Journal of Informetrics 11 (2017) 1031–1042

Contents lists available at ScienceDirect

Journal of Informetrics

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

Regular article

Bridging centrality as an indicator to measure the 'bridging role' of actors in networks: An application to the European Nanotechnology co-publication network

Laurent Bergé [a,*], Thomas Scherngell [b,*], Iris Wanzenböck [c,*]

[a] Centre for Research in Economics and Management, University of Luxembourg, Luxembourg
[b] Center for Innovation Systems & Policy, AIT Austrian Institute of Technology, Austria
[c] Copernicus Institute of Sustainable Development, Utrecht University, Netherlands

ARTICLE INFO

Article history:
Received 9 January 2017
Received in revised form 30 June 2017
Accepted 12 September 2017
Available online 12 October 2017

Keywords:
Bridging centrality
Nanoscience and nanotechnology
Co-publications
R&D networks
Actor positioning
S&T indicators

ABSTRACT

In the recent past, we can observe growing interest in STI studies in the notion of *positioning indicators*, shifting emphasis to actors in the innovation process and their R&D inter-linkages with other actors. In relation to this, we suggest a new approach for assessing the positioning of actors relying on the notion of bridging centrality (BC). Based on the concept of bridging paths, i.e. a set of two links connecting three actors across three different aggregate nodes (e.g. organisations, or regions), we argue that triangulation in networks is a key issue for knowledge recombinations and the extension of an actor's knowledge base. As bridges are most often not empirically observable at the individual level of research teams, we propose an approximated BC measure that provides a flexible framework for dealing with the aggregation problem in positioning actors. Hereby, BC is viewed as a function of an aggregate node's (i) participation intensity in the network, (ii) its openness to other nodes (i.e. the relative outward orientation of network links), and iii) the diversification of links to other nodes. In addition, we propose a generalised version of the BC measure that accounts for different node categories. An illustrative example on the European Nanotechnology co-publication network observed at the level of organisations demonstrates the usefulness and complementary interpretation power in comparison to conventional centrality measures.

1. Introduction

Over the past decade, we have observed considerable progress in the advancement and application of Science, Technology and Innovation (STI) indicators (see, e.g., OECD, 2005). Three central drivers may be distinguished that have significantly triggered this development: *First*, there is wide agreement, both in the scientific and in the policy realm, that Research & Development (R&D) activities play a crucial role for economic growth in the long-run. New growth theory, for instance, emphasises that knowledge production and innovation of firms and other agents contribute to economic productivity gains because of the existence of industry wide knowledge externalities (see, e.g., Grossman and Helpman, 1991). *Second*,

* Corresponding authors.
E-mail addresses: laurent.berge@uni.lu (L. Bergé), thomas.scherngell@ait.ac.at (T. Scherngell), i.wanzenbock@uu.nl (I. Wanzenböck).

http://dx.doi.org/10.1016/j.joi.2017.09.004

research and innovation processes have become increasingly characterised by an interactive and networked nature (see, e.g., Scherngell, 2013), where innovating actors are impelled to tap and incorporate external knowledge sources in their R&D activities (see, e.g., Fischer 2001). Lepori et al. (2008) have provided a systematic discussion on respectively emerging needs for STI indicators in this context. The latter have for a long time – and actually still – put stronger emphasis on single entity specific characteristics of e.g. organisations, regions or countries, rather than their inter-linking and relative positioning in the local and global innovation system. *Third*, significant technical advancements in the production and handling of systematic data on STI have had a complementary effect in this context, enabling researchers to apply new indicators which has ignited a more intensive discussion on their strengths and weaknesses, and, by this, a more targeted and purposeful advancement of these indicators.

In relation to the perception of the interactive nature of knowledge production, the notion of *positioning indicators* has gained increased interest in science and innovation studies. It originates from considerations on new requirements imposed to the production of indicators in terms of their adaption from classical input-output to a positioning indicators framework, focusing on flows and linkages between research actors in the innovation system (Lepori 2008). These linkages may take the form of more formalised collaborations in R&D, such as joint R&D projects (see, e.g., Scherngell and Barber, 2011; Scherngell and Lata, 2013), joint publication activities (see, e.g., Glänzel and Schubert, 2004; Abbasi, Hossain, & Leydesdorff, 2012 Newman, 2001; Rodriguez and Pepe, 2008), and researchers mobility (see e.g. Edler, Fier, & Grimpe, 2011), or they appear as unformalised knowledge flows – often referred to as disembodied knowledge flows (see, e.g., Fischer, Scherngell, & Jansenberger, 2006). However, though a number of works have provided evidence on the crucial importance of such linkages between researching organisations for the successful generation of new knowledge, most indicators remain rooted in a classical, linear conception of the research and innovation process.

