Scientometrics
https://doi.org/10.1007/s11192-022-04414-6

Connections matter: a proxy measure for evaluating network membership with an application to the Seventh Research Framework Programme

Matteo Cinelli[1,2] · **Giovanna Ferraro**[3] · **Antonio Iovanella**[4]

Received: 18 August 2021 / Accepted: 27 April 2022
© The Author(s) 2022

Abstract
Although the topic of networks has received significant attention from the scientific literature, it remains to be seen whether it is possible to quantify the degree to which an organisation benefits from being part of a network. Starting from the concept of network value and that of Metcalfe's Law, this paper introduces and defines the collective network effect (CNE). CNE is based on the concept that a network member is not only affected by its friends but also by the friends of its friends. By taking into account network connection patterns, CNE provides a proxy for quantifying the benefit of network membership. We computed the CNE for the nodes of a large network built using the whole set of common projects among the participants of the 7th Framework Programme for Research and Technological Development of the European Commission. The obtained results show that nodes with a higher CNE have access to substantially more conspicuous fundings than nodes with a lower CNE. In general, such a measure could supplement other centrality measures and be useful for organisations and companies aiming to evaluate both their current situation and the potential partners they should link with in order to extract the highest benefits from network membership.

Keywords Network membership · Collective network effect · Metcalfe's Law · Research funding · European projects

✉ Matteo Cinelli
matteo.cinelli@uniroma1.it

Giovanna Ferraro
giovanna.ferraro@unipr.it

Antonio Iovanella
antonio.iovanella@unint.eu

[1] ISC-CNR Uos "Sapienza", Via dei Taurini, 19, 00185 Rome, Italy

[2] Department of Computer Science, Sapienza University of Rome, Viale Regina Elena, 295, 00161 Rome, Italy

[3] Department of Engineering and Architecture, University of Parma, Parco Area delle Scienze 181/A, 43124 Parma, Italy

[4] School of Economics, Università degli Studi Internazionali di Roma - UNINT, Via Cristoforo Colombo, 200, 00147 Rome, Italy

Published online: 18 June 2022

Introduction

Being part of a network is generally considered beneficial for companies and organisations; a notation that is reflected by several real-world examples as well as by the increasingly widespread interest of the scientific community in investigating the effects of network membership. The importance of network membership has been discussed in terms of its beneficial role for network organisation and management (Choi et al., 2010; Ebers, 2015; Etemad et al., 2001), in terms of its capability in fostering firms' ability to sustain and develop their competitive advantages (Ford et al., 2011) and in terms of its importance for accessing the fundamental resources needed to survive the market, as in the case of small and medium-sized enterprises (SME) (Cova et al., 1994; Henrikki, 1998; Partanen et al., 2008).

The benefits of network membership have been widely investigated also from a structural point of view. For instance, the effect of the network structure on certain functions was studied in (Choi et al., 2010; Cinelli et al., 2019) and (Cowan & Jonard, 2004) in order to shed lights on processes such as the development and the diffusion of innovation. The intensity of collaborations among members was also studied from a structural point of view considering centrality measures, focusing on the effects of bridging members (Bergé et al., 2017), in studying collaborative regional networks (Calignano et al., 2019; Calignano & Quarta, 2015; Hazır et al., 2018) and collaboration paths through layered networks (Liu et al., 2013). Nevertheless, there is still a lack of a quantitative perspective capable of providing the tools to measure the benefits of network membership.

The most remarkable effort in such a direction was that of computing the value of the network under the assumption that the higher the value of the whole network the higher the benefit for a node to be part of it. The rule used for quantifying the network value was called the Metcalfe's law (Gilder, 1993).[1] Such a law derives from the world of telecommunication networks, e.g., the ethernet, where all the nodes are equally important in terms of their potentiality to be connected to others (Odlyzco & Tilly, 2005). Relying on such an assumption, the Metcalfe's law estimates the network value proportional to the total number of possible connections, that is asymptotically proportional to the square of the network size.

In response to certain criticisms regarding the simplicity of Metcalfe's law, other laws such as Sarnoff's law (linear growth of network value) and Reed's law (Reed, 1999) (exponential growth of network value), and Odlyzko's law (Briscoe et al., 2006), have been proposed. Notwithstanding its simplicity and certain limitations (Swann, 2002; Briscoe et al., 2006), Metcalfe's law remains a reasonable rule of thumb (Madureira et al., 2013; Van Hove, 2016), having been used in many practical applications such as figuring out Facebook's network value (Metcalfe, 2013; Zhang et al., 2015) and, more recently, estimating the values of cryptocurrency networks (Alabi, 2017) such as Bitcoin, Ethereum and Dash.

Essentially, the main limitations of the Metcalfe's law are two: the potential value of a network is a biased measure of its actual value since real networks are not fully connected; the assumption that all the nodes contribute to the total value of the network equally is not realistic. These two limitations are quite evident, for instance, in the application of Metcalfe's law to Facebook in which the users are neither fully interconnected nor identical.

[1] The name "Metcalfe's law" derives from the work of (Gilder, 1993) that mentioned the intuition that Robert Metcalfe had about 10 years before.

A partial solution to the aforementioned biases consists in considering the contribution of each node as proportional to its number of connections (Barabási, 2016). However, it still doesn't cover an important issue that is considering the quality of connections. For example, being connected only to hubs—i.e., nodes with many connections—strongly differs to being connected only to spokes—i.e., nodes with few connections. The concept of quality of a node is not only attributable to its number of connections, but also to some other aspects such as the experience, the performance, the ability in acquiring more resources and on the reputation, as shown in (Bol et al., 2018; Egghe et al., 2013; Liao, 2021).

In order to be able to provide a proxy measure for the benefits of the network memberships that doesn't suffer from the same issues of network value, we propose a new measure called the *collective network effect*, that extends the discussion introduced in (Arpetti & Iovanella, 2020).

Starting from the concept of network value, as described by the Metcalfe's law and by its recent improvements, we extend it by considering the heterogeneity of the network structure in order to provide a measure that focuses on the node-level perspective. Considering the set of the node's connections in combination with the connections of its neighbours results in a measure that can be considered a proxy of the network benefit rather than the network value. Interestingly, the integral of the collective network effect (CNE) goes far beyond the value estimated by the actual number of connections; an aspect in line with the criticism against Metcalfe's assumption of connection homogeneity (Odlyzco & Tilly, 2005) and somewhat closer to the view of a network as a complex system where the total is more than the sum of its parts.

From a technical perspective, our measure, despite being correlated with other centrality measures, could be easily integrated in a wider framework aimed at evaluating the prominence of actors in social networks.

