Journal of Scientometric Res. 2019; 8(1):1-8
http://www.jscires.org

Research Article

Ego Citation Networks Considered as Domination Networks

Yuxian Liu[1,2], Ronald Rousseau[3,*]

[1]Tongji University Library, Tongji University, Siping Street, Yangpu, Shanghai, CHINA.
[2]School of History and Archives, Yunnan University, Kunming, CHINA.
[3]University of Antwerp, Faculty of Social Sciences, Middelheimlaan, Antwerpen, and KU Leuven ECOOM, Naamsestraat, Leuven, BELGIUM.

ABSTRACT

In this article we continue our study of power structures. A dynamic quantitative theory, including the measurement of power or dominance structures is applied to citation networks. A curve somewhat similar to the Lorenz curve for inequality measurement is used. We calculate the D-measure, a normalized measure expressing the degree of dominance in a network. The D-measure of a citation network consisting of the ego (the original article) with n references and m received citations is obtained. When n=m the network is symmetric and the D-measure is 0.5; when the article did not yet receive a citation the D-measure is n/(n+1). When the article has no references then the D-measure is 1/(m+1). A real-world case is described and the evolution of its D-measure over time is shown. Our work is one way of describing an evolving ego-citation network in a quantitative way.

Keywords: Power structure, Network dynamics, Citation graphs, Time evolution, Network properties.

Correspondence

Ronald Rousseau

Faculty of Social Sciences, University of Antwerp, Middelheimlaan 1, 2020 Antwerpen, BELGIUM.

Email: ronald.rousseau@uantwerpen.be

Received: 08-10-2018
Revised: 11-12-2018
Accepted: 28-01-2019
DOI: 10.5530/jscires.8.1.1

INTRODUCTION

It goes without saying that network theory is an essential part of contemporary science. In the field of Library and information science (LIS) article citation networks, author citation networks, author collaboration networks, bibliographic coupling and co-citation networks are among the best known.[1] When collecting data and studying networks one may focus on one particular actor and their properties within the corresponding ego network (a network build from the perspective of one actor, the ego) or one may focus on the network as a whole, including all relations of all actors belonging to this network. Of course, one may also study global properties of ego networks. A typical example of an ego network is White's description of Eugene Garfield's research network.[2] In this contribution we study a global property, called dominance, of networks, focussing on citation networks. The included real-world example is an ego network.

This article applies the notion of dominance as introduced in an earlier article.[3] In this earlier work we discussed global dominance as a special network property based on the idea of zero-sum arrays. One may say that the 'dominance' idea is a way to operationalize ubiquitous terms such as 'top', 'leading' or 'superior'. Practically, we try to contribute to a quantitative description of the changing – global - structure of an evolving citation network. Our approach starts from graph theory and can be applied to real data of any kind. At this level there is no necessity for data to conform to certain modeling aspects such as scaling, self-organization or power law behavior. Experience with such data and resulting numerical values may in the future lead to a way to distinguish between different types of evolutions of citation networks. Parts of this contribution were presented during the ISSI conference in Wuhan (China).[4]

Zero-sum arrays and D-curves

In this section we briefly recall the definitions and main results of[3] as these will be necessary to understand the developments presented further on.

Definition: arrays

If X is a (finite) array, i.e. an N-tuple ($N > 1$), then the j-th element of X is denoted as $(X)_j = x_j$, where x_j is a real number. In this article components of any array are assumed to be ranked in decreasing order.

Definition: zero-sum array

If $X = (x_1, \ldots, x_N)$ is a real-valued array, i.e. an N-tuple, such that $\sum_{i=1}^{N} X_i = 0$, then X is called a zero-sum array. The set of all zero-sum arrays is denoted as **Z.**

Constructing a pseudo-Lorenz curve for zero-sum arrays

Definitions and notations

Given a zero-sum array X, we set

$$I_+(X) = \{i \in \{1,\ldots,N\} \text{ such that } x_i > 0\}$$
$$I_0(X) = \{i \in \{1,\ldots,N\} \text{ such that } x_i = 0\} \text{ and}$$
$$I_-(X) = \{i \in \{1,\ldots,N\} \text{ such that } x_i < 0\}.$$

As in[3] we assume that X is not the trivial zero-array, hence $I_0 \neq \{1,\ldots, N\}$. This implies that $I_+(X)$ and $I_-(X)$ are always non-empty, but they may have different numbers of elements. When it is clear about which array we are talking or when it does not matter we simply write I_+, I_0 or I_-.