In that regard, we employ a *Network Science* perspective in this study, relying on the notion of *R&D networks* defined as a set of nodes representing knowledge producing actors interlinked via edges representing knowledge flows (see, e.g., Scherngell 2013). From a network theoretical perspective, the *positioning* of actors is most often described by the concept of *centrality*, rooted in Social Network Analysis (SNA). Centrality measures try to capture a certain function and/or role of a node, i.e. an actor, due to its inter-relations to other nodes (Borgatti 2005). Thus, the concept of *centrality* can be related to the concept of *positioning indicators*. Up to now, a number of science and innovation studies have started to utilize the centrality concept. However, these studies employ the most basic analytical concepts (Yan, Zhai, & Fan, 2013), such as degree, betweenness or eigenvector centrality. Usually they neglect conceptual problems that arise for networks defined at the aggregate level of (large) organisations, rather than the relevant level, i.e. individual researchers or a research team within an organisation. In particular, such problems are related to the loss of information regarding the structure of network relations and with that, information on the real channels through which knowledge flows. Furthermore, the used centrality measures are not conceptually adapted to the context of science and innovation studies, accounting for theoretical considerations on, e.g., the recombination of knowledge or relational capacity.

Thus, we propose an alternative measurement approach as a significant and complementary extension to common centrality measures, based on conceptual and theoretical debates in science and innovation studies. It is intended to provide a flexible framework enabling us to overcome problems related to the (often implicit) aggregation of nodes in R&D networks, *i*.e. the aggregation of links between research teams to often very large research organisations, or even regions and countries. The concept of 'bridging paths' provides us a good entry point in this context. A *bridging path* denotes an indirect connection between two nodes via a third 'bridging node', e.g. a research team *A* located in one organisation at one specific geographical place connects two other research teams, *B* and *C*, located in other organisations and places. Based on the individual bridging paths, we define a *Bridging Centrality* (BC) measure accounting for the 'bridging role' of an aggregate node in the network driven by the underlying inter-linking activities by the micro-actors of a that node.

Since link information at the level of research teams is often empirically not observable, in particular for large-scale R&D networks (e.g. co-patent-, co-publication or project-based R&D networks), empirical network studies are usually conducted at the level of organisations, or even regions and countries. Thus, we propose an approximation measure that, both conceptually and mathematically, relates to the number of bridging paths that we can expect for an aggregate node. The objective of this study is to discuss the conception and formal derivation of the basic BC measure first introduced by Bergé et al. (2017) for the specific case of regions, and to apply it to a highly interesting real-world context for science and innovation studies that is the European Nanotechnology co-publication network. Further we aim to advance the measure by generalising it to situations in which we are interested in the bridging centrality of organisations between different types of nodes, e.g. between organisations from science and industry, or national vs. international ones.

The remainder of the study is structured as follows: Section 2 introduces the network definition to elaborate in some more detail on the notion of bridges. Afterwards we outline the formal definition of our basic and generalised BC measure in Section 3, and derive the three components of Bridging Centrality: *participation intensity*, *relative outward orientation* and *diversification*. It shows how we conceptually perceive the number of bridges of a node starting from these three components, before we generalise the measure to account for different node categories. Then we shift attention to an illustrative example in Section 4, applying our measures to the European Nanotechnology co-publication network observed at the level of organisations, and comparing results with conventional centrality measures. The final section closes with some concluding remarks and some concrete ideas for a future research agenda.

2. Network definition and the concept of bridges

In social sciences, analytical strategies employed to deal with the divide between individualistic and holistic approaches for describing social systems are referred to as multilevel analysis (see, e.g., Lazega and Snijders 2015). In traditional sociological literature, this is aptly described as the phenomenon of *ecological fallacy*, pointing to logical failures in the inference of statistical data observed at an aggregated level on the nature and characteristics of individuals (see Robinson 1950). Social Network Analysis (SNA) faces, on the one hand, often similar problems, in particular in a context of science and innovation studies (see, e.g., Wanzenböck, Scherngell, & Brenner, 2014), while on the other hand entails promising potentials to overcome such analytical problems (Snijders 2016).

Concerning the focus of this study, we argue that these aggregation problems prominently occur in the measurement of the network positioning of actors in STI studies. SNA provides a rich toolbox to evaluate the positioning and, by this, the role of actors in different types of spatial or sectoral innovation systems. The concept of centrality is fundamental in this respect, usually adopted to assign a value to each actor in an innovation system – represented as network comprising nodes, usually organisations, regions or countries, inter-linked by edges (R&D linkages) – depending on their position within the network (Wasserman and Faust 1994). However, most SNA measures of centrality have been developed for the analysis of social systems, where the nodes of the network are usually identified in terms of individual persons. Accordingly, the original meaning borne by the SNA centrality measures as well as respective interpretations rely on assumptions on the social behaviours of individual persons, and how these persons might influence each other by these behaviours.

