedron with constraints of the form $a_H^{\top}x \geq 1$, where a_H is a characteristic binary vector representing a graph indexed by H. Let a set $B(\mathscr{C})$ of subsets of nodes V such that (intersection) $H \cap C \neq \varnothing$ for all $C \in \mathscr{C}$ and (minimality) if $K \subset H$ then $K \cap C = \varnothing$ for some $C \in \mathscr{C}$ [11]. The set $B(\mathscr{C})$ is called a blocking clutter and $\mathscr{C}$ is called a clutter. Then $H \in B(\mathscr{C})$ is a subset of the nodes V which represents the edges that connect a partition V into S and its complement $\overline{S}$. For any path from $s \in S$ to $t \in \overline{S}$ the vector $(a_H)_i = 1$ if it connects one node from S to another node from $\overline{S}$ and 0 otherwise. Let $(x_C)_i = 1$ if the edge is on a path for some node $s \in S$ to $t \in \overline{S}$ and 0 otherwise. Then the characteristic vectors a_H with $H \in B(\mathscr{C})$ and x_C with $C \in \mathscr{C}$ are a blocking clutter and a clutter, respectively. Since a path from $s \in S$ to $t \in \overline{S}$ for some $C \in \mathscr{C}$ coincides with an edge that connects S to $\overline{S}$ for some $H \in B(\mathscr{C})$ (intersection property), we have that $a_H^{\top}x_C \geq 1$ for some $C \in \mathscr{C}$. This means that the s-metric problem is a set-covering problem $\min\{c^{\top}x : x \in Q \cap \mathbb{B}^p\}$, where $Q = \{x \in \mathbb{R}^p : a_H^{\top}x \geq 1 \text{ for all } H \in B(\mathscr{C})\}$.

If we show that the extreme points of $\operatorname{conv}(Q \cap \mathbb{Z}^p)$ are exactly the paths from a node $s \in S$ to a node $t \in \overline{S}$, represented by the vectors x_C with $C \in \mathscr{C}$, then we are done.

By the intersection property we find that if $x \in Q \cap \mathbb{Z}^p$, then there is an $H \in B(\mathscr{C})$ such that $a_H^\top x_C = 1$ for all $C \in \mathscr{C}$. This follows because a path from S to $\overline{S}$ coincides with exactly one edge for an $H \in B(\mathscr{C})$. This shows that the incidence vectors x_C are the extreme points of $\operatorname{conv}(Q \cap \mathbb{Z}^p)$. Since all characteristic vectors x_C are incidence vectors and so in $\mathbb{B}^p$, the claim of the theorem follows. □

This result, together with the fact that $P = \{x \in \mathbb{R}_+^p : x \text{ satisfies (10), (11)}\}$ is not integral, tells us that we may consider a relaxation of the s-metric problem by dropping the degree constraints (10) and considering optimizing only over $Q = \{x \in \mathbb{R}_+^p : x \text{ satisfies (11)}\}$. One way of doing this is to incorporate the degree constraints (10) in the objective function; this is a Lagrangian relaxation of the degree constraints. This is similar to the approach of Held and Karp [13].

4. Relaxation of the Integer Problem

The linear relaxation of the s-metric problem with the cutset formulation means dropping constraint (12) that x is 0/1 [17]. The LP problem is then $z_{\text{LP}} = \min\{c^\top x : x \in P\}$. Linear relaxation, however, does not guarantee an integral solution, and in fact we know that P is not integral (see Example 4). As in the traveling salesman problem, the combinations of the degree (equality) constraints and the cutset constraints make optimization difficult (see, e.g., [12, 13, 18, 19]). Here we consider Lagrangian relaxation of the degree constraints. We use the subgradient method to find the optimal solution with the step size for the subgradient as suggested in Held and Karp [13] (see also Fisher [18]).

4.1. Lagrangian Relaxation. We consider Lagrangian relaxation where the degree constraints (10) are considered hard and so are entered as a penalty in the function to optimize. For fixed $u \in \mathbb{R}^n$ (positive or negative) we have the Lagrangian relaxation:

$$z_{\text{LR}}(u) = \min\left\{c^\top x + u^\top (d - Dx) : x \in Q \cap \mathbb{Z}^p\right\}. \tag{19}$$

Since we are dealing with equality constraints of degree $Dx = d$ we have that a solution $x(u)$ for the Lagrange multiplier u is also a solution to the original IP, because $d - Dx = 0$ for this solution. It follows that $z_{\text{LR}}(u) \le z_{\text{IP}}$. To obtain the bound closest to the original problem, the greatest lower bound is of interest, which is the Lagrangian dual $\max\{z_{\text{LR}}(u) : u \in \mathbb{R}^n\}$. Combining this gives the Lagrangian dual [11]:

$$z_{\text{LD}} = \min\left\{c^\top x : Dx = d, x \in \operatorname{conv}\left(Q \cap \mathbb{Z}^p\right)\right\}. \tag{20}$$

This means that we can use an LP to find the Lagrangian dual. We know from Theorem 5 that this results in a binary solution, which is required; we cannot drop the integrality constraint in general since Q is not integral. One way to obtain a solution to the Lagrangian dual is by subgradient optimization.