We note that $\sum_{i \in I_+} x_i = -\sum_{i \in I_-} x_i$. Next we put and $\Sigma_+ = \sum_{i \in I_+} x_i$ $\forall i = 1,\ldots,N : a_i = \frac{x_i}{\Sigma_+}$. With each zero-sum array X, we associate a corresponding A-array, denoted A_X and equal to $A_X = (a_1, \ldots, a_N)$. Also A_X is a zero-sum array. Related to the array X, we will further need the array Q_X, with $(Q_X)_j = q_j = \sum_{k=1}^{j} |a_k|$ where |.| denotes the absolute value of a number. Clearly $q_N = 2$.

Construction of a D-curve

D-curves (D for dominance) of a zero-sum array X, have been introduced in.[3] A D-curve of a zero-sum array X is defined as the polygonal line connecting the points

$$(0,0) \to \left(\frac{1}{N}, a_1\right) \to \ldots \to \left(\frac{i}{N}, \sum_{j=1}^{i} a_j\right) \to \ldots \to \left(\frac{|I_+|}{N}, 1\right) \to$$

$$\left(\frac{N-|I_-|}{N}, 1\right) \to \ldots \to \left(\frac{k}{N}, \sum_{j=1}^{k} |a_j|\right) \to \ldots \to (1,2)$$

where $i \in I_+$, $k \in I_-$.

This curve can be described as a function, denoted as $D_X(t)$: for $t \in [0,1]$, we have:

$$D_X(t) = \begin{cases} N a_1 t, & t \in [0, \frac{1}{N}] \\ \sum_{k=1}^{i} a_k + N a_{i+1}\left(t - \frac{i}{N}\right), & t \in [\frac{i}{N}, \frac{i+1}{N}], \quad i = 1,\ldots,N - \#(I_-) - 1 \\ \sum_{k=1}^{i} |a_k| + N|a_{i+1}|\left(t - \frac{i}{N}\right), & t \in [\frac{i}{N}, \frac{i+1}{N}], \quad i = N - \#(I_-),\ldots,N-1 \end{cases}$$

where # denotes the number of elements in a set. We see that a D-curve is partly concave (namely when the a's are positive) and partly convex (the part where the a's are negative), as illustrated in Figure 1.

If $\#(I_+) = N - \#(I_-)$ then the D-curve has no horizontal part. If $I_0 \neq \emptyset$ then it has a horizontal middle part, at height 1.

An example.

The D-curve of (6,2,0,0,-1,-1,-2,-4), N=8, has A-array (array of a-values) equal to $A = \left(\frac{6}{8}, \frac{2}{8}, 0, 0, \frac{-1}{8}, \frac{-1}{8}, \frac{-2}{8}, \frac{-4}{8}\right)$, hence connects points with ordinates $\left(0, \frac{6}{8}, \frac{8}{8}, \frac{8}{8}, \frac{8}{8}, \frac{9}{8}, \frac{10}{8}, \frac{12}{8}, \frac{16}{8} = 2\right)$.

This D-curve is shown in Figure 1.

Definition. Equivalent zero-sum arrays

Zero-sum arrays that lead to the same D-curve are said to be equivalent. The arrays (5,1,0,0,-2,-4), (10,2,0,0,-4,-8) and $\left(\frac{5}{6}, \frac{1}{6}, 0, 0, \frac{-2}{6}, \frac{-4}{6}\right)$ are equivalent. Equivalent zero-sum arrays of length N all have the same A-array. Also (4,1,0,-5) and (4,4,1,1,0,0,-5,-5) are equivalent zero-sum arrays, but with different length.

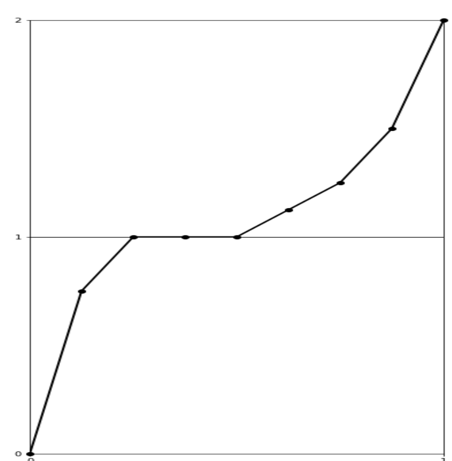

Figure 1: D- curve of the array (6,2,0,0,-1,-1,-2,-4).