As a proof of concept, we apply the proposed measure to quantify the benefit for each organisation in joining the 7th Framework Programme for Research and Technological Development (FP7) in terms of network membership. FP7—one of the main research policy instruments of the European Commission—came into effect in 2007 and remained operative until 2013, bringing together all the initiatives designed to consolidate the European Research Area and to promote strategic actions for achieving targets in scientific excellence, growth, competitiveness and employment. FP European programmes were largely investigated under the network point of view (Balland et al., 2019; Heller-Schuh et al., 2011). Specifically, FP7 dataset was largely investigated under different perspectives; for instance, the degree of participation in the network was analysed in (Calignano, 2021), the evolution of the EU research network across countries was investigated in (Balland et al., 2019; de Arroyabe et al., 2021).

Using FP7 dataset, we build up a network in which nodes are organisations and two organisations are linked if they co-participated in at least one funded project. We find that the CNE is a good proxy for the quality of the nodes neighbourhood. Indeed, by fixing the number of connections, nodes with CNE higher than expected have access to more conspicuous fundings. Such an effect is disproportional to the number of connections the node has; in other words, weakly connected nodes are more likely to benefit from having an high network effect than hubs. In general, the collective network effect could be acknowledged as an additional measure useful for institutions and companies aiming to evaluate both their current situation and the potential partners to be linked with in order to extract the highest benefits from network membership.

The paper is organised as follows: "Methodological prerequisites" section contains some methodological prerequisites: "Collective network effect" section presents the definition of CNE; "Empirical application" section includes the analysis on the FP7 network; and "Conclusions" section gives some final remarks and suggests directions for future research.

Methodological prerequisites

It is a settled practice to consider a graph $G = (V, E)$ as the mathematical abstraction for a network; in such representation is intended that V is the set of n nodes and E is the set of m links representing the relationships among the nodes. For example, if we consider an inter-organisational network, the set V contains the organisations in the system, while the set E contains all collaborations among such organisations.

A node in V is generally indicated by its index i, while the set of links in E is given by the adjacency matrix $A = (a_{ij})_{i,j=1,\dots,n}$, in which the generic element a_{ij} is equal to 1 if the link between i and j exists or 0, otherwise.

In this paper, we take into consideration undirected and unweighted networks, hence $a_{ij} = a_{ji} \in \{0, 1\}$, for each $i, j = 1, \dots, n$. In network analysis, a fundamental quantity is given by the number of relationships owned by a node, i.e. by the number of nodes that are present in the *neighbourhood* of i—in symbols, $N(i)$. For a generic node i, such a quantity is measured by the *degree centrality* indicated as k_i.

As the network science approach includes Social Network Analysis (Scott & Carrington, 2011), this paper considers some of the main centrality measures, which represent the relative importance of a node within a network, with the assertion that the higher the centrality index of a node, the higher its perceived centrality in the graph. Moreover, centrality measures assess the involvement of nodes in contributing to the cohesiveness of the network (Borgatti & Everett, 2006; White & Harary, 2001).

The concept of centrality has an inherent ambiguity; there is no point in including all measures in one method (Rowley, 1997). Deciding which option to choose requires consideration of the specificity of the measures and the requirements of different applications. There are several quantities describing the centrality that depend on the type of statistics on which they are based; the most commonly used are reported in Table 1.

Beside such measures, another useful one that represents the overall importance of the neighbourhood of node i is the average degree of $N(i)$ that is computed as (Pastor-Satorras et al., 2001):

$$k_{nn,i} = \frac{1}{k_i} \sum_j k_j a_{ij} \tag{1}$$

Such a quantity can be also generalised in order to quantify the overall prominence of all the nodes with degree value k using the following equation:

$$k_{nn}(k) = \frac{\sum_{i:k_i=k} k_{nn,i}}{\sum_{i:k_i=k} 1} \tag{2}$$

The trend of $k_{nn}(k)$ characterises two distinctive structural organisations of networks in which different degree sets show diverse features in the local connectivity structure. In particular, when $k_{nn}(k)$ displays an increasing trend, then the network exhibits *assortative mixing*, that is high-degree nodes tend to link with other high-degree nodes. Conversely, a

Table 1 A short glossary of centrality measures

Measure	Definition	Meaning
Degree centrality (k_i)	The number of links incident upon a node, which can be interpreted as the neighbourhood size of each member within the network.	This value highlights the immediate risk of a node catching whatever is flowing through the network. It quantifies how well it is connected to the other elements of the graph. The degree centrality is an indicator of the spread of node connectivity along the graph and is a crucial gauge in defining the network organisation
Closeness centrality (C_C)	The natural distance between all pairs of nodes defined by the length of their shortest paths. Thus, the more central a node is, the lower its distance is to all other nodes.	This value measures how long it takes to spread information from a member to all others sequentially
Betweenness centrality (C_B)	The number of times a node acts as bridge along the shortest path between two other nodes.	This measure reveals the intermediary members that are essential for connecting different regions of the network
Eigenvector centrality (C_E)	The influence of a node in a network according to the number and the quality of its connections.	Indeed, a node with a smaller number of high quality links has more power than one with a larger number of mediocre contacts
Clustering coefficient (C_i)	For any node i, is the fraction of the connected neighbours of i.	It determines the capacity of link creations among neighbours, i.e. the tendency in the network to create stable groups

Please refer to (Scott & Carrington, 2011) for a complete description and formulation

Table 2 Table of notation

Symbol	Meaning
k_i	Degree of node i
$N(i)$	Neighbourhood of node i
$k_{nn,i}$	Average degree $N(i)$
$k_{nn}(k)$	Average degree of nodes with degree k
k_{nn}	Average degree of the neighbourhood for non-correlated networks
NS_i	Network share of the node i
CNE_i	Collective network effect of node i
$CNE_i(k)$	Collective network effect of node i with degree k
NE_G	Network effect of network G
$\sqsubseteq$	Metcalfe's network value

decreasing trend of $k_{nn}(k)$ indicates *disassortative mixing*, that is high-degree nodes tend to connect to low-degree nodes. Empirical networks display degree-degree correlation, for instance social networks display positive degree correlation (Newman, 2002), while technological networks display somewhat negative correlation (Newman, 2003).

When a network is neither disassortative nor assortative, there is not degree-degree correlation and it holds that (Catanzaro et al., 2005):

$$k_{nn} = \frac{\langle k^2 \rangle}{\langle k \rangle} \tag{3}$$

$\langle k \rangle$ being the average degree value and $\langle k^2 \rangle$ the average of the squared degree values.