4.2. Subgradient Optimization. For differentiable functions, minimization over convex functions is relatively straightforward [20]. If $f : \mathbb{R}^p \to \mathbb{R}$ is differentiable and convex, then

$$f(x) \ge f(x^*) + \nabla f(x^*)(x - x^*) \tag{21}$$

for all $x, x^* \in \mathbb{R}^p$, and $\nabla f(x^*)$ is the gradient of f at x^*. If x^* is a local minimum then the gradient is 0 and it follows from convexity that x^* is also a global minimum. For nonsmooth and/or nondifferentiable functions the gradient cannot be defined everywhere [18]. In the subgradient algorithm the gradient vector is replaced by a subgradient. Let the vector $u \in \mathbb{R}^p$ be a subgradient at x^* if

$$f(x) \ge f(x^*) + u^\top (x - x^*) \tag{22}$$

for all $x \in \mathbb{R}^p$. Since the function $z_{\text{LR}}(u)$ is piecewise linear and convex, we can define a subgradient by the degree constraints $d - Dx(u)$ [11, 21].

The essential step in this algorithm is to determine how to change the subgradient u correctly. Like in the gradient method the subgradient is adjusted at each iteration in the direction most favorable to the constraints. A sequence of subgradients (u^k) is generated by

$$u^{k+1} = u^k - \theta^k \left(d - Dx\left(u^k\right)\right). \tag{23}$$

We then have that when $d - Dx(u^k) = 0 \in \partial u$, the set of all subgradients (subdifferential), we have obtained the global minimum. The subdifferential ∂u contains 0 if the step size $\theta^k \to 0$ [22]. The most often used way to determine the step size is by the so-called Held-Karp method (see, e.g., [18, 21]):

$$\theta^k = \gamma^k \frac{z\left(u^k\right) - w}{\left\|d - Dx(u^k)\right\|^2}, \tag{24}$$

where $0 < \gamma^0 \le 2$, $w \le z_{\text{LD}}$, and $\|\cdot\|$ denotes the Euclidean norm. The sequence γ^k is halved whenever $z(u^k)$ remains unchanged for some number of iterations [18].

The subgradient algorithm for the Lagrangian dual is presented in Algorithm 1.

4.3. Computational Results. Computations were done with *lpSolve* (version 5.5) in R [23]. The *lpSolve* software is a mixed integer linear programming solver that uses (revised) simplex methods for linear programming and branch and bound methods for integer programming [24].

We generated 100 connected random graphs (Erdös-Renyi) of nodes with size $n = 100$. Edges between each pair of nodes were determined by a probability of around 0.4 (requiring that the graph be connected), resulting in graphs with $|E|$ approximately 4470.

For the subgradient algorithm we used $\theta^0 = 2$ and $\gamma^0 = 1.8$, which was halved if the objective value did not improve [18]. The Lagrange multiplier starting value was $u^0 = \max_i\{d_i\}1_n$, the maximum of degrees for all nodes.

Similar to Fisher [18] we report the percentage of problems for which the IP solution is exactly the same as

(1) $u \leftarrow u^0 \in \mathbb{R}^n, \theta \leftarrow \theta^0 \in \mathbb{R}$, and $k \leftarrow 0$
(2) **while** $s \neq 0$ **do**
(3) solve
$$z(u^k) = \min\left\{c^\top x + \left(u^k\right)^\top (d - Dx) : x \in \operatorname{conv}\left(Q \cap \mathbb{Z}^p\right)\right\}$$
and $x(u^k)$ is the optimal solution
(4) $s \leftarrow d - Dx(u^k)$
(5) $u^{k+1} \leftarrow \max\{u^k - \theta^k s, 0\}$, where
$$\theta^k = \gamma^k \frac{z(u^k) - w}{\|d - Dx(u^k)\|^2}$$
and $w \leq z_{\text{LD}}$
(6) $k \leftarrow k + 1$
(7) **end while**
(8) the optimal solution is $x(u^{k-1})$
(9) **end**

Algorithm 1: Subgradient algorithm for Lagrangian dual.

the LD solution, the ratio of the average of the LD solutions and the average of the IP solutions, to indicate the sharpness of the bound, and the maximum of the subgradient, which should be (close to) 0. For the 100 simulations the percentage of exact solutions $z_{\text{IP}} = z_{\text{LD}}$ is 81.48 and the ratio $100 \times \text{ave}(z_{\text{LD}})/\text{ave}(z_{\text{IP}})$ is 99.96. This suggests that in many cases the result is exact and that if it is not exact then the bound using the Lagrangian dual is sharp. The subgradient was in all cases 0 or close to 0; the maximum subgradient over all simulations was $2.665 \cdot 10^{-15}$, indicating that the degree constraints were satisfied in all cases.