Partial orders for zero-sum arrays

Definition: the dominance relation $\leq_D$ in **Z**

Let X and Y be zero-sum arrays, not necessarily of the same length, then we say that X is D-smaller than Y (or Y is D-larger than X), denoted as $X \leq_D Y$ (or $Y \geq_D X$) if, for each $t \in [0,1]$, $D_X(t) \leq D_y(t)$. X is strictly D-smaller than Y, denoted as $X <_D Y$, if, for each $t \in [0,1]$, $D_X(t) \leq D_Y(t)$ and there is at least one point t_0, hence infinitely many, where $D_X(t_0) < D_Y(t_0)$. When $X \leq_D Y$ it is clear that the D-curve of X lies nowhere strictly above the D-curve of Y. The relation $\leq_D$ determines a partial order in the set of all equivalence classes of zero-sum arrays. Formally we write:

$X \leq_D Y$ if and only if $\forall t \in [0,1] : D_X(t) \leq D_Y(t)$

Moreover, X = Y (as equivalence classes) if and only if $\forall t \in [0,1] : D_X(t) = D_Y(t)$.

As the dominance relation $\leq_D$ is only a partial order, some arrays cannot be compared: they are said to be intrinsically incomparable.

It is clear that the idea of a D-curve is inspired by the idea of a Lorenz curve, but has different properties.[5-7]

Maximum and minimum D-curves

Maximum D-curves

For fixed N the maximum D-curve occurs when (0,0) is linearly connected to the point with coordinates (1/N,1) and then linearly connected to the endpoint (1,2). This D-curve corresponds to all zero-sum arrays of the form X = (s,-t,…,-t), with s, t > 0 and s = (N-1)t.

Considering N as a variable, the line y = x + 1, passing through the points (0,1) and (1,2) is an upper bound for all these D-maximum N-curves.

Minimum D-curves

For fixed N, a minimum D-curve is obtained by connecting the origin (0,0) linearly to the point with coordinates $\left(\frac{N-1}{N},1\right)$ and then further linearly to the point (1,2). This minimum D-curve corresponds to all arrays of the form Y = (u,…, u, -v), with u, v > 0 and v = (N-1)u. The line y = x, passing through the origin and the point (1,1) is a lower bound for all these minimum D-curves. We note that if N increases then the maximum D-curve becomes larger too (in the partial order of D-curves). Similarly, minimum D-curves become smaller.

A measure respecting the dominance relation $\leq_D$ in **Z**

The area between the D-curve and the line y = x respects the D partial order. This area is denoted as $AR_D(X)$. For any zero-sum array this area takes values on the interval [0,1]. We will refer to this area as the D-measure. The D-measure is denoted AR_D and is calculated as:

$$AR_D(X) = \frac{1}{N}\left(\sum_{i=1}^{N} q_i\right) - \frac{N+2}{2N}.$$

where the q-values are the components of the array Q_X with $(Q_X)_j = q_j = \sum_{k=1}^{j} |a_k|$

An example: for X = (5,2,0,-3,-4) the D-measure is equal to:

$$AR_D(X) = \frac{1}{5}\left(\frac{5}{7}+\frac{7}{7}+\frac{7}{7}+\frac{10}{7}+\frac{14}{7}\right) - \frac{7}{10} = \frac{37}{70}$$

Maximum D-curves have D-measures equal to $\frac{N-1}{N}$, while minimum D-curves have D-measures equal to $\frac{1}{N}$. Although some D-curves are intrinsically incomparable, their D-values, being positive real numbers are always comparable.

Applications to directed networks, such as direct citation networks

Graph theory

We briefly recall the basic graph theoretical terminology needed in the sequel and note that in this article the terms graph and network are considered to be synonyms. A directed graph (in short: digraph), denoted G(V,E) consists of a set of vertices or nodes, denoted as V or V(G) and a set of edges or links, denoted as E or E(G). Nodes will be denoted by lower case letters such as i and j. An edge is an ordered pair of the form (i,j) where i and j are nodes, hence belong to V. Node i is called the initial node and node j is the terminal node of edge (i,j). A directed path, or chain, from node i to node k is a sequences of edges $(v_n)_{n=1,\ldots,M}$ such that the terminal node of edge v_n coincides with the initial node of edge v_{n+1} and such that node i is the initial node of edge v_1 and node k is the terminal node of edge v_M. If node i coincides with node k the directed path is a directed circuit or loop. A directed graph is called acyclic or loopless if it does not contain directed circuits. A directed graph is weakly connected if a path exists between any two nodes in the underlying undirected graph. We will always assume that the graphs we study have a finite number of nodes, at least two and are acyclic and weakly connected.

