Networks and their degree distribution, leading to a new concept of small worlds

Peer-reviewed author version

EGGHE, Leo (2024) Networks and their degree distribution, leading to a new concept of small worlds. In: Journal of informetrics, 18 (3) (Art N° 101554).

DOI: 10.1016/j.joi.2024.101554
Handle: http://hdl.handle.net/1942/43395

Networks and their degree distribution, leading to a new concept of small worlds

Leo Egghe

Hasselt University, Hasselt, Belgium

E-mail: leo.egghe@uhasselt.be

ORCID: 0000-0001-8419-2932

Abstract

The degree distribution, referred to as the delta-sequence of a network is studied. Using the non-normalized Lorenz curve, we apply a generalized form of the classical majorization partial order.

Next, we introduce a new class of small worlds, namely those based on the degrees of nodes in a network. Similar to a previous study, small worlds are defined as sequences of networks with certain limiting properties. We distinguish between three types of small worlds: those based on the highest degree, those based on the average degree, and those based on the median degree. We show that these new classes of small worlds are different from those introduced previously based on the diameter of the network or the average and median distance between nodes. However, there exist sequences of networks that qualify as small worlds in both senses of the word, with stars being an example. Our approach enables the comparison of two networks with an equal number of nodes in terms of their "small-worldliness".

Finally, we introduced neighboring arrays based on the degrees of the zeroth and first-order neighbors.

Keywords: network theory; Lorenz curves; generalized Lorenz majorization, small worlds; degrees; comparison of networks; trees; neighboring array

1. Introduction

Consider an undirected network or graph G = (V,E), where V denotes the set of vertices or nodes, and E denotes the set of edges or links. In this text, the terms graph and network refer to the same mathematical concept and are used interchangeably. A path of length n is a sequence of vertices (v_0, …v_k, v_{k+1}, …, v_n) such that {v_0, …, v_{n-1}} and {v_1, …, v_n} are sets (being sets each consist of different elements) and for k= 0,…, n-1, v_k is adjacent to v_{k+1}. A cycle is a path for which the starting point v_0 coincides with the endpoint v_n. A graph is connected if there exists (at least one) path between any two vertices. If #V = N, then the degree of node i, i = 1, …, N, i.e., the number of edges connected to node i, is denoted as . In this article we always assume that G is connected, hence all degrees are strictly larger than zero. A shortest path between two nodes in G is a path with the smallest length. The length of a shortest path defines a distance in the mathematical sense of the word in the set of network nodes. Then the diameter of a network is defined as the supremum of the distances between its nodes. As there is no natural order among the nodes in a network we assume that their degrees are ranked in decreasing order. We note that we use the term 'decreasing order' also when the decrease is not strict.

Notation

The array of degree values of the nodes in a network G with N nodes is denoted as

. (1)

where . We will informally refer to such an array as a delta-sequence, consisting of delta-values. Clearly, , a notion which is known as the total degree of the network. It is easy to see that 2(N-1) N(N-1). The lower bound is obtained e.g., for a chain consisting of N nodes, see further, while the upper bound is obtained for a complete graph where each node is connected to each other node.

As real-world networks are often dynamic we will work, as we did in a previous study on small-world networks (Egghe & Rousseau, 2024), within the context of a sequence of finite node sets of networks , i.e., # = N Of course, also all edge sets E_N are finite.

Before moving on to examples and theory we recall the following definitions.

1.1 Definition: Free or unrooted tree (Knuth, 1973, p. 363)

A free or unrooted tree is a connected graph with no cycles. Equivalently it is a connected graph such that removing any edge makes it disconnected. Another equivalent definition states that if v and v′ are different vertices, then there exists exactly one path from v to v′. As we will never use the notion of the root of a tree, we will just use the term 'tree' for "free tree".

1.2 Definition. Isomorphic graphs

Two graphs G and G′ are isomorphic if there exists a bijection f between the vertices of G and G′ such that there is an edge between vertices u and v in G if and only if there is an edge between the vertices f(u) and f(v) in G′.

1.3 Definition: Spanning tree of a connected graph

A spanning tree of an N-node connected graph is a set of N-1 edges that connects all nodes of the network and contains no cycles. A graph may have different (non-isomorphic) spanning trees.

1.4 Examples of networks and their delta-sequences

1.4.1 The complete network on N nodes

The delta-sequence of an N-node complete network is

In this case N(N-1) and its diameter is 1.

1.4.2 The N-star

The N-star consists of a central node and N-1 peripheral nodes, each with one link, namely to the center. Then

Its sum is 2(N-1) and its diameter is 2.

1.4.3 The N-polygon

The N-polygon consists of N nodes forming a simple path of different nodes, that links to its starting point. Then

Its sum is 2N and its diameter is N/2 for even N and (N-1)/2 for odd N.

1.4.4 The N-chain

The N-chain consists of one path of N different nodes. Then

Its sum is 2(N-1) and its diameter is N-1.

1.4.5 Trees

It is obvious that there is no delta-sequence applicable to all trees, but we do have the following lemma.

Lemma (Knuth, 1976, p. 363)

An N-node connected network is a tree if and only if it has N-1 edges and hence a total degree equal to 2(N-1).

We note that stars and chains are all special trees.

2. The delta-sequence and a generalized majorization partial order

2.1 The standard Lorenz curve and Gini index

As the delta-sequence does not have a fixed sum we first consider its (standard) Lorenz curve. The highest Lorenz curve for a network with N nodes is obtained for the star. It always starts by connecting the origin to the point with coordinates . In general, its standard Gini index (Rousseau et al., 2018, formula (4.19)) is (N-2)/2N with a limiting value of 0.5. The lowest Lorenz curve, with Gini index zero, is obtained for a delta-sequence consisting of the same numbers, such as for any complete network, but also for any polygon.

2.2 The non-normalized Lorenz curve

In (Egghe & Rousseau, 2023a) we used the so-called non-normalized Lorenz curve in a continuous context. This study and its follow-up (Egghe & Rousseau, 2023b) led to a rigorous definition of the notion of global impact. We would say, based on (Egghe & Rousseau, 2023b), that the notion of majorization in a network does not only depend on the number of links, but also on their concentration.

In the discrete context, the non-normalized Lorenz curve is defined as follows.