5. Application to Psychopathology and the DSM IV

The diagnostic statistical manual (DSM) IV contains criteria to diagnose people as having certain disorders like general anxiety (GA) or major depression (MD) disorder. These criteria are a set of symptoms which need to be ascertained for a candidate patient. Many of the symptoms are part of different disorders and so it is "easy" to have multiple disorders. For instance, both GA and MD have a set of nine symptoms of which any five must be valid for a person to be suffering from that disorder. But GA and MD have four symptoms in common, and hence, many patients suffering from GA also suffer from MD (see [25] for an elaborate discussion on this).

A network can be created from the DSM IV by connecting symptoms that belong to the same disorder [6]. This network contains $n = 439$ nodes and $m = 2621$ edges. In this network approximately 47% of the nodes is connected, giving rise to a giant component. This giant component has $n = 208$ nodes and $m = 1944$ edges. The network is shown in Figure 5. The range of the degree sequence is between 0 and 71 with median 17, and its distribution seems to fit an exponential distribution [6]. The degree correlation of the giant component is 0.18, while the normalized s-metric is 0.56, suggesting that the graph is assortative. Again the

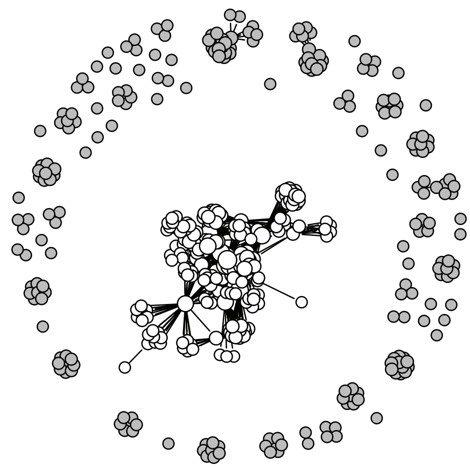


Figure 5: The DSM IV graph constructed by connecting all symptoms that belong to the same disorder. The 208 white nodes belong the giant component and the remaining 231 gray nodes consist of 28 components.

degree correlation and normalized s-metric obtain different numbers for the same graph, due to the comparison in either the general class of graphs for the degree correlation or in the restricted class of graphs for the normalized s-metric.

The DSM IV graph constructed by connecting all symptoms that belong to the same disorder. The 208 white nodes belong the giant component and the remaining 231 gray nodes consist of 28 components.

6. Discussion

Assortativity is for certain classes of graphs best calculated by the normalized s-metric [8]. The existing algorithms by Li et al. [7] and van Mieghem et al. [2] to determine the normalized s-metric are not optimal in the sense that the first one does not always obtain the desired degree sequence and the second does not necessarily result in a connected graph. Here we used a linear program to calculate the minimal and maximal value graphs with respect to the s-metric. This is a binary integer program which we showed can be solved by a linear program. Any solution satisfies the constraints, and hence, the obtained minimal or maximal graph with respect to the s-metric is simple and connected with specified degree.

In this paper we considered the class of simple and connected graphs with specified degree sequence. Although the ideas used here can be used for other types of graphs, generalization is not immediate for all variations. For instance, if multiple edges were allowed between any pairs of nodes, then showing that a linear program can be used is less straightforward, indicating that some solutions are difficult to find. Also, generalizations to other metrics, like shortest path betweenness, are not straightforward because to compute the shortest path length is not linear in the adjacency matrix.

Since there exists a class of graphs for which both the degree correlation and the s-metric are approximately zero [8], it may be more appropriate to consider local assortativity as introduced by Piraveenan et al. [14]. In Piraveenan et al. [3] it is shown that many networks that are disassortative have many hubs that are assortative, so revealing interesting structure at a local scale.

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

References

[1] M. E. J. Newman, "Assortative mixing in networks," *Physical Review Letters*, vol. 89, no. 20, Article ID 208701, 2002.