In[3] we introduced a local and a global dominance theory. In this contribution, we restrict ourselves to the global theory, in short GDT. In this theory we use arrays of the form $\Sigma = (\sigma_1, \sigma_2, \ldots, \sigma_N)$, defined as follows:

$$\sigma_i = \sigma_i^+ - \sigma_i^-$$

σ_i^+ = the sum of the lengths of the chains that start in node i

σ_i^- = the sum of the lengths of all chains that end in node i

From these definitions we see that also Σ leads to a zero-sum array and hence we can apply the zero-sum theory. Such a zero-sum array will be referred to as a global flow array. This array consists of global flow numbers. We want to characterize a weakly connected, acyclic network in terms of how much dominance is present.

Definitions

Definition: a local source of a digraph is a node having in-degree zero and strictly positive out-degree.

If a local source can reach any other node in a digraph it is called a network source. Since we have assumed that there are no loops in a network, we see that if a network source in a digraph exists it is necessarily unique hence it becomes the network source.

Definition: a local sink of a digraph is a node having out-degree zero and strictly positive in-degree.

If a local sink can be reached by any other node it is called a network sink. Also a network sink, if it exists, is necessarily unique and hence is referred to as the network sink.

It is obvious that a local source has a strictly positive flow number and that a local sink has a strictly negative flow number. Moreover, adding a node in a digraph which is linked only to the network source (linked from the new node to the original network source), makes this new node the network source. Next we recall the following definitions.

Definition: global dominance nodes

A node with the highest global flow in a D-graph is called a global dominance node.

Networks corresponding to maximum and minimum D-curves

For fixed N (the number of nodes), the graph shown in Figure 2 yields the only graph corresponding to a maximum global D-array.

Similarly, for fixed N, the graph shown in Figure 3 yields the only graph corresponding to a minimum D-array.

Terminology and meaning: hierarchies versus power (dominance)

Let us reconsider the digraph shown in Figure 2. There is not much hierarchical structure in this digraph, but it reflects a very strong power structure: one ruler and many equally

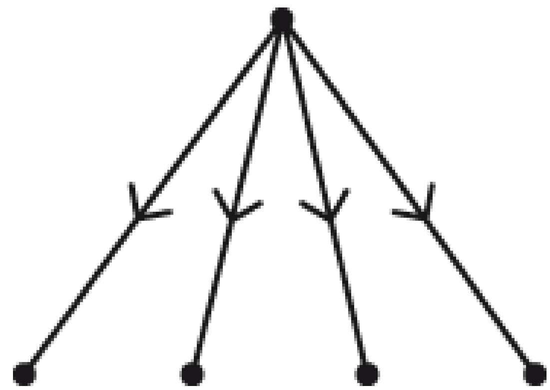

Figure 2: N-node graph corresponding to a maximum D-curve (N=5).

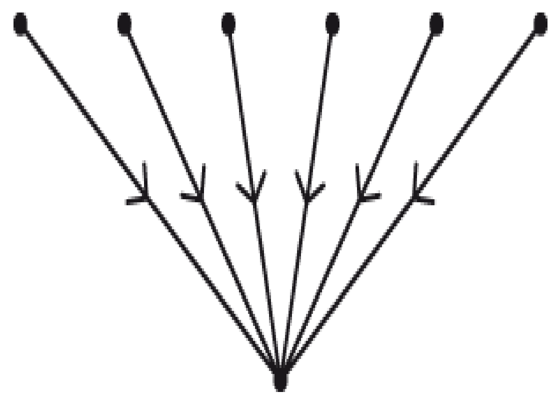

Figure 3: A graph corresponding to the minimum D-curve (N=7).

powerless subordinates. In applications of D-curves to institutes, research groups or scientists as nodes we want to gauge the power structure that is present. The more inequality among nodes with a positive flow the more powerful the order relation is. But also: the more even (in the sense of evenness)[8] the nodes with a negative flow, the more powerful the order.

Dynamic aspects of networks and properties of D-curves

When a new measure is proposed one usually derives theoretical properties and explains the possible benefits of using such a measure. This has been done in.[3] Studying dynamic aspects of networks and their corresponding dominance measures is a next and essential step for potential applications in fields such as business management, politics and social interactions. Here we restrict ourselves to applications in citation analysis.