Differently from the empirical case of Eq. (1), k_{nn} is independent from the node's degree but dependent exclusively on $\langle k \rangle$ and $\langle k^2 \rangle$ that are global network's characteristics. In addition to being straightforwardly computable, such measures give a global view of the network, making k_{nn} a simple and powerful overall measure.

Note that k_{nn} is greater than the average degree of a node. In fact, $\langle k^2 \rangle / \langle k \rangle - \langle k \rangle = (\langle k^2 \rangle - \langle k \rangle^2)/\langle k \rangle = \sigma^2/\langle k \rangle > 0$, since: (i) the variance is positive unless the network displays the same degree for all its nodes; and (ii) $\langle k \rangle$ is greater than zero unless all the nodes have zero degree.

Therefore, $\langle k^2 \rangle / \langle k \rangle > \langle k \rangle$, revealing a scenario named the *friendship paradox*, according to which 'your friends have more friends than you do' (Feld, 1991). This bias is caused by an over-representation of high-degree versus low-degree nodes.

For the sake of completeness, in Table 2 we resume the notations used thus far and hereafter.

Collective network effect

In this section, we introduce the CNE starting from the concept of 'network value' defined by Metcalfe's law. It states that the value $\sqsubseteq$ of a network G is proportional to the square of the number of nodes, i.e. $\sqsubseteq \propto n^2$ (Gilder, 1993). This definition originates from the context of the ethernet networks, and considers that nodes are potentially connected by mutual relationships; that is, the n nodes are linked to other $n-1$ nodes; in such a case, each node

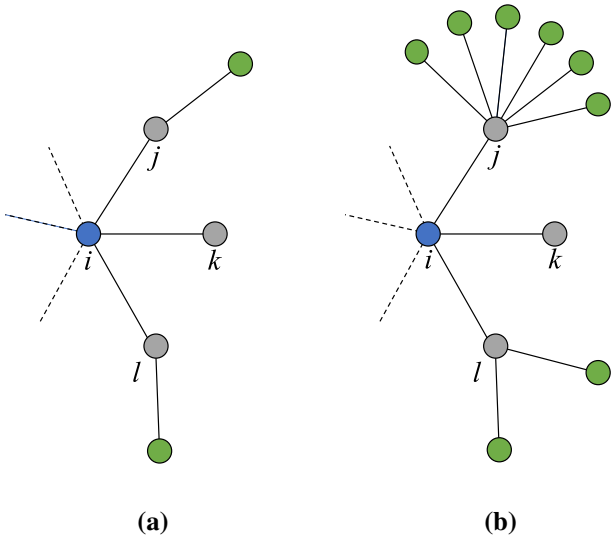

Fig. 1 Schematic representation of the neighbourhood of node i when its neighbours display different shares of the network

has degree $\binom{n}{2}$, i.e., of the same order of n^2. Therefore, the potential value of the network maps into the case of a complete network, i.e. one with density $d = \frac{2m}{n(n-1)} = 1$.

In real applications, networks are rarely densely connected, meaning that estimating the current value (as much as the potential value) of a network as $\frac{n(n-1)}{2}$ appears unrealistic. In our setting, it is reasonable to consider that not all nodes are mutually linked to one another. Therefore, we consider the whole set of connections that are present in a certain moment rather than all the possible relationships.

By means of the handshaking lemma, a well-known result in graph theory (e.g. see Bollobás, 2013), p. 4), we obtain that the sum of all networks' degrees is equal to twice the numbers of links, i.e. $\sum_{i \in V} k_i = 2m$. Therefore, it is possible to redefine Metcalfe Law's in order to estimate the current value of a network, such as $\sqsubseteq \propto 2m$. Note that the value $\sqsubseteq \propto n^2$ remains still valid for a complete network.

Additionally, the global network value can be dissected in its nodal contributions, meaning the Network Share *NS* of the generic node i is:

$$NS_i = \frac{k_i}{2m} \tag{4}$$

Even if the network share seems well defined, it fails to capture a fundamental aspect of being part of a network; that is, the quality and the influence of the neighbourhood. We can reasonably assume that the benefits for a node of being part of a network are not only given by its network share but also by the share of its neighbours. For example, as shown in Fig. 1, assuming that (a) and (b) are different portions of a network having m links, than node i displays the same *NS* in both cases, while the *NS* of j, k and l differs in the two cases. In case (b) nodes j and l have higher network share than in case (a), thus node i is

characterised as having a higher value of the average degree of neighbourhood $k_{nn,i}$. Therefore, we can easily state that node i may benefit more from being part of the network in case (b) than in case (a).

In such a vein, it is possible to introduce the collective effect of neighbours (referred to as the collective network effect CNE_i) of a node i, as the sum of its network share NS_i with the network share NS_j of all the nodes j in its neighbourhood $N(i)$:

$$CNE_i = NS_i + \sum_{j \in N(i)} NS_j = \frac{k_i}{2m} + \sum_{j \in N(i)} NS_j = \frac{k_i}{2m} + \frac{1}{2m} \sum_{j \in N(i)} k_j \tag{5}$$

Equation (5) can be rewritten using the identity $\sum_{j \in N(i)} k_j = k_i \cdot k_{nn,i}$, that is k_i multiplied by the average degree of a neighbour. Thus:

$$CNE_i = \frac{k_i}{2m} + \frac{k_i}{2m} k_{nn} = \frac{k_i}{2m}(1 + k_{nn,i}) \tag{6}$$

Similarly to the case of Eq. (1), we can generalise Eq. (6) considering all the nodes i with fixed k using Eq. (2), obtaining:

$$CNE_i(k) = \frac{k_i}{2m} + \frac{k_i}{2m} k_{nn}(k_i) = \frac{k_i}{2m}\left(1 + k_{nn}(k_i)\right) \tag{7}$$

In this case, the CNE is a function of a given degree. Note that for a given degree k, $CNE_i(k)$ is the value averaged on all CNE_i of nodes i having degree k.

Also in this case, we exploit the properties of the configuration model, holding for $n \to \infty$, by substituting in Eq. (6) the result of Eq. (3) obtaining:

$$CNE_i = \frac{k_i}{2m}\left(1 + \frac{\langle k^2 \rangle}{\langle k \rangle}\right) \tag{8}$$

Overall, by looking at Eqs. (6) and (8), for a node i we can state that its CNE is given by the value of its network share NS_i multiplied by a value that captures the level of connectedness given either as the average of the neighbours' degrees or as a measure of the whole network. In that latter case, it represents a specific measure concerning the entire network G. Indeed, we call such a measure *Network Effect* NE_G of the network G:

$$NE_G = \left(1 + \frac{\langle k^2 \rangle}{\langle k \rangle}\right) \tag{9}$$

Since $\langle k^2 \rangle / \langle k \rangle > 0$ and with the exclusion of the networks with all nodes with the same degree, we have that $NE_G > 1$. NE_G is a baseline value that expresses the benefit of being part of a network and it can exceed or be disregarded by empirical values of CNE_i depending on the heterogeneity of the neighbours of i.