Definition: Non-normalized Lorenz curves

Let be a decreasing N-array of non-negative real numbers, then the corresponding non-normalized Lorenz curve is the polygonal line connecting the origin (0,0) with the points , j= 1, …, N. This curve ends at the point with coordinates .

Definition: The non-normalized (or generalized) majorization order for N-arrays

If X and Y are decreasing N-arrays of non-negative real numbers, then X is majorized by Y, denoted as X Y if

(2)

The relation is only a partial order as non-normalized Lorenz curves (just like standard Lorenz curves) may intersect, see further. If then obviously X Y, but the opposite relation does not hold.

Definition. Acceptable measures

If **X** denotes the set of all decreasing N-arrays of non-negative real numbers, then a function m: **X** is an acceptable measure for the relation if X Y implies that m(X) ≤ m(Y).

It is important to note that the definitions of non-normalized Lorenz curves and in particular the notion of the generalized Lorenz majorization order, denoted as , and the corresponding acceptable measures are generally applicable to all decreasing N-arrays of non-negative real numbers.

Hence, the above definitions can be applied to the set of delta-sequences and the corresponding networks, leading to expressions such as for two N-node networks, but also to the gamma-sequences, introduced later in this text. We already note that if T(G) denotes a spanning tree of the network G, then T(G) G.

As an illustration of the importance of the generalized majorization order, we recall the following definition.

Definition: Network density (Wasserman & Faust, 1994)

The density D of an undirected network G with N nodes is defined as

(3)

Clearly, network density is just a normalized total degree. Two N-networks with the same density D have non-normalized Lorenz curves with the same endpoint, but D does not say anything about the exact relation between the two non-normalized Lorenz curves. In this sense, the majorization partial order applied to delta-sequences refines the notion of network density.

Let now and be the degree sequences of the N-node networks G and H, then the following theorem holds.

2.4 Theorem 1

(i)

(ii)

(iii) Md() ≤ Md() and neither does it imply that Md() ≥ Md(), where Md stands for the median of a sequence.

(iv) The reverse implications do not hold

Proof. (i) and (ii) follow trivially from the definition of the majorization relation .

(iii) We provide two counterexamples (N=5)

The 5-chain G has a degree sequence = (2,2,2,1,1) and H (see Fig. 1) has a degree sequence = (3,2,1,1,1). Then but M() = 2 > M() = 1.

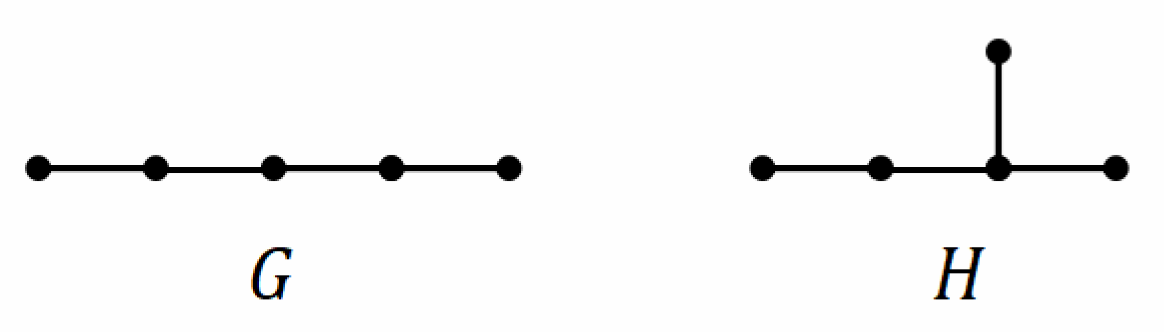

Fig. 1. Networks G and H illustrating part (iii)

For G_1 and H_1 (see Fig. 2) we have = (3,3,2,2,2) and = (4,3,3,2,2). Then and M() = 2 < M() = 3.

Fig. 2 Networks G_1 and H_1 illustrating part (iii)

(iv) For the opposite of case (i) we consider the networks H_1 and H_2 (see Fig. 3) with = (5,3,3,3,3,3) and = (4,4,4,3,3,2). Then neither , nor , illustrating that is a partial, not a complete, order. Moreover, < , = = and □

Fig.3 The networks H_1 (left) and H_2 (right) used in part (iv)

From our earlier investigation (Egghe & Rousseau, 2023a), we know that the following are acceptable measures for generalized majorization among delta-sequences:

A. The Gini index: Gini() =

B. The entropy or Theil measure: Th() =

C. The power measure: P() =

Being applied to generalized majorization, these measures themselves are generalizations of the original ones.

2.5 The network with the lowest non-normalized Lorenz curve.

Theorem 2. An N-node chain is the lowest connected network in the generalized majorization partial order (fixed N).

Proof. By the Lemma in 1.4. 5, the endpoint of the generalized Lorenz curve of any network that is not a tree is situated strictly above that of a tree. Hence, the lowest possible network must be a tree. Among all trees, the N-chain has the lowest generalized Lorenz curve.

Remark. We note that it is not even possible for a general network to have a generalized Lorenz curve that at any place is situated below that of an N-chain. Indeed, when the endpoint is fixed, then the lowest generalized Lorenz curve is the one whose classical Lorenz curve is the diagonal. As the lowest possible endpoint is 2N, this corresponds e.g., to the N-polygon, with delta-array (2,2,…, 2). Its cumulative array is (2,4,6,…., 2N-4, 2N-2, 2N). Yet, the corresponding array for the N-chain is (2,4,6,…., 2N-2, 2N-1,2N), showing that it is not possible to be situated (locally) strictly under the generalized Lorenz curve of the N-chain.

Remark further that the largest generalized Lorenz curve of an N-node network is the one corresponding to the N-complete network.