We already know that adding a node in a digraph which dominates the network source, makes this new node the network source, hence it becomes a global dominance node. Linear structures are clear hierarchies but they are not interesting in the context of power structures: such structures are always intrinsically incomparable and because of their symmetry their D-measures are always equal to 0.5. In this article we study some aspects of evolving ego-citation networks

An ego citation network

A simple example

An author chooses which articles to include in his/her reference list. For this reason one can consider a citing article to be in a dominating position with respect to the article which receives a citation. We note though that scientifically one may argue that the cited article is the superior one as the citing article recognizes it as an authority. From the abstract point of view of the network research presented here it does not really matter which point of view is taken (it is just a matter of reversing the arrows in a directed network), but, of course, socially and emotionally it does.

We consider an article that has four references and which receives step by step more and more (direct) citations, making the structure more and more top-heavy. In this simple example we only consider the position of the original article with respect to its references and articles citing it, neglecting other nodes in the citation network. The different steps are shown in Table 1. Such a citation network always begins with a maximum D-curve.

A general formula

It is possible to derive a general formula for the global D-measure in the case that the original article has n references and received m citations, hence N = n+m+1. Note that possible relations between cited and cited, citing and citing or cited and citing articles are not taken into account in this calculation. This case is illustrated in Figure 4.

The corresponding array is:

$$X = \left(\underbrace{1+2n,\dots,1+2n}_{m\ times}, n-m, \underbrace{-(1+2m),\dots,-(1+2m)}_{n\ times}\right)$$

If m ≥ n then the |A|- array (array of absolute values of the A-array) is: $\left(\underbrace{\frac{1+2n}{(1+2n)m}}_{m\ times}, \frac{m-n}{(1+2n)m}, \underbrace{\frac{1+2m}{(1+2n)m}}_{n\ times}\right)$ and the Q-array becomes:

Table 1: A short citation history of a fictitious article with four references.

Number of citing articles	N	Array	Value of the global D-measure
0	5	(4,-1,-1,-1,-1)	0.8
1	6	(9,3,-3,-3,-3,-3)	0.708
2	7	(9,9,2,-5,-5,-5,-5)	0.621
4	9	(9,9,9,9,0,-9,-9,-9,-9)	0.5
8	13	(9,9,9,9,9,9,9,9,-4,-17,-17,-17,-17)	0.357

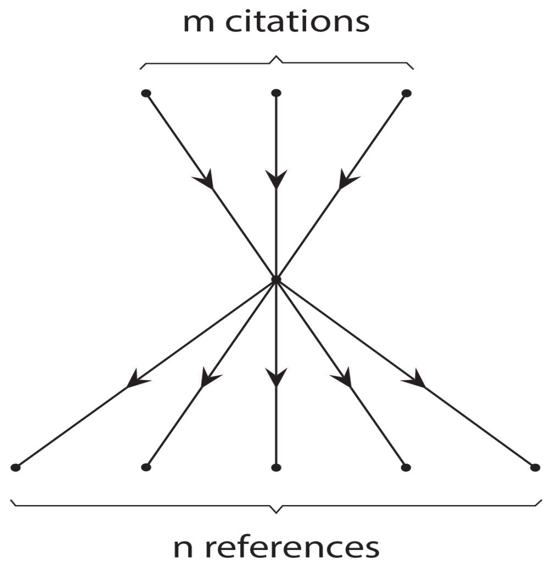

Figure 4: The general case: an article with n references and m received citations.

$$\left(\underbrace{j\left(\frac{1+2n}{m(1+2n)}\right)}_{j=1,\dots,m}, 1+\frac{m-n}{m(1+2n)}, \underbrace{\frac{m(1+2n)+m-n+k(1+2m)}{m(1+2n)}}_{k=1,\dots,m}\right)$$

$$AR_D(X) = \frac{1}{N}\left(\sum_{i=1}^{N} q_i\right) - \frac{N+2}{2N}$$

$$= \frac{1}{m+n+1}\left(\sum_{j=1}^{m}\frac{j}{m} + \left(1+\frac{(m-n)}{m(1+2n)}\right) + \sum_{k=1}^{n}\left(1+\frac{(m-n)+k(1+2m)}{m(1+2n)}\right)\right) - \frac{m+n+3}{2(m+n+1)}$$

$$= \frac{1}{m+n+1}\left(\frac{m+1}{2} + \left(1+\frac{(m-n)}{m(1+2n)}\right) + n + \frac{n(m-n)}{m(1+2n)} + \frac{(1+2m)n(n+1)}{2m(1+2n)}\right) - \frac{m+n+3}{2(m+n+1)}$$

$$= \frac{1}{m+n+1}\left(\frac{m^2(1+2n)+m(6n^2+12n+5)-n(n+1)}{2m(1+2n)}\right) - \frac{m+n+3}{2(m+n+1)}$$

$$= \frac{m(4n^2+5n+2)-n(n+1)}{2m(m+n+1)(2n+1)}$$

One can easily check that, if m = n, the value of $AR_D(X)$ is equal to 0.5. If n = 0, corresponding with an article without references, then the result is 1/ (m+1). This case clearly corresponds with a decreasing power structure for increasing m.