In a complete network—where each node is linked to the others—the number of connections is $m = n(n-1)/2$, and for each node we obtain that $NS_i = 1/n$, while $CNE_i = 1$. Indeed, in this case, each node has degree $k_i = n - 1$, then $NS_i = k_i/2m = (n-1)/n(n-1) = 1/n$, while $\langle k^2 \rangle = (n-1)^2$ and $\langle k \rangle = n - 1$ thus $CNE_i = (1/n)(1 + (n-1)^2/(n-1)) = 1, \quad \forall i$.

Such a result means that when all the nodes are mutually connected, then the CNE on each node reaches the maximum and is equal across all nodes. In other terms, each node is affected by any other in the network.

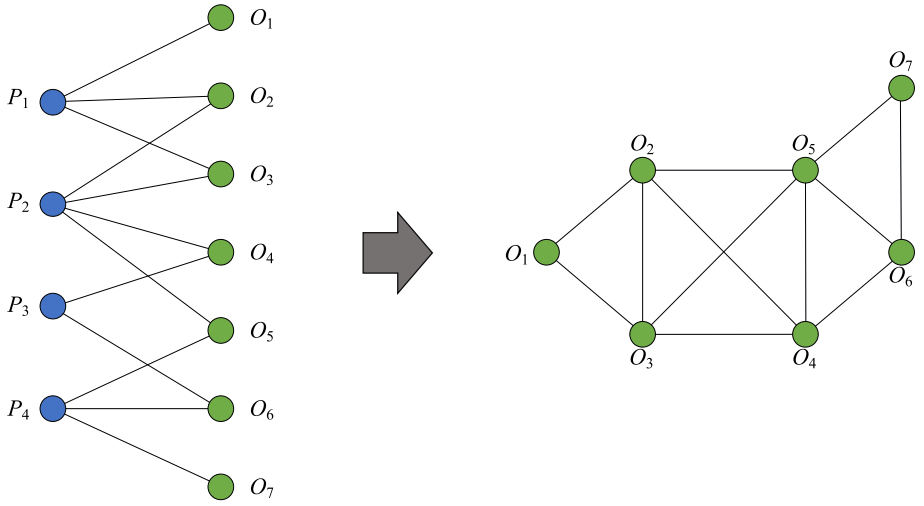

Fig. 2 An example of bipartite graph with four projects (P_1, P_2, P_3, P_4) and the corresponding participating organisations ($O_1, O_2, \ldots, O_7$) (left panel); the resulting one-mode projection (right panel)

Empirical application

We test the effectiveness of the *CNE* by using data provided by the European Commission regarding the 7th Framework Programme for Research and Technological Development (FP7). FP7 had the active strategic objective of fostering scientific and technological development across Europe and it was active from 2007 to 2013 with a total budget of over €50 billion.[2]

FP7 data comes in the form of table with columns listing project acronyms and respective participants and its network properties have been matter of recent investigations (Heller-Schuh et al., 2011; de Arroyabe et al., 2021). Such columns are employed in order to realize a bipartite network in which one partition is made up of financed projects while the other is made up of participants to such projects (see an example in Fig. 2). A link between the partitions exists if an organisation participated in a project. The bipartite network is then projected onto one of its partitions (by means of an operation called one-mode projection (Newman, 2018)) thus obtaining another network in which two organisations are connected if they participated in the same project (connections are considered without weights). The projected network has $n = 293{,}85$ nodes and $m = 685{,}319$ links and we consider as node metadata (i.e. non-structural information about the node) the contribution of the European Commission to each organisation, measured in euros. The FP7 network displays heterogeneous degree distribution—having a shape close to a lognormal distribution—indicating the presence of hubs as shown in Fig. 3. Hub nodes are actually very active institutions that participated in a high number of projects throughout the programme, as reported in Table 3.

[2] European Commission Research & Innovation, FP7 in brief: https://wayback.archive-it.org/12090/20191127213419/https:/ec.europa.eu/research/fp7/index_en.cfm; accessed: October 27, 2020.

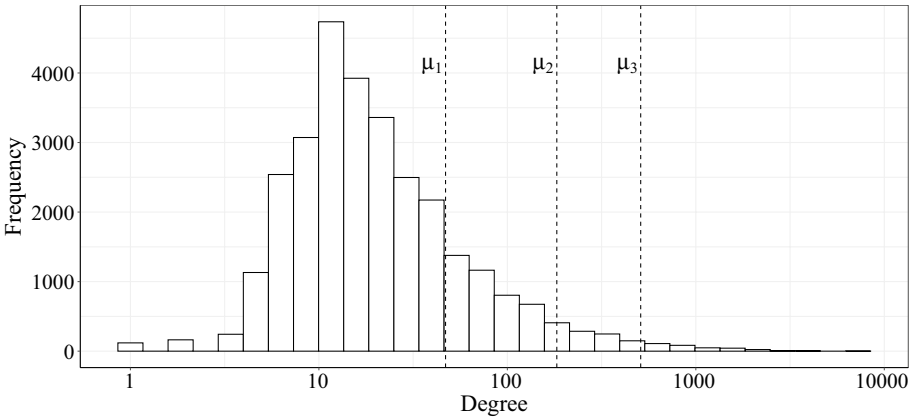

Fig. 3 Degree distribution for FP7 Cordis network. Thresholds μ_1, μ_2 and μ_3 are three mean degree values that are used as thresholds as indicated in Table 4. In more detail, the value μ_1 is the average degree of all the nodes in the network, the value μ_2 is the average degree of all the nodes in the network that have degree higher than μ_1 and μ_3 is the average degree of all the nodes in the network that have degree higher than μ_2

Table 3 The first ten institutions in the network of FP7 ranked by their degree

Institutions	Country	Degree	CNE_i
Fraunhofer-Gesellschaft zur Förderung der angewandten Forschung e.V.	DE	7261	0.624
Centre National de la Recherche Scientifique—Cnrs	FR	4557	0.564
Consiglio Nazionale delle Ricerche	IT	3956	0.533
Commissariat á L'Énergie Atomique et aux Énergies Alternatives	FR	3842	0.533
Nederlandse Organisatie Voor Toegepast-Natuurwetenschappelijk Onderzoek—TNO	NL	3469	0.458
Agencia Estatal Consejo Superior de investigaciones Científicas	ES	3286	0.491
Danmarks Tekniske Universitet	DK	2870	0.454
Fundacion Tecnalia Research & Innovation	ES	2852	0.399
Katholieke Universiteit Leuven	BE	2838	0.445
Teknologian Tutkimuskeskus VTT	FI	2762	0.412