2.6 Examples of non-comparable networks

We already remarked that the relation is only a partial order, implying that some N-node networks are not comparable. Theorem 1 already provided some examples. Here we provide some more examples.

a) Case N =5

Consider the networks in Fig. 4

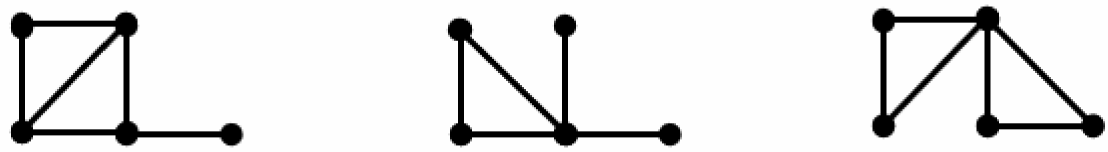

Fig.4. Incomparable 5-node networks: G_1, G_2 and G_3

For the first and the second network, we have delta-sequences (3,3,3,2,1) and (4,2,2,1,1). Hence and . These networks have a different number of links. Consider now G_1 and G_3 with the same number of links. The delta-sequence of G_3 is (4,2,2,2,2). Then, clearly and . We further note that $D(G_1) > D(G_2)$, and but and .

b) Case N = 6 and beyond.

Consider again the networks shown in Fig.3.

The delta-sequences of H_1 and H_2 are respectively (5,3,3,3,3,3) and (4,4,4,3,3,2). Their sums are equal to 20. Yet, and . If we add chains of equal length to a node with degree 3, we may obtain incomparable networks with any larger number of nodes.

c) The cases $2 \leq N \leq 4$.

i) There is only one network with N = 2, hence any two networks are comparable.

ii) The case N=3. Then there are only 2 non-isomorphic networks, shown in Fig.5.

Fig. 5. The two non-isomorphic networks with three nodes.

The delta-sequence of the chain on the left is (2,1,1), while the delta-sequence of the polygon on the right is (2,2,2). Clearly, the chain is strictly smaller than the polygon.

iii) The case N = 4. There are six non-isomorphic connected 4-node networks. See Fig. 6.

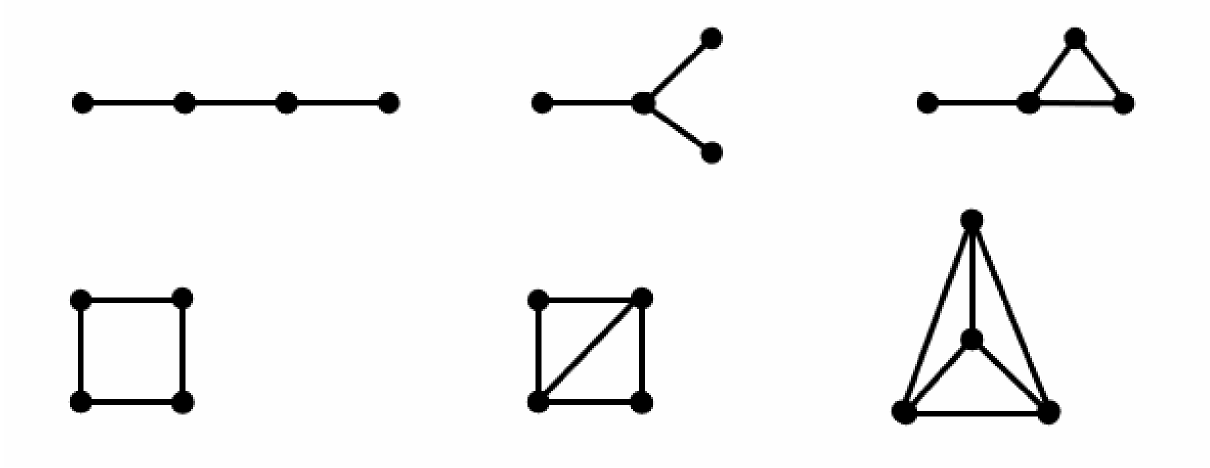

Fig. 6. The six non-isomorphic networks with degree 4.

Referring to these networks, from left to right and from the first row to the second, as G_1, G_2, G_3, G_4, G_5, and G_6 we obtain the following delta-sequences:

= (2,2,1,1); = (3,1,1,1); = (3,2,2,1); = (2,2,2,2); = (3,3,2,2); and = (3,3,3,3). The corresponding cumulative distributions are: (2,4,5,6), (3,4,5,6), (3,5,7,9), (2,4,6,8), (3,6,8,10) and (3,6,9,12). Hence we have the following relations between these networks (Fig. 7).

Fig.7 Relations between delta-sequences of networks with 4 nodes

We see that only G_2 and G_4 are incomparable. We note that the (generalized) Gini-indices for these networks are: Gini(G_1) = 17, Gini(G_2) = 18, Gini(G_3) = 24, Gini(G_4) = 20, Gini(G_5) = 27 and Gini(G_6) = 30, illustrating the fact that the Gini index is an acceptable measure for the generalized majorization order.

3. Classical small worlds

Starting from the late nineties, researchers like Watts and Strogatz (1998), Albert et al. (1999), Barabási and Albert (1999), and Newman and Watts (1999) began associating their studies on networks with the small-world phenomenon, also known as the "six degrees of separation" (Milgram, 1967). In essence, a small-world network is defined by its short average distance between nodes.

In a previous article (Egghe & Rousseau, 2024) we studied small worlds in this sense and in two variants thereof. In the following, we assume we are given a sequence of finite sets and a distance function d defined on each individual level . The term "small world" was used there for a sequence of finite sets satisfying one of the properties defined below.

3.1 Small worlds based on the diameter (SWD)

If for each , is the diameter of, defined as

(4)

then is a SWD if there exists a finite constant $C \geq 0$ such that

(5)

Note that is short for diam().

3.2 Small worlds based on the average distance (SWA)

If , denotes the average distance between two different elements in :

(6)

then   is an SWA if there exists a finite number $C \geq 0$ such that

(7)

3.3 Small worlds based on the median distance (SWMd)

If , denotes the median distance between two different elements in :

(8)

then is a SWMd if there exists a finite number $C \geq 0$ such that

(9)

Note that {{…}} in (12) refers to a multiset (Rousseau et al., 2018, 5.13.1), i.e. a "set" in which elements may occur more than once. It is obvious that if a sequence of finite sets is an SWD then it is also an SWA and an SWMd (Egghe & Rousseau, 2024, section 2.4).

We recall from (Egghe, 2024) that if (G), j = 1,…, N-1, denotes the number of times distance j (the shortest distance between two nodes) occurs in the network G, then the array is called the – array of the network G.