If now m ≤ n then the positive part is equal to n (2m+1) and

$$AR_D(X) = \frac{1}{N}\left(\sum_{i=1}^{N} q_i\right) - \frac{N+2}{2N}$$

$$= \frac{1}{m+n+1}\left(\sum_{j=1}^{m} \frac{j(1+2n)}{n(1+2m)} + \left(\frac{m(1+2n)+(n-m)}{n(1+2m)}\right) + \sum_{k=1}^{n}\left(1+\frac{k}{n}\right)\right) - \frac{m+n+3}{2(m+n+1)}$$

$$= \frac{1}{m+n+1}\left(\frac{m(m+1)(1+2n)}{2n(1+2m)} + 1 + n + \frac{n+1}{2}\right) - \frac{m+n+3}{2(m+n+1)}$$

$$= \frac{1}{m+n+1}\left(\frac{m^2(2n+1)+m(6n^2+8n+1)+3n(n+1)}{2n(1+2m)}\right) - \frac{m+n+3}{2(m+n+1)}$$

$$= \frac{m^2+m(4n^2+n+1)+2n^2}{2n(m+n+1)(2m+1)}$$

Also, here the expression becomes equal to 0.5 when m = n. If m = 0, corresponding with an article that has not been cited yet, then the result is n/ n/(n+1).

A real-world example

As a real-world example we consider the article[9] as the ego in a citation network. This article's full bibliographic data are: Empirical series of journal h-indices: The JCR category *Horticulture* as a case study

By: Liu, Yuxian; Rao, I. K. Ravichandra; and Rousseau, Ronald

Scientometrics, Volume: 80 Issue: 1 Pages: 59-74; 2009

All necessary data are provided in the Appendix. We start with the network consisting of the ego and its references (links between references are not included) and expand the network, year by year, by articles citing the ego. If these citing articles cite nodes (articles) already present in the network, these links are added too. Hence this example differs from the theoretical example in a previous section as more nodes and links are included. Data are obtained from the WoS (up to and including the year 2017). This leads to a network consisting of the ego, its 17 references and 7 citing articles. For each year we calculate the corresponding D-measure. Results are shown in Table 2. Figure 5 illustrates the final configuration, from which all other configurations can be derived.

We note that the original D-value is (N-1)/N=17/18=0.944. D-values for the whole network decrease over time: by adding new nodes on top (which have just a few links), the total network dominance decreases. Finally, the D-value for the EGO (the smallest positive value in the X-arrays) decreases from 17 in 2009 to 8 in 2017. In theory this value may even become negative, when more and more citing articles are included. We recall here that the EGO in this example is just a node used to build a network. Moreover, the D-value

Table 2: Evolution of the citation network and corresponding D-values over time.

Cumulative number of citing articles - year	N	X-array	Value of the D-measure
0 - 2009	18	$\left(17, \underbrace{-1,\cdots,-1}_{17\text{times}}\right)$	0.944
1 - 2010	19	$\left(39, 16, \underbrace{-3}_{13\text{ times}}, \underbrace{-4}_{4\text{ times}}\right)$	0.881
4 – 2011, 2012, 2013	22	$\left(39, 37, 37, 36, 13, \underbrace{-9}_{11\text{ times}}, \underbrace{-10}_{4\text{ times}}, -11, -12\right)$	0.772
5 - 2014	23	$\left(93, 39, 37, 36, 36, 10, \underbrace{-14}_{11\text{ times}}, \underbrace{-15}_{4\text{ times}}, -17, -20\right)$	0.761
6 – 2015, 2016	24	$\left(93, 39, 37, 36, 36, 36, 9, \underbrace{-16}_{11\text{ times}}, \underbrace{-17}_{4\text{ times}}, -19, -23\right)$	0.732
7 - 2017	25	$\left(93, 39, 39, 37, 36, 36, 36, 8, \underbrace{-18}_{10\text{ times}}, \underbrace{-19}_{4\text{ times}}, -20, -22, -26\right)$	0.704