In Fig. 4 we report the distribution of CNE_i for the FP7 network, noting how the CNE_i positively correlates with the node degree. It follows that most of the nodes with low degree values display a low CNE. In spite of this aspect, it is indeed admissible that a node i with degree k can assume values of CNE_i greater than any node of degree at least $k+1$. In other words, having a lower number of connections doesn't imply having a lower value of the CNE. Additionally, we note a relatively strong scattering of the values of CNE_i around their average value given by $CNE_i(k)$, as shown in the inset of Fig. 4, meaning that some nodes display a CNE_i higher than expected while others display a CNE_i lower than expected. This means that not all nodes are alike in terms of the quality of their neighbourhoods; indeed, some outperform the average while others do not.

Finally, as reported in Table 3, the rankings of nodes with respect to degree and CNE do not match perfectly. The two variables are not perfectly correlated: the *CNE* is capable of capturing patterns of connections that goes beyond the first neighbours. Therefore, the

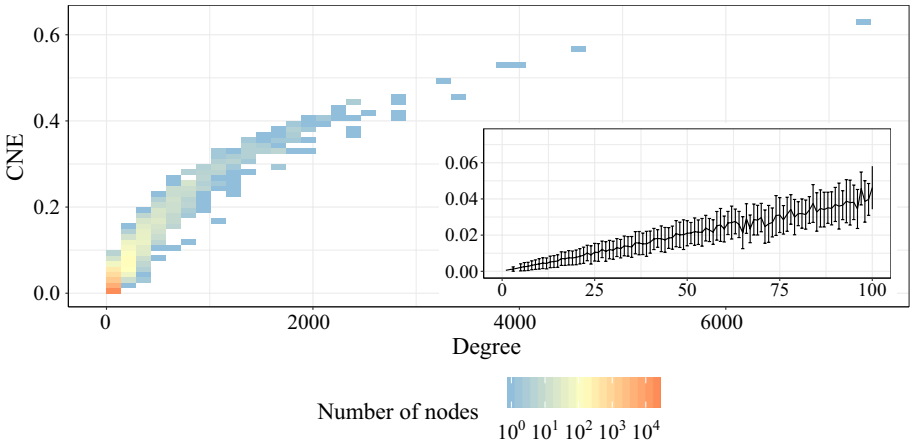

Fig. 4 Scattering of the values of CNE_i for nodes in the FP7 network. The inset shows the values of CNE_i for degree until $k_i = 100$, showing that some nodes display a CNE_i higher than expected while others display a CNE_i lower than expected

value of the collective network effect is not strictly dependent on the number of connections retained by the node.

Association between CNE and funding

The evaluation of business performance is generally a complex task (Chandler & Hanks, 1993). However, the scientific literature concerned with the study of funded projects, generally agrees in considering funding a key indicator in association with other relevant aspects such as: collaboration patterns (Defazio et al., 2009; Ma et al., 2015), innovation and technology transfer (de Arroyabe et al., 2021), growth (Piekkola, 2007; Staniewski et al., 2016) and other elements related to performance (Cooke & Wills, 1999).

In order to test the capacity of the proposed measure to quantify the benefits of being embedded in a high-quality collaboration pattern, we correlate the CNE with the contribution (in Euros) that nodes received from the European Commission for participating in projects, by answering to the following question:

- *Do nodes with CNE_i higher than expected receive, on average, a higher amount of funds than nodes with CNE_i lower than expected*?

In other words, we aim at validating the CNE_i by testing if it can be considered a good proxy for discerning nodes that benefited significantly from being part of their neighbourhood.

In order to answer the aforementioned question in a favorable manner still distinguishing between different classes of nodes we cluster them by their degree values and then we compute the aforementioned probability within each cluster. The clustering is performed by means of a non-parametric method developed in the field of bibliometrics to identify disjoint sets of academics by partitioning the distribution of their number of publications. The method, introduced in Schubert et al. (1987) and referred to as Characteristics Scores and Scales (CSS), can be consistently applied to partition distributions that are not necessarily related to the number of publications (Cinelli, 2020).

Table 4 Classes of degrees and relative results for *Pr*

Classes	Characteristic scores	No. of nodes in the cluster	No. of over-performing nodes	*Pr*
LK	$k_i \leq 47$	24,068	10,637	0.94
AK	$48 \leq k_i \leq 183$	4119	2041	0.67
HK	$184 \leq k_i \leq 510$	857	383	0.53
VHK	$k_i > 510$	341	50	0.55

Note that being the network connected the class ZK is empty

The technique involves reiterated truncation of a frequency distribution according to mean values μ called "characteristic scores". After truncating the overall distribution at its mean value, the mean of the sub-population above the first mean is recalculated; the sub-population is again truncated, and so on until the procedure is stopped (Abramo et al., 2017). Applying the CSS method, with three characteristic scores, the following five classes of nodes C can be obtained:

- Zero degree (ZK): $k = \mu_0 = 0$
- Low degree (LK): $\mu_0 < k \leq \mu_1$
- Average degree (AK): $\mu_1 < k \leq \mu_2$
- High degree (HK): $\mu_2 < k \leq \mu_3$
- Very high degree (VHK): $k > \mu_3$

The different values of μ and the resulting clusters are reported in Fig. 3.

The aforementioned research question can be answered by computing, for each class C, the probability that the average contribution $\overline{EC}_C$ received by over-performing nodes, i.e. those with $CNE_i > CNE_i(k)$, is higher than the average contribution received by under-performing nodes, i.e. those with $CNE_i \leq CNE_i(k)$. In formula:

$$Pr\left(\overline{EC}_C(CNE_i > CNE_i(k)) > \overline{EC}_C(CNE_i \leq CNE_i(k))|k \in C\right) \quad \forall\, i, k, C \tag{10}$$

In Table 4 we summarise the results for *Pr* where column 1 reports the classes, column 2 the different values for the characteristic scores μ_1, μ_2 and μ_3, column 3 the number of nodes falling in the corresponding cluster, column 4 the number of over-performing nodes, i.e. nodes with $CNE_i > CNE_i(k)$, and, finally, in column 5 the values of *Pr*.