4. Degree sequences and small worlds

4.1 Introduction to this section

In this section, we introduce a new class of small worlds, namely those based on degrees of nodes in a network. Similar to a previous study, small worlds are defined as sequences of networks with certain limiting properties. We distinguish between different types of small worlds and show that these new classes of small worlds are different from those introduced previously (Egghe & Rousseau, 2024). However, there exist sequences of networks that qualify as small worlds in both senses of the word, with stars being an example.

As suggested by a referee we explain why we think that this new approach to small worlds is worth pursuing. The idea of a "degree small world" (in short: a DSW) comes from the observation that the and sequences are basic in network theory and that the notion of a small world has (so far) only been defined for the alpha-sequence. When considering both sequences, we have the intuitive feeling that higher degrees are related to smaller diameters, hence smaller worlds. This is evidenced by several observations. First, the sum of all delta-values is equal to $2\alpha_1$ as a trivial consequence from the definitions of the alpha- and delta-sequences. This shows that higher degrees yield more occurrences of direct links, i.e., of distance 1 (with N fixed). This, in turn, implies since the sum of all alpha-values is equal to N(N-1)/2 (N fixed) that the higher indices of the alpha's, i.e., the larger distances, diminish in occurrence.

For this reason, we are convinced that a degree variant of the notion of a small world makes sense. Since now we focus on high degrees instead of small distances, we look for a function f(N) so that the limit for N tending to infinity of a delta-related measure divided by f(N) is equal to infinity, instead of being strictly smaller than infinity as in the alpha-case. In the next sections, we used δ_1 and two other delta-related measures. We realize that there is not a unique function f(N) that could be used, but we can already exclude some functions by the following argument.

It is clear that DSW differs from SW, otherwise, we wouldn't have to make a new study. Yet, we still want extreme network cases (such as the complete network and the chain) to yield the same qualification in the DSW as in the SW case. For the complete network (the "smallest" world) we still must have that it is DSW (as it is SW). Now $\Delta = (N-1, \ldots, N-1)$ (N times), and hence a function such as f(N) = N is excluded, while a function such as f(N) = ln(N) is acceptable. Using f(N) = ln(N) is also acceptable for the chain, consisting of N-2 times the value 2 and two times the value 1. Indeed, with f(N) = ln(N) taking the limit yields that the chain is not DSW (and it is also not SW).

The same argument applies to the N-polygon consisting of N times the value 2. A star is DSW, and it also is SW.

We will show further that our new approach enables the comparison of two networks with an equal number of nodes in terms of their "small-worldliness".

4.2. Definitions of small worlds derived from the degree distribution

Because we will define here small worlds derived from the degree distribution we will use the abbreviation DSW.

4.2.1 Small worlds derived from the largest degree (DSWL)

Let be an infinite sequence of networks with the Nth network having N nodes.

If ,, is the delta-sequence of, then is a degree small world based on the largest degree if

(10)

We informally say that is DSWL.

4.2.2 Small worlds derived from the average degree (DSWA)

If , denotes the average degree in network ,

(11)

then is a degree small world based on the average degree if

(12)

In the same vein as above we say that is DSWA.

Remark

Strictly speaking it is not necessary to assume that the limits in the above formulae exist: the use of e.g., limsup in (5), (7) and (9) and e.g., liminf in (10), (12) and (13) could serve as well for the idea of a small world (SW). But since the former formulae are conform with the ideas of a SW in the literature we keep on using lim in the above formulae.

Also, in the (counter-)examples, the existence of the limits are verified (see also further on). In addition, in these limits, we have that N is the number of nodes in G_N, and hence is also the variable in the expression lim $_{N+\infty}$. We admit that some examples use a subsequence of , i.e., , with , where , but this is not a major issue and, in our opinion, not enough reason for extending the theory to include limits such as which is rather cumbersome and non-elegant.

4.2.3 Small worlds derived from the median degree (DSWMd)

If , denotes the median degree of the network G_N, then is a degree small world based on the median degree if

(13)

We say that is DSWMd.

4.3.Proposition 1

a) If is DSWMd, then is DSWA

b) If is DSWA, then is DSWL

c) If is DSWMd, then is DSWL

d) the reverse relations do not hold.

Proof. Implication a) follows from the Markov property (Chow & Teicher, 1978), which states that. Hence, if is DSWMd, then is DSWA.

Implication b) follows from the fact that .

Implication c) follows immediately from implications a) and b), or by noticing that

d) Consider the following sequence of star networks   (Fig. 8)

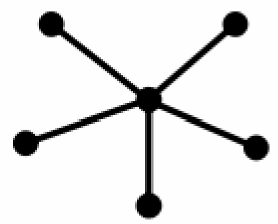

Fig. 8. Star network (illustrated for N=6)

Then . Clearly is DSWL, but not DSWA or DSWMd.

Finally, we have to show that DSWA does not imply DSWMd.

Consider Fig.9, to which we refer as an M-spider, denoted as S_M (Fig.9 shows a spider with M =5). It consists of a complete M-node graph, where each node has an extra two links. Hence N= 3M.

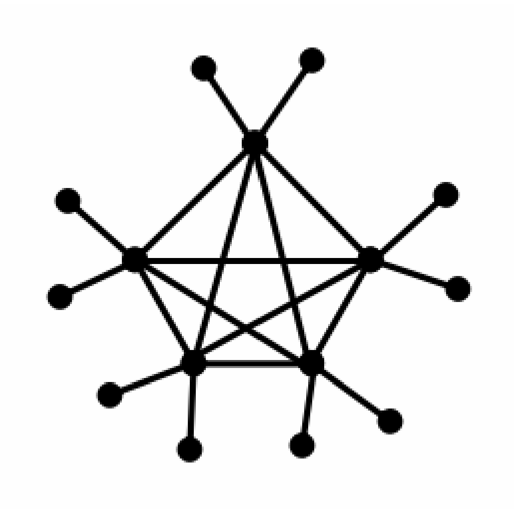

Fig.9. 5-spider

The delta-sequence of an M-spider is . The average and hence , showing that is DSWA. As) = 1, this shows that is not DSWMd.□

4.4 Examples

We have already considered the sequence of stars. Now, we have a look at complete graphs, chains, and polygons.