[2] P. van Mieghem, H. Wang, X. Ge, S. Tang, and F. A. Kuipers, "Influence of assortativity and degree-preserving rewiring on the spectra of networks," *The European Physical Journal B*, vol. 76, no. 4, pp. 643–652, 2010.

[3] M. Piraveenan, M. Prokopenko, and A. Y. Zomaya, "Classifying complex networks using unbiased local assortativity," in *Proceedings of the Alife 12th Conference*, pp. 329–336, Odense, Denmark, 2010.

[4] A.-L. Barabási and R. Albert, "Emergence of scaling in random networks," *Science*, vol. 286, no. 5439, pp. 509–512, 1999.

[5] A. Barrat, M. Barthelemy, and A. Vespignani, *Dynamical Processes on Complex Networks*, Cambridge University Press, 2008.

[6] D. Borsboom, A. O. J. Cramer, V. D. Schmittmann, S. Epskamp, and L. J. Waldorp, "The small world of psychopathology," *PLoS ONE*, vol. 6, no. 11, Article ID e27407, 2011.

[7] L. Li, D. Alderson, J. C. Doyle, and W. Willinger, "Towards a theory of scale-free graphs: definition, properties, and implications," *Internet Mathematics*, vol. 2, no. 4, pp. 431–523, 2005.

[8] D. L. Alderson and L. Li, "Diversity of graphs with highly variable connectivity," *Physical Review E*, vol. 75, no. 4, Article ID 046102, 2007.

[9] L. Li, D. Alderson, W. Willinger, and J. Doyle, "A first-principles approach to understanding the internet's router-level topology," in *Proceedings of the 2004 Conference On Applications, Technologies, Architectures, And Protocols For Computer Communications (SIGCOMM '04)*, pp. 3–14, ACM, *Computer Communication Review*, vol. 34, no. 4.

[10] A. Schrijver, *Theory of Linear and Integer Programming*, Wiley-Interscience Series in Discrete Mathematics, John Wiley & Sons, Chichester, UK, 1986.

[11] G. Nemhauser and L. Wolsey, *Integer and Combinatorial Optimization*, John Wiley & Sons, 1999.

[12] S. Boyd, *The subtour polytope of the travelling salesman problem [Ph.D. thesis]*, University of Waterloo, 1986.

[13] M. Held and R. M. Karp, "The traveling-salesman problem and minimum spanning trees," *Operations Research*, vol. 18, no. 6, pp. 1138–1162, 1970.

[14] M. Piraveenan, M. Prokopenko, and A. Y. Zomaya, "Local assortativity and growth of Internet," *European Physical Journal B—Condensed Matter and Complex Systems*, vol. 70, no. 2, pp. 275–285, 2009.

[15] D. Bertsimas and J. Tsitsklis, *Introduction to Linear Optimization*, Athena Scientific and Dynamic Ideas, Belmont, Mass, USA, 1997.

[16] S. Krumke, Integer Programming, 2007.

[17] A. Schrijver, "Polyhedral proof methods in combinatorial optimization," *Discrete Applied Mathematics*, vol. 14, no. 2, pp. 111–133, 1986.

[18] M. L. Fisher, "The lagrangian relaxation method for solving integer programming problems," *Management Science*, vol. 50, no. 12, pp. 1861–1874, 2004.

[19] S. Boyd and P. Elliott-Magwood, "Structure of the extreme points of the subtour elimination problem of the stsp," in *Combinatorial Optimization and Discrete Algorithms*, S. Iwata, Ed., pp. 33–47, 2010.

[20] J. R. Magnus and H. Neudecker, *Matrix Differential Calculus with Applications in Statistics and Econometrics*, Wiley Series in Probability and Statistics, John Wiley & Sons, Chichester, UK, 1999.

[21] M. Held, P. Wolfe, and H. P. Crowder, "Validation of subgradient optimization," *Mathematical Programming*, vol. 6, pp. 62–88, 1974.

[22] M. Held and R. M. Karp, "The traveling-salesman problem and minimum spanning trees. Part II," *Mathematical Programming*, vol. 1, no. 1, pp. 6–25, 1971.

[23] R Development Core Team, *R: A Language and Environment for Statistical Computing*, R Foundation for Statistical Computing, Vienna, Austria, 2012, http://www.R-project.org/.

[24] M. Berkelaar, K. Eikland, and P. Notebaert, *lpsolve: Open Source (Mixedinteger) Linear Programming System*, GNU Lesser General Public License, 2004.

[25] A. O. J. Cramer, L. J. Waldorp, H. L. J. van der Maas, and D. Borsboom, "Comorbidity: a network perspective," *Behavioral and Brain Sciences*, vol. 33, no. 2-3, pp. 137–150, 2010.