We note that low-degree nodes with $CNE_i \geq CNE_i(k)$ are almost sure to receive a higher amount of funds than nodes with $CNE_i < CNE_i(k)$. We also note that over-performing with respect to CNE_i almost guarantees access to better resources. Therefore, CNE_i is a good proxy for the quality of the neighbourhood and consequently for the benefit of network membership.

When the degree starts to grow, the effect of CNE_i increasingly weakens until randomness (i.e., *Pr* 0.5), suggesting that other factors beyond the quality of the neighbourhood increasingly gain importance.

In general, from the analysis of the FP7 network, we can conclude that in networks of organisations small nodes can actually benefit from careful selection of their

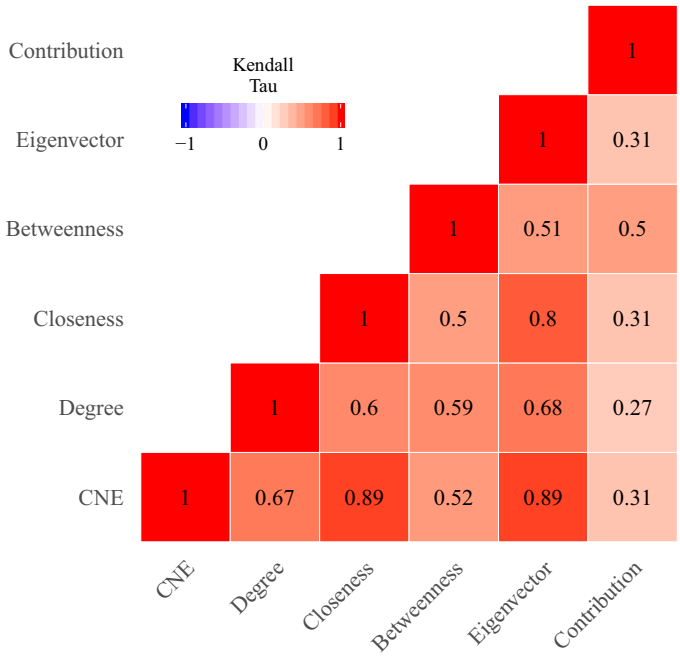

Fig. 5 Correlation matrix between different centrality measures

connections, providing further evidence to the notion that moving first within an environment is fundamental for gaining a competitive advantage.

Comparison with other centrality measures

In this section, we provide quantitative evidence for the difference among the CNE and other centrality measures by computing the concordance between the rankings provided by such measures using the Kendall rank correlation coefficient:

$$\tau = \sum_{i<j} \frac{\text{sgn}(x_i - x_j)\text{sgn}(y_i - y_j)}{\binom{n}{2}}, \tag{11}$$

where the vectors x and y are the rankings provided by the centrality measures and n is the length of the two vectors. The Kendall τ lies in the range $[-1, 1]$ and it has value 1 in the case of perfect agreement, 0 in the case of independence and -1 in the case of perfect disagreement.

In Fig. 5 we note that eigenvector centrality is the measure that provides the most similar ranking to CNE, together with the closeness centrality. While the former matching is somewhat expected (although with some differences since the two variables are not perfectly correlated) the latter seems quite surprising since the concept of shortest path, that is core to closeness centrality, is not the conceptual focus of our measure (except for the fact that the neighbours of a node are at distance one from it). In general,

Table 5 Classes of degrees and relative results for *Pr* considering eigenvector centrality. Note that being the network connected the class ZK is empty

Classes	Characteristic scores	No. of nodes in the cluster	No. of over-performing nodes (eigenvector)	*Pr*
LK	$k_i \leq 47$	24,068	11,078	0.94
AK	$48 \leq k_i \leq 183$	4119	2024	0.67
HK	$184 \leq k_i \leq 510$	373	383	0.52
VHK	$k_i > 510$	341	51	0.55

Table 6 Classes of degrees and relative results for *Pr* considering closeness centrality

Classes	Characteristic scores	# of nodes in the cluster	No. of over-performing nodes (closeness)	*Pr*
LK	$k_i \leq 47$	24,068	12,211	0.98
AK	$48 \leq k_i \leq 183$	4119	2109	0.73
HK	$184 \leq k_i \leq 510$	419	383	0.55
VHK	$k_i > 510$	341	49	0.34

Note that being the network connected the class ZK is empty

a high correlation between the obtained rankings is somewhat expected (Valente et al., 2008) and the usefulness of different centrality measures, including CNE, should be found in their capability to identify nuances in the concept of importance.

Given such results we repeated our analysis finding slightly different results in the case of eigenvector centrality (Table 5) and quite different results in the case of closeness centrality (Table 6).

The stronger similarities in terms of the indicator *Pr* occur in the case of the eigenvector even if the CNE seems to display a slightly higher probability in the case of nodes belonging to the class HK. The difference with the closeness centrality is, in general, more evident and mostly localised in the class VHK whereas nodes over-performing with respect to average closeness seems do not have a better access to funding with respect to under-performing ones. When we consider nodes with lower degree values, the association with funding of over-performing nodes is higher than what observe for CNE.

Comparing these results with those obtained in "Association between CNE and funding" section we note that the declining effect of our measure in the case of high degree is observed also for all the others centrality measures and perhaps it could be associated to certain node features that both the network structure and the FP7 dataset do not fully capture. From a quantitative perspective, we noted that the network is slightly disassortative ($r = -0.107$) with respect to degree—i.e., high degree nodes tend to be connected to low(er) degree ones. This for instance could be partially explained by the fact that the CNE follows a saturation-like curve, as shown in Fig. 4. In other words, the correlation profile of degree and CNE is not linear. Furthermore, when the network is disassortative the k_{nn} decreases with respect to k and for very high degree nodes the CNE is likely to be determined more by their own network share than by the quality of their neighbours. Such a theoretical observation could explain the fluctuations in the actual quality of neighbours of high degree nodes and, in turn, be among the factors causing a decreasing trend of

Pr. Such fluctuations are also expected in systems characterised by heterogenous elements such as the FP7 network.

Conclusions

Several studies document the advantages that being part of a network can bring to an organisation. Indeed, when an organisation joins a network, it expects to obtain better performance due to collaborations, knowledge sharing, technology acquisitions, increased capabilities and so on. Moreover, it is possible to increase opportunities, to help in implementing sustainable development and to assist in assimilating, utilising and reconfiguring external knowledge.