4.4.1. The sequence of complete graphs.

For each N, . Hence complete graphs are DSWMd and hence also DSWA and DSWL.

4.4.2. The sequence of chains of length N, . Then . Chains are not small worlds based on their degree sequences.

4.4.3. Polygons . These are formed by connecting begin and end nodes of chains. Their delta-sequences are: . We see that polygons too are not small worlds based on their degree sequences.

5. The relation between SWs and DSWs

In this section, we will find out if being an SW (Egghe & Rousseau, 2024) based on the so-called alpha-sequence, implies also being a DSW or vice versa. It will be shown that such implications do not exist, which implies that the notions of SW and DSW are different concepts.

Two sequences of relations are already known: SWD SWA SWMd, (Egghe & Rousseau, 2024) and DSWMd DSWA DSWL (see above). We will prove now that there is, in general, no relation between these sequences of implications.

5.1 Theorem 3

(i) SWD DSWL

(ii) DSWMd SWMd

Proof.(i). We first construct a network for fixed N > 7. Consider a chain with nodes (step 1). Each of these points has descendants (see Fig. 10) (step 2).

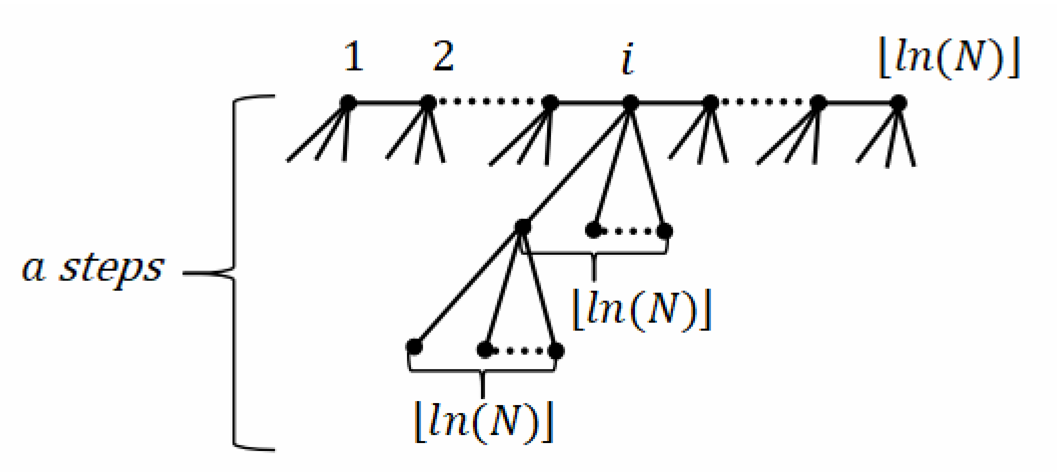

Fig.10 Sketch of the construction of a counterexample, used in Theorem 3 (i)

We continue this construction until at step a, we have . We see that a is smaller than or equal to any b for which . Hence we take b = . We see that this network's diameter satisfies the equality . Hence

which proves that is SWD. However, is not DSW because each delta-value is smaller than or equal to +2, so that cannot be equal to +∞.

(ii). We will construct a kite consisting of an M-complete network and a tail consisting of M-1 nodes (Fig. 11). Hence N = 2M-1. For fixed N the delta-sequence of this kite, K_N, is . We see that .

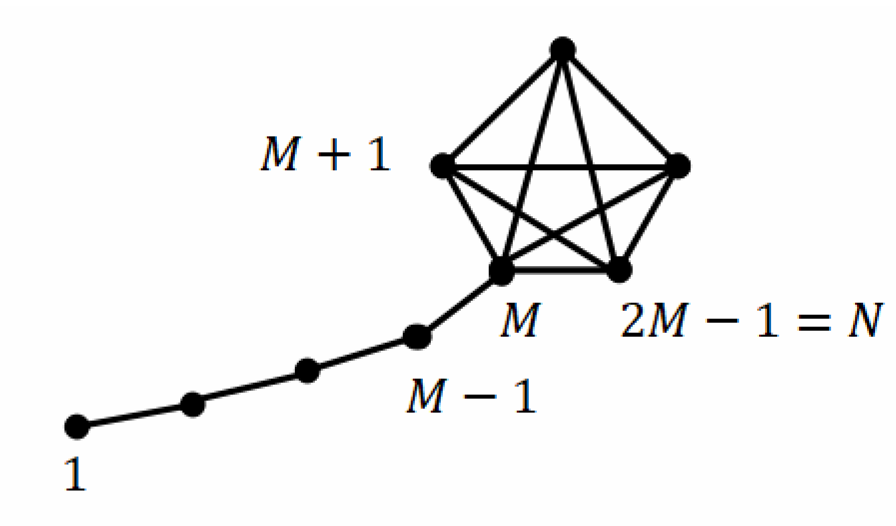

Fig. 11. Kite with N=2M-1 nodes

Now, the kite's alpha-sequence is: . One may check that = and that .

Now we see that = =

Then the median is that natural number i such that i is the first number for which (*) > . This means that

and thus . The plus sign is not possible as otherwise i > N, hence

>

Hence: .

Consequently: which proves that is not(SWMd).

We further see that the median of the delta-sequence of (denoted here as) is M-1 = (N-1)/2. Hence:

which shows that the sequence is DSWMd. □

Corollary. If is DSWL then it is not necessarily SWD.

Proof. Assume that if is DSWL then it is also SWD. Now DSWMd implies DSW (Proposition 1) from which we would know that is SWD, from which it would follow by (Egghe & Rousseau, 2024) that were SWMd, which is a contradiction (by Theorem 3).

Proposition 2. Consider Z_1 = and Z_2 = , then .

Proof. It suffices to give one sequence in the intersection . We know (Egghe & Rousseau, 2.6.2) that the sequence of stars is SWD and we also know that this sequence is DSWL. This proves this proposition.

Proposition 3. Consider Z_3 = and Z_4 = , then .

Proof. Again it suffices to give one sequence in the intersection. The sequence of N-chains is situated in the intersection, see Example 4.4.2. Note that also the sequence of N-polygons provides another example, see Example 4.4.3.