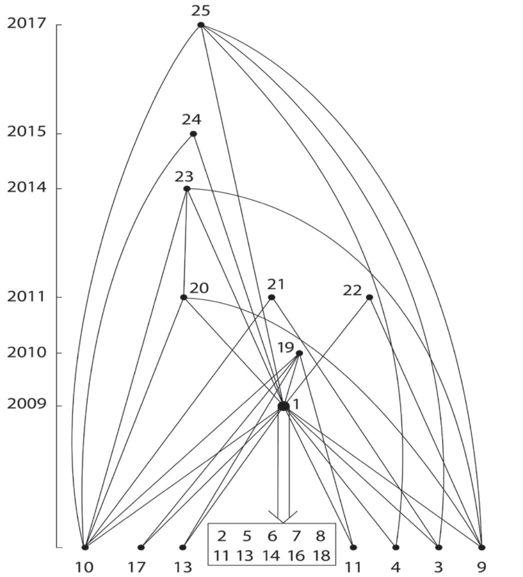

Figure 5: The final network for the real-world example (nodes 2,5,6,7,8,12,13,14,16 and 18 are represented in one box as they have the same role in the network and hence have the same flow values).

characterizes networks as a whole and has little relation with a specific node, even when this node was used to build the network.

CONCLUSION

This articles deals with an indicator, called the D-measure, related to a global property of a directed network. In particular we applied our earlier domination theory to citation networks consisting of the ego (the original article) with n references and m received citations. When n=m the network is symmetric and the D-measure is 0.5; when the article did not yet receive a citation the D-measure is n/(n+1). When the article has no references then the D-measure is 1/(m+1). The theory was illustrated by a real-world example.

Our work is one way of describing an evolving ego-citation network in a quantitative way. Practically, we try to contribute to a quantitative description of the changing – global - structure of an evolving citation network. Our approach starts from a graph-theoretical theory and applies it to real data, skipping any necessity for data to conform to certain modeling aspects such as scaling, self-organization or power law behavior. Experience with such data and resulting numerical values may in the future lead to a way to distinguish between different types of evolutions of citation networks. Being rather mathematical, we consider our contribution as reflecting an aspect of the formalization or "hardening" of the social sciences as mentioned by Herbert Simon.[10]

ACKNOWLEDGEMENT

The authors thank anonymous reviewers for thoughtful observations. They further thank Leo Egghe (University Hasselt, Belgium) for helpful discussions. This work was supported by NSSFC via 71173154, National Social Sciences Foundation of China. Finally, we also thank Mats Benner, Joshua Eykens and Raf Guns for helpful information.

CONFLICT OF INTEREST

The authors declare that their is no conflict of interest.

ABBREVIATIONS

GDT: Global Dominance Theory.

REFERENCES

1. Rousseau R, Egghe L, Guns R. Becoming Metric-Wise: A Bibliometric Guide for Researchers, Cambridge (MA) and Kidlington (UK): Chandos-Elsevier. 2018.
2. White HD. Toward ego-centered citation analysis. In: The Web of Knowledge. A Festschrift in Honor of Eugene Garfield Medford (NJ): Information Today. 2000;475-96.
3. Liu YX, Rousseau R, Egghe L. Partial orders for zero-sum arrays with applications to network theory. Journal of Informetrics. 2017;11(1):257-74.
4. Liu YX, Rousseau R. Further examples of dominance structure measurement. Conference Proceedings of the 16th Conference on Scientometrics and Informetrics, 16-20 October Wuhan University, Wuhan (China). 2017;1024-34.
5. Lorenz MO. Methods of measuring concentration of wealth. Journal of the American Statistical Association. 1905;9(70):209-19.
6. Marshall AW, Olkin I, Arnold BC. Inequalities: Theory of Majorization and its Applications (2nd.). New York: Springer. 2011.
7. Rousseau R. Lorenz curves determine partial orders for comparing network structures. DESIDOC: Journal of Library and Information Technology. 2011;31(5):340-7.
8. Nijssen D, Rousseau R, Van Hecke, P. The Lorenz curve: a graphical representation of evenness. Coenoses. 1998;13(1):33-8.
9. Liu YX, Rao IKR, Rousseau R. Empirical series of journal h-indices: The JCR category Horticulture as a case study. Scientometrics. 2009;80(1):59-74.
10. The Sverige Riksbank. 1978. https://www.nobelprize.org/nobel_prizes/economic-sciences/laureates/1978/simon-bio.html