This paper contributes to this debate by providing a new measure that quantifies the network effect through the concept of network value and the patterns of relationships around specific nodes. In order to test the effectiveness of the proposed measure we considered as a case study the network built on the whole set of participants at the FP7 initiative of the European Commission. The measure highlighted two main groups of organisations; those having a network effect higher than expected due to the presence of a valuable neighbourhood and those having a network effect lower than expected due to a weak neighbourhood. To evaluate the capacity of our measure to quantify the quality of the neighbourhood of a member, as well as the implication of being part of one group or another, we considered the financial contributions received from the European Commission. Results show that, for small nodes, it is crucial to be part of the group of over-performing nodes because in this case an organisation is almost sure to receive more conspicuous funding than if they had belonged to a group of under-performing nodes. Disproportionate to the node degree, this effect increasingly weakens until randomness, meaning that other factors beyond the quality of the neighbourhood gain importance. Furthermore, in the FP7 programme, high degree nodes are mostly research organisations or universities that are usually well organised for initiatives such as funded projects, while many small or medium degrees nodes are SMEs. This is not surprising since SMEs constitute the majority of enterprises in the EU. As reported in Staniewski et al. (2016), despite their strong motivation to set up a business, SMEs face many different difficulties, due to a lack of professional experience and, indeed, funding. In our analysis, we obtain a post-hoc information that confirm that is beneficial for an organisation to join a partnership with high quality nodes. This aspect is partially highlighted by other centrality measures, since the degree is basically the number of partners times the number of projects, the closeness shows communication patterns and the eigenvector do not discriminates classes of nodes, unless when employing more sophisticated cluster methodologies.

With our measure we are mostly interested in pointing out the benefit of being part of a network with more accuracy than the Metcalfe's Law. When we consider our case study, we obtain (post-hoc) the information that the network effect act on small organisations in a way that the selection of a good partnership brings to better funding opportunities. In summary, our measure presents two main advantages. First, it provides an overall information on the network more accurate than the Metcalfe's Law. Indeed, it provides a macro measure that can be considered as a baseline in order to express the benefit of being part of the network. For instance, an organisation willing to join a network could potentially choose among alternatives on the basis of various centrality measures among which the network effect. Second the CNE displays and interesting micro feature. Namely, when a node has

joined a network, it can evaluate its network share in a way that is more accurate than the Metcalfe's Law, since it can compute the value with respect to the real number of links in the network, as well as it can compare its performance with nodes having the same class of degree.

Under these perspectives, the measure was not conceived in order to be the determinant of some endogenous characteristic as, for instance, the amount of funding, but to draw to attention on the fact that being part of a network is not beneficial on its own but the potential benefit also depends on the pattern of connections in which the node is embedded.

Such attempt to determine a measure for the network effect still has some limitations. The first is that networks can exist in a large assortment of forms, possibly mutating their nature over time; they can represent systems at different levels of scale (micro, meso and macro), and they can have different scopes. Furthermore, members behaviours are difficult to categorise, information and processes stemming from connections are particularly heterogeneous, different management styles are possible, and so on. The second issue is that, in many real cases, the network structure is not available or it is impossible to retrieve.

From a methodological perspective, further research should be devoted to the application of the proposed measure in other domains in order to analyse its potentialities as well as its effectiveness. In all such cases, contextualisation will be particularly important in order to reveal the correlations of our measure with the case under observation. More in general and coherently with the correlation analysis discussed in "Comparison with other centrality measures" section, we remark that the CNE should not be considered as substitute to other centrality measures but rather as a new element able to enrich the set of tools that allow scientist and practitioners to capture the many nuances elicited by the structure of complex networks.

From an empirical perspective, our case study could be enriched by performing a longitudinal analysis of the whole series of Framework Programmes, including H2020. A longitudinal analysis could provide interesting insights on the evolution of the network structure and on the process of growth in terms of new organisations getting funds for their project. A potential research effort pursuing such direction could be related to studying the assortativity of the FP7 network (i.e., the tendency of similar nodes to connect) that we find to be (against our expectations) slightly negative and equal to $r = -0.107$. The tendency of high degree nodes to connect to low degree ones more than expected at random could be reverted in previous/future snapshots indicating the propensity of nodes to form clubs of prominent elements to get funding for their projects.

Funding Open access funding provided by Università degli Studi di Roma La Sapienza within the CRUI-CARE Agreement.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

References

Abramo, G., D'Angelo, C. A., & Soldatenkova, A. (2017). An investigation on the skewness patterns and fractal nature of research productivity distributions at field and discipline level. *Journal of Informetrics, 11*(1), 324–335.

Alabi, K. (2017). Digital blockchain networks appear to be following Metcalfe's law. *Electronic Commerce Research and Applications, 24*, 23–29.

Arpetti, J., & Iovanella, A. (2020). Towards more effective consumer steering via network analysis. *European Journal of Law and Economics, 50*(3), 359–380.

Balland, P.-A., Boschma, R., & Ravet, J. (2019). Network dynamics in collaborative research in the EU, 2003–2017. *European Planning Studies, 27*(9), 1811–1837.

Barabási, A.-L. (2016). *Network science*. Cambridge university press,

Bergé, L., Scherngell, T., & Wanzenböck, I. (2017). Bridging centrality as an indicator to measure the 'bridging role' of actors in networks: An application to the european nanotechnology co-publication network. *Journal of Informetrics, 11*(4), 1031–1042.

Bol, T., de Vaan, M., & van de Rijt, A. (2018). The Matthew effect in science funding. *Proceedings of the National Academy of Sciences, 115*(19), 4887–4890.

Bollobás, B. (2013). *Modern graph theory* (Vol. 184). Springer.

Borgatti, S. P., & Everett, M. G. (2006). A graph-theoretic perspective on centrality. *Social networks, 28*(4), 466–484.

Briscoe, B., Odlyzko, A., & Tilly, B. (2006). Metcalfe's law is wrong. *IEEE Spectrum, 43*(7), 34–39.

Calignano, G. (2021). Better connected, more reputable? On the association between node centrality and academic reputation in the european union research and innovation networks. *European Policy Analysis, 7*(1), 240–262.

Calignano, G., Fitjar, R. D., & Hjertvikrem, N. (2019). Innovation networks and green restructuring: Which path development can EU Framework Programmes stimulate in norway? *Norsk Geografisk Tidsskrift-Norwegian Journal of Geography, 73*(1), 65–78.

Calignano, G., & Quarta, C. A. (2015). The persistence of regional disparities in Italy through the lens of the European Union nanotechnology network. *Regional Studies, Regional Science, 2*(1), 470–479.

Catanzaro, M., Boguñá, M., & Pastor-Satorras, R. (2005). Generation of uncorrelated random scale-free networks. *Physical Review E, 71*(2), 027103.