6. Delta-sequences and small worlds derived from degree distributions

6.1 Theorem 4

Consider the network sequences and , such that for each , have the same number of nodes. If now, there exists , such that for each : , then

(a) is DSWL implies that is DSWL, and hence DSWA.

(b) is DSWA implies that is DSWA.

(c) it does not follow that is DSWMd implies that is DSWMd.

(d) the opposite relations of (a) and (b) do not hold.

Proof. Results (a) and (b) follow from the definitions of DWDL and DSWA and Proposition 1 in 3.3. The opposite relations of (a) and (b) do not hold, because if they did then we would have an equivalence in that proposition, which does not hold.

Finally, we prove part (c). Inspired by the spider S_M we construct the following networks. We consider three positive natural numbers M, a, and b (a and b stay fixed) and construct two networks with N = 2M+a+b nodes. For the first one, denoted as $S_{1,N}$, we take b < a. It consists of a complete (M+a) network, where, moreover, on (M+b) of these nodes we add one node (by a single link), see Fig. 12.

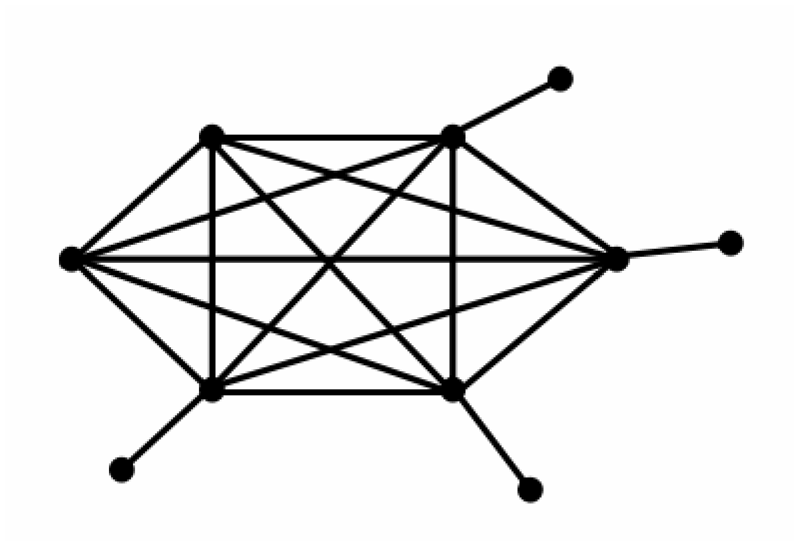

Fig. 12 Case M = 3, a = 3, b= 1

Then .

For the second network, denoted as $S_{2,N}$, we take b > a. It again consists of a complete M+a network, on each of these nodes we add a singly-linked node, while moreover on b-a nodes we add a second, single-linked node, see Fig. 13.

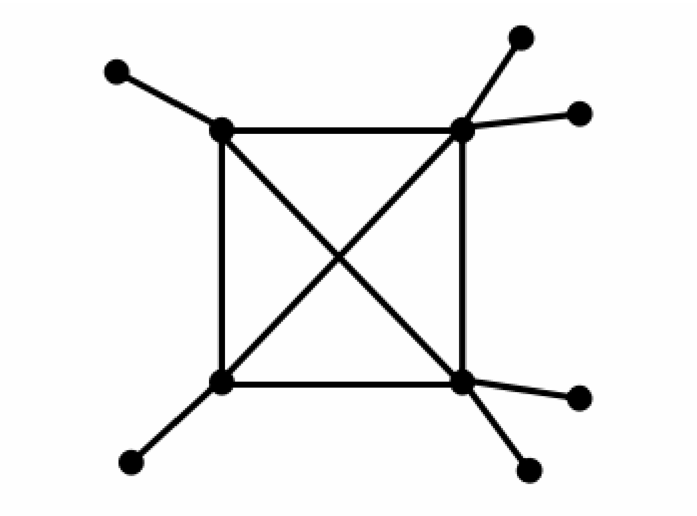

Fig. 13. Case M =3, a = 1, b=3

Then .

Clearly, for each N. Now = M+a-1, which tends to infinity for N (or M) . This shows that is DSWMd. Further, (as b > a), which shows that is NOT(DSWMd).

6.2 Definition

Based on Theorem 4, we propose the following definition:

Given two networks G and H with the same number of nodes and with delta-sequences and , such that then we say that network H is a smaller world than network G in the degree majorization sense.

We explain this statement: one cannot say that a network is a small world (at least not in our view as we defined the notion of a small world only for network sequences), but it is possible to compare two networks in terms of small-worldliness. This is done based on the generalized Lorenz curve. This is similar to the use of the classical Lorenz curve for comparing e.g., income inequality.

Theorem 4 shows that the uses of the terms "small world", "smaller world" and "small-worldliness" are consistent.

7. The neighboring array and the neighboring index

7.1 Definitions

Let G be a connected network with N nodes. If v is a node in G, then the neighborhood of v, denoted A(v), is the set {w in G: v and w are linked}. Let

Then the neighboring (or gamma) array of G, is denoted as = , where the numbers are the numbers , defined above, arranged in decreasing order.

Stated otherwise, the gamma-value of a node in a network is equal to the sum of the degrees of its zeroth and first-order neighbors. Next we define (G) as the neighboring index of G.

7.2 A characterization of the neighboring array.

Let A be the adjacency matrix of an N-node network and let e = be the unit array. Then we have the following matrix multiplication result.

Theorem 5.

Proof. It is easy to check (and well-known) that . It is also well-known that the elements (i,j) of matrix A^2, denoted as , yield the number of paths from node i to node j with length 2. Then

□

7.3 Proposition 4

(14)

Proof. Every value occurs times for its first-order neighbors, plus one more time for its zeroth-order neighbor (itself). □

Remark. By (14) follows from , but when is given, it is still an open problem if it is possible to construct .

7.4 Examples (with N nodes); we assume that gamma-values are given in decreasing order.

7.4.1 The complete N-node network

7.4.2 The star

7.4.3 The polygon (N>2)

7.4.4 Chain (N>3)

7.4.5 The non-isomorphic networks (N=6).

Consider the networks shown in Fig. 14. Their alpha-sequences are the same, namely and so are their delta-sequences: , but their gamma-sequences are different, showing that these networks are not isomorphic: while =. Note that

This example shows that gamma-sequences in combination with alpha- and delta-sequences are stronger than the combination of alpha- and delta-sequences alone in detecting isomorphism classes. Yet, the combination of alpha-, delta- and gamma-sequences is not enough to detect isomorphism as shown in the next example.