Appendix: Articles in the real-world citation network example EGO

1. Liu, YX., Rao, I. K. R. and Rousseau, R. (2009). Empirical series of journal h-indices: The JCR category Horticulture as a case study. Scientometrics, 80(1), 59-74. EGO's references (17 articles)
2. BANKS, M. G. (2006). An extension of the Hirsch index: indexing scientific topics and compounds. Scientometrics, 69 :161–168.
3. Braun T, Glänzel W, Schubert A. A Hirsch-type index for journals. Scientometrics. 2006;69(1):169-73.
4. Burrell QL. Hirsch's h-index: a stochastic model. Journal of Informetrics. 2007;1(1):16-25.
5. Burrell QL. Hirsch's h-index and Egghe's g-index. In: D. TORRES-SALINAS, H. F. MOED (Eds), Proceedings of ISSI 2007, Madrid, CSIC. 2007;162-9.
6. Burrell QL. Hirsch index or Hirsch rate?. Some thoughts arising from Liang's data. Scientometrics. 2007;73(1):19-28.
7. Egghe L. Dynamic h-index: the Hirsch index in function of time. Journal of the American Society for Information Science and Technology. 2007;58(3):452-4.
8. Egghe L, ROUSSEAU R. Introduction to Informetrics. Amsterdam: Elsevier. 1990.
9. Egghe L, Rousseau R. An informetric model for the h-index. Scientometrics. 2006;69(1):121-9.
10. Hirsch JE. An index to quantify an individual's scientific research output. Proceedings of the National Academy of Sciences of the United States of America. 2005;102(46):16569-72.
11. Jin BH, Liang LM, Rousseau R, Egghe L. The R- and AR-indices: complementing the hindex. Chinese Science Bulletin. 2007;52(6):855-63.
12. Kelly CD, Jennions MD. The h index and career assessment by numbers. Trends in Ecology and Evolution. 2006;21(4):167-70.

13. Liang LM. H-index sequence and h-index matrix: constructions and applications. Scientometrics. 2006;69(1):153-9.
14. Liu YX, Rousseau R. Hirsch-type indices and library management: the case of Tongji University Library. In: D. TORRES-SALINAS, H. F. MOED (Eds), Proceedings of ISSI 2007, Madrid, CSIC. 2007;514-22.
15. Liu YX, Rousseau R. Definitions of time series in citation analysis with special attention to the h-index. Journal of Informetrics. 2008;2(3):202-10.
16. Rao IKR. Distributions of Hirsch-index and g-index: an empirical study, In: D. TORRES-SALINAS, H. F. MOED (Eds), Proceedings of ISSI 2007, Madrid, CSIC. 2007;655-8.
17. Rousseau R. A case study: evolution of JASIS' h-index. Science Focus. 2006;1(1):16-7. (in Chinese). English version available at: E-LIS: ID-code 5430.
18. The Stimulate-6 Group. The Hirsch index applied to topics of interest to developing countries. First Monday. 2007;12(2). http://www.firstmonday.org/issues/issue12_2/stimulate/Articles citing the EGO (7 articles, mentioned chronologically)
19. Mahbuba D, Rousseau R, Srivastava D. A scientometric analysis of health and population research in South Asia: focus on two research organizations. Malaysian Journal of Library and Information Science. 2010;15(3):135-47.
20. Egghe L, Guns R, Rousseau R. Thoughts on uncitedness: Nobel laureates and Fields medalists as case studies. Journal of the American Society for Information Science and Technology. 2011;62(8):1637-44.
21. Hodge DR, Lacasse JR. Ranking disciplinary journals with the Google Scholar h-index: A new tool for constructing cases for tenure, promotion and other professional decisions. Journal of Social Work Education. 2011;47(3):579-96.
22. Lin F. A study on power-law distribution of hostnames in the URL references. Scientometrics. 2011;88(1):191-8.
23. Yuan XY, Hua WN, Rousseau R, Ye FY. A preliminary study of the relationship between the h-index and excess citations. Canadian Journal of Information and Library Science – Revue canadienne des Sciences de l'Information et de Bibliothéconomie. 2014;8(2):127-44.
24. Hsu WC, Tsai CF, Li JH. A hybrid indicator for journal ranking: An example from the field of Health Care Sciences and Services. Online Information Review. 2015;39(7):858-69.
25. Bertoli-Barsotti L, Lando T. A theoretical model of the relationship between the h-index and other simple citation indicators. Scientometrics. 2017;111(3):1415-48.