Chandler, G. N., & Hanks, S. H. (1993). Measuring the performance of emerging businesses: A validation study. *Journal of Business venturing, 8*(5), 391–408.

Choi, H., Kim, S.-H., & Lee, J. (2010). Role of network structure and network effects in diffusion of innovations. *Industrial Marketing Management, 39*(1), 170–177.

Cinelli, M. (2021). Ambiguity of network outcomes. *Journal of Business Research, 129*(1)*,* 555–561.

Cinelli, M., Ferraro, G., & Iovanella A. (2019). Evaluating relevance and redundancy to quantify how binary node metadata interplay with the network structure. *Scientific Reports*, *9*, 11404.

Cooke, P., & Wills, D. (1999). Small firms, social capital and the enhancement of business performance through innovation programmes. *Small Business Economics, 13*(3), 219–234.

Cova, B., Mazet, F., & Salle, R. (1994). From competitive tendering to strategic marketing: an inductive approach for theory-building. *Journal of Strategic Marketing, 2*(1), 29–48.

Cowan, R., & Jonard, N. (2004). Network structure and the diffusion of knowledge. *Journal of Economic Dynamics and Control, 28*(8), 1557–1575.

de Arroyabe, J. C. F., Schumann, M., Sena, V., & Lucas, P. (2021). Understanding the network structure of agri-food FP7 projects: An approach to the effectiveness of innovation systems. *Technological Forecasting and Social Change, 162*, 120372.

Defazio, D., Lockett, A., & Wright, M. (2009). Funding incentives, collaborative dynamics and scientific productivity: evidence from the EU framework program. *Research Policy, 38*(2), 293–305.

Ebers, M. (2015). International encyclopedia of the social & behavioral sciences. In J. D. Wright (Ed.), *Interorganizational relationships and networks* (2nd ed., pp. 621–625). Elsevier.

Egghe, L., Guns, R., & Rousseau, R. (2013). Measuring co-authors' contribution to an article's visibility. *Scientometrics, 95*(1), 55–67.

Etemad, H., Wright, R. W., & Dana, L. P. (2001). Symbiotic international business networks: collaboration between small and large firms. *Thunderbird International Business Review, 43*(4), 481–499.

Feld, S. L. (1991). Why your friends have more friends than you do. *American Journal of Sociology, 96*(6), 1464–1477.

Ford, D., Gadde, L.-E., Hakansson, H., & Snehota, I. (2011). *The business marketing course: Managing in complex networks*. Wiley.

Gilder, G. (1993). Metcalfe's law and legacy. *Forbes ASAP, 152*(6), 158–166.

Hazır, C. S., LeSage, J., & Autant-Bernard, C. (2018). The role of R &D collaboration networks on regional knowledge creation: Evidence from information and communication technologies. *Papers in Regional Science, 97*(3), 549–567.

Heller-Schuh, B., Barber, M., Henriques, L. M., Paier, M., Pontikakis, D., Scherngell, T., Veltri, G., & Weber, M. (2011). *Analysis of networks in European Framework Programmes (1984–2006)*. Technical report. Joint Research Centre (Seville site).

Henrikki, T. (1998). The network approach in analyzing international marketing and purchasing operations: a case study of a european SME's focal net 1992–95. *Journal of Business & Industrial Marketing, 13*(2), 109–131.

Liao, C. H. (2021). The Matthew effect and the halo effect in research funding. *Journal of Informetrics, 15*(1), 101108.

Liu, Y., Rousseau, R., & Guns, R. (2013). A layered framework to study collaboration as a form of knowledge sharing and diffusion. *Journal of Informetrics, 7*(3), 651–664.

Ma, A., Mondragón, R. J., & Latora, V. (2015). Anatomy of funded research in science. *Proceedings of the National Academy of Sciences, 112*(48), 14760–14765.

Madureira, A., den Hartog, F., Bouwman, H., & Baken, N. (2013). Empirical validation of Metcalfe's law: How internet usage patterns have changed over time. *Information Economics and Policy, 25*(4), 246–256.

Metcalfe, B. (2013). Metcalfe's law after 40 years of ethernet. *Computer, 46*(12), 26–31.

Newman, M. E. J. (2002). Assortative mixing in networks. *Physical Review Letters, 89*(20), 208701.

Newman, M. E. J. (2003). Mixing patterns in networks. *Phys. Rev. E, 67,* 026126.

Newman, M. E. J. (2018). *Networks*. Oxford University Press.

Odlyzco, A., & Tilly, B. (2005). A refutation of Metcalfe's law and a better estimate for the value of networks and network interconnections. Technical report, Digital Technology Center University of Minnesota, March 2005

Partanen, J., Möller, K., Westerlund, M., Rajala, R., & Rajala, A. (2008). Social capital in the growth of science-and-technology-based SMEs. *Industrial Marketing Management, 37*(5), 513–522.

Pastor-Satorras, R., Vázquez, A., & Vespignani, A. (2001). Dynamical and correlation properties of the internet. *Physical Review Letters, 87*(25), 258701.

Piekkola, H. (2007). Public funding of R &D and growth: firm-level evidence from finland. *Economics of Innovation and New Technology, 16*(3), 195–210.

Reed, D. P. (1999). That sneaky exponential—beyond Metcalfe's law to the power of community building. *Context Magazine* 2(1).

Rowley, T. J. (1997). Moving beyond dyadic ties: A network theory of stakeholder influences. *Academy of management Review, 22*(4), 887–910.

Schubert, A., Glänzel, W., & Braun, T. (1987). A new methodology for ranking scientific institutions. *Scientometrics, 12*(5–6), 267–292.

Scott, J., & Carrington P. J. (2011). *The SAGE handbook of social network analysis*. SAGE.

Staniewski, M. W., Szopiński, T., & Awruk, K. (2016). Setting up a business and funding sources. *Journal of Business Research, 69*(6), 2108–2112.

Swann, G. P. (2002). The functional form of network effects. *Information Economics and Policy, 14*(3), 417–429.

Valente, T. .W., Coronges, K., Lakon, C., & Costenbader, E. (2008). How correlated are network centrality measures? *Connections (Toronto, Ontorio), 28*(1), 16.

Van Hove, L. (2016). Testing Metcalfe's law: Pitfalls and possibilities. *Information Economics and Policy, 37*, 67–76.

White, D. R., & Harary, F. (2001). The cohesiveness of blocks in social networks: Node connectivity and conditional density. *Sociological Methodology, 31*(1), 305–359.

Zhang, X.-Z., Liu, J.-J., & Xu, Z.-W. (2015). Tencent and Facebook data validate Metcalfe's Law. *Journal of Computer Science and Technology, 30*(2), 246–251.