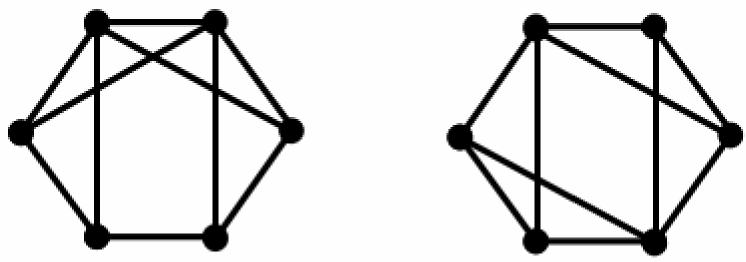

Fig. 14 Two non-isomorphic networks G (left) and G′ (right)

7.4.6 An example of two non-isomorphic networks with equal alpha-, delta-, and gamma-sequences.

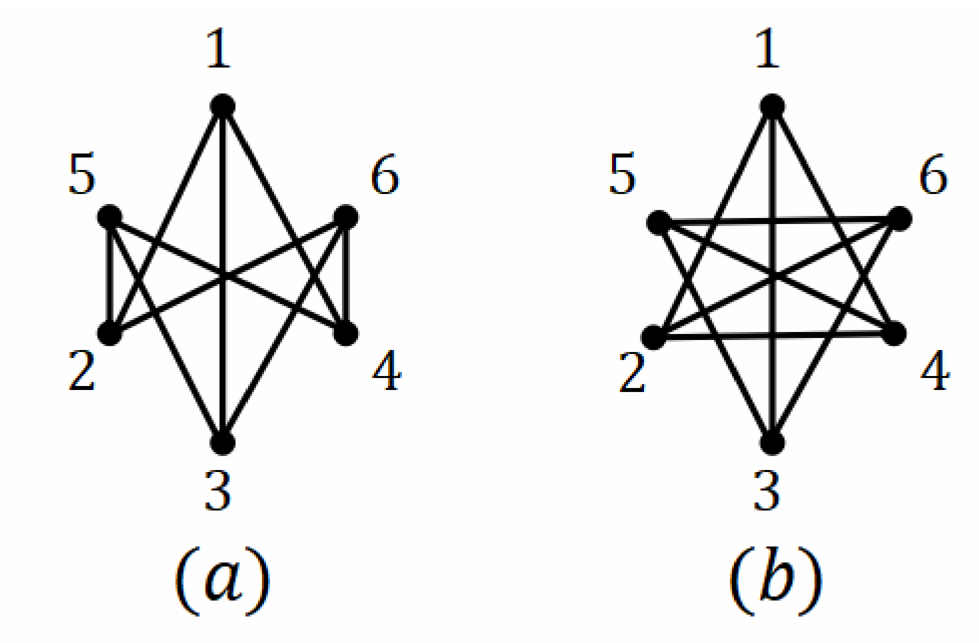

Fig.15. Two non-isomorphic networks with equal alpha-, delta- and gamma-sequences

For both networks in Fig. 15, we have AF=(9,6,0,0); = (3,3,3,3,3,3) and Γ = (12,12,12,12,12,12). Yet, they are not isomorphic as Fig. 15(a) has triples that are not connected, such as {1,5,6}, while such triples do not exist in Fig.15(b).

Remarks

(a) If two networks have the same total degree then they do not necessarily have the same neighboring index .

The following networks, shown in Fig.16 (note that they are trees) have the same = 8, but the chain has = 22, while the other one has = 24.

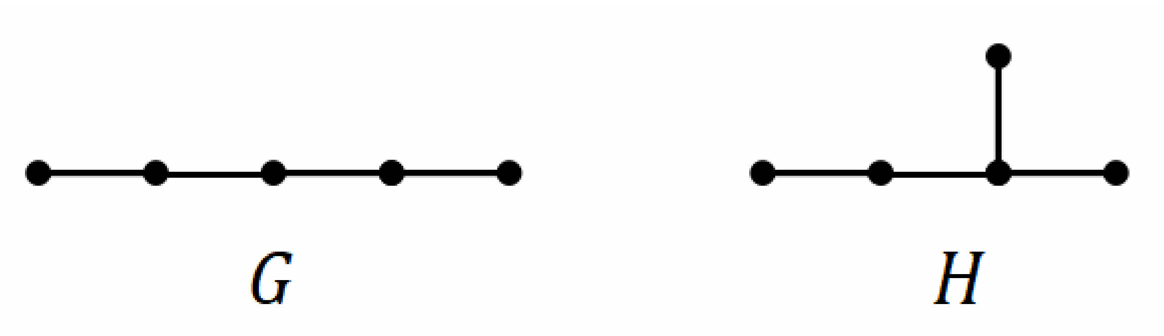

Fig. 16. Networks with the same total degree but different neighboring index (same example as in Fig. 1)

(b) If two networks have the same neighboring index then they do not necessarily have the same total degree .

The networks shown in Fig. 17 (with N = 6) have the same neighboring index , namely 48, but different total degrees , namely 14 and 12.

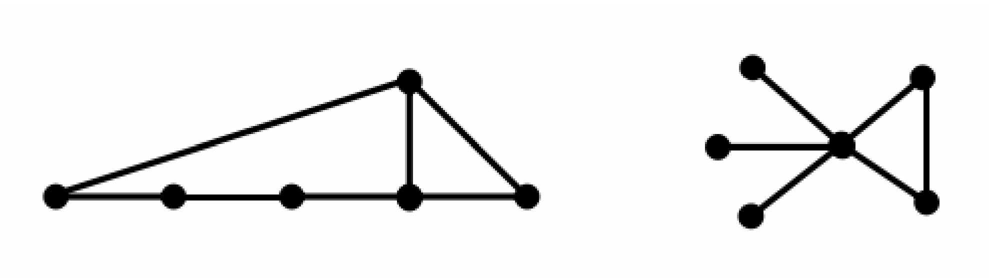

Fig. 17. Two networks with the same neighboring index but different total degree

We close this section by providing examples that the majorization relation is not kept between Δ and Γ.

Concretely: Δ Δ′ Γ Γ′ nor Γ′ Γ (15)

and Γ Γ′ Δ Δ′ nor Δ′ Δ. (16)

Indeed, for (15) we consider the networks (N=6), see Fig. 18.

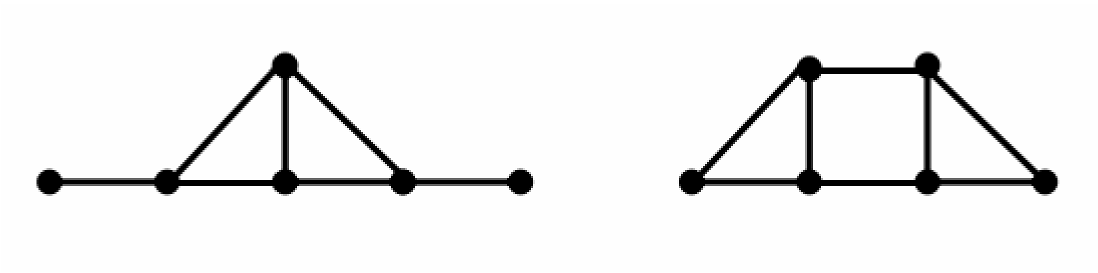

Fig.18. Illustration for (15)

The upper network has Δ = (3,3,3,3,1,1) and Γ = (12,12,10,10,4,4) while the lower one has Δ′ = (3,3,3,3,2,2) and Γ′ = (11,11,11,11,8,8). Then Δ Δ′, but neither Γ Γ′ nor Γ′ Γ holds.

For the case (16) (with N =4) we consider Fig.19.

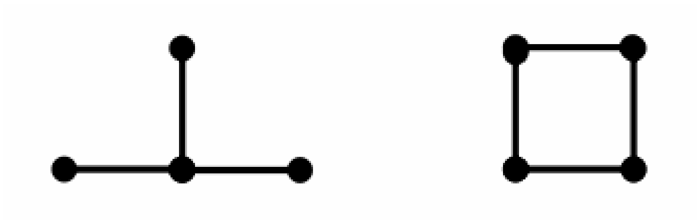

Fig. 19. Illustration for (16)

For the upper network, we have Δ = (3,1,1,1) and Γ = (6,4, 4,4) while the lower one has Δ′ = (2,2,2,2) and Γ′ = (6,6,6,6). Consequently: Γ Γ′, but neither Δ Δ′, nor Δ′ Δ holds.

In this context, we have the following open problem (OP).

OP1: does Γ = Γ′ implies Δ = Δ′ ? Even for trees, this problem is open. We conjecture that the answer is no.

8 Conclusion

We examined the degree distribution of a network and presented several examples. Utilizing the non-normalized Lorenz curve, we employed a generalized form of the majorization partial order. It is important to highlight that this represents a novel and fundamental application of the generalized Lorenz partial order. Additionally, we introduced measures, including a Gini-type index, that respect the generalized Lorenz partial order.

We introduced a new class of small worlds, namely those based on the degrees of nodes in a network. Similar to a previous study, small worlds are defined as sequences of networks with certain limiting properties. We distinguish between three types of small worlds: those based on the highest degree, those based on the average degree, and those based on the median degree. We show that these new classes of small worlds are different from those introduced previously based on the diameter of the network or the average and median distance between nodes. However, there exist sequences of networks that qualify as small worlds in both senses of the word, with stars being an example. Our approach enables the comparison of two networks with an equal number of nodes in terms of their "small-worldliness". This comparison uses generalized Lorenz curves and the corresponding notion of generalized Lorenz majorization.

Extending the idea of delta- and alpha-sequences we introduced gamma-sequences, gave examples, showed their relation with delta-sequences, and showed that there exist non-isomorphic networks with the same alpha-, delta- and gamma-sequences.

We end this article by stating two more open problems (OP):

OP2. Apply the generalized Lorenz order to the gamma-sequence.

OP3. Define and study Small Worlds in terms of the gamma-sequence.

Acknowledgment. The author thanks Li Li (Beijing) for drawing excellent illustrations, and Ronald Rousseau for stimulating discussions. He further thanks both referees for their correct and helpful remarks. He especially thanks Referee 1 for pointing out a mistake in the original version of the paper.

References

Albert, R., Jeong, H., & Barabási, A.-L. (1999). Diameter of the World Wide Web. *Nature*, 401(6749), 130-131.

Barabási, A.-L. & Albert, R. (1999) Emergence of scaling in random networks. *Science*, 286(5439), 509-512.

Chow, Y.S. & Teicher, H. (1978). *Probability Theory. Independence – Interchangeability – Martingales*. New York: Springer.

Egghe, L. (2024). Extended Lorenz majorization and frequencies of distances in an undirected network. *Journal of Data and Information Science,* 9(1), 1-10.

Egghe, L., & Rousseau, R. (2023a). Global impact measures. *Scientometrics*, 128(1), 699-707.

Egghe, L., & Rousseau, R. (2023b). Global informetric impact: A description and definition using bundles. *Journal of Informetrics*, 17(1), art. 101366.

Egghe, L. & Rousseau, R. (2024). The small-world phenomenon: a model, explanations, characterizations, and examples. Preprint: arXiv:2402.10233

Knuth, D.E. (1973). *The Art of Computer Programming (second ed.) Vol.1 Fundamental Algorithms*. Reading: Addison-Wesley.

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

Newman, M.E.J., & Watts, D.J. (1999). Scaling and percolation in the small-world network model. *Physical Review E*, 60, 7332-7342.

Rousseau, R., Egghe, L., & Guns, R. (2018). *Becoming Metric-Wise. A Bibliometric Guide for Researchers*. Kidlington: Chandos-Elsevier.

Wasserman, S. & Faust, K. (1994). *Social Network Analysis: Methods and Applications*. Cambridge: University Press.

Watts, D.J., & Strogatz, S.H. (1998). Collective dynamics of 'small-world' networks. *Nature*, 393(6684), 440-441.