Thus, the interpretation of traditional SNA centrality measures based on observations at the organisational, regional or even country level of analysis raises important conceptual issues. Most importantly, it implies that every individual actor bounded under a specific *aggregate node* (e.g. organisation, region or country) would homogeneously benefit from the R&D linkages to other aggregate nodes, irrespective of who establishes the relations and the strength of these relations. Such an approach is based on the assumption that node-internal knowledge flows are in the 'air'. However, this assumption appears heroic and can hardly be made; it remains unclear how the individual actors belonging to a specific aggregate node benefit from the centrality observed at the aggregated level.

In this context, we propose a flexible analytical approach to overcome – or at least address – conceptual problems related to unobserved micro-level structures and dynamics of the observed network at an aggregate level. Hereby, we shift attention to the concept of 'bridging path' denoting a form of indirect connection between aggregate nodes, i.e. nodes such as organisations or regions that are indirectly connected in the network through their micro-level individual actors. This idea can be directly related to important concepts from network science, like the discussion on structural holes (Burt 1992) or on weak ties (Granovetter 1973;Borgatti and Halgin 2011). Considering the structural hole idea, it can be argued that a peculiar 'bridging position' of an aggregate node points to a brokerage function between groups or clusters of nodes. 'Bridging nodes' may have the important function to 'close' structural holes between aggregate nodes. Connecting arguments from the discussion on weak ties, it can be assumed that nodes with a higher Bridging Centrality are likely to act as brokers for specifically precious information. They exert a peculiar role for knowledge diffusion and transmitting information throughout the network as bridges are more likely to convey non-redundant information than other linkages.

For a formal definition of the bridging path concept, consider a network observed at the level of aggregate nodes, e.g. organisations, and the connections between the aggregate nodes represent the R&D linkages between their individual actors.[1] This represents a weighted network where we define g_{ij} as the number of R&D linkages (*i.e.* micro-level links) between aggregate nodes i and j. Further, each micro-level link between two aggregate nodes is denoted by y_{ik}^{a}, representing the a^{th} link between aggregate nodes i and j with $a \in \left\{1, \ldots, g_{ij}\right\}$. A bridging path is then regarded as a set of two links at the micro level connecting three individual actors from three different aggregate nodes. In this sense, a bridging path is defined as a set of two links from two different aggregate nodes, say i and j, with a third one, say k, so that the actors from i and j are both connected to the same actor in k. This means that a pair of links (y_{ik}^{a}, y_{jk}^{b}) forms a bridging path if, and only if, y_{ik}^{a} and y_{ik}^{b} are connected to the same actor of aggregate node k.

This concept is depicted by Fig. 1 representing a network of three aggregate nodes, which we assume are organisations in this illustration (see Guns and Rousseau 2015 for related ideas in the context of grouping network nodes[2]). In this figure, the pair of links (y_{ik}^{2}, y_{jk}^{1}) is a bridging path between aggregate nodes, *i.e.* organisations, i and j stemming from k because the actor from k maintains both links y_{ik}^{2} and y_{jk}^{1}. Although both aggregate nodes j and k do have links with aggregate node i, there is no bridging path between them because the actors from i of the links y_{jk}^{1} and y_{ik}^{2} neither connected to y_{ij}^{1}, y_{ij}^{2} nor y_{ij}^{3}. Hence, aggregate node i does not provide any bridging path between aggregate nodes j and k in this illustrative example.

[1] Note that we do explicitly not speak about the notion of multi-level networks here, in which network links are observed at different levels of aggregation. In this study we refer to a situation where we observe network links only at a higher aggregation level, and later derive a respective measure applied at this aggregate level (in our case organisations participating in R&D networks). We do not observe the links at different levels (e.g researchers and organisations), i.e. we cannot trace multi-level network structures.

[2] Note that the definition of bridging paths used in this study differs, both analytically and conceptually, from the bridging path concept introduced by Hwang et al. (2006) or Guns and Rousseau (2015), with the latter using the notion of Gefura measures. Both refer to derivations of the betweenness concept, trying to capture a nodeís bridging function between different groups of nodes.

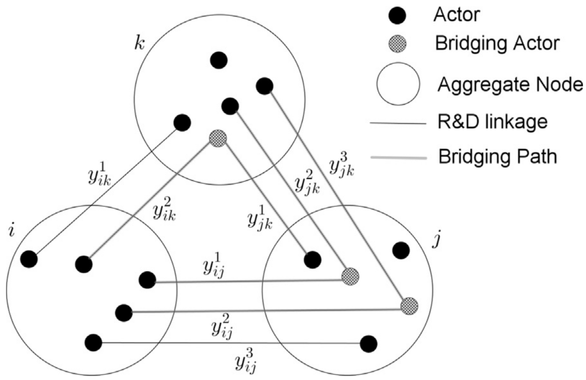

Fig. 1. Illustration of the notion of bridging path.
Notes: The figure is based on Bergé et al. (2017) and depicts three bridging paths formed by the following pairs of links: (y_{ik}^2, y_{jk}^1), $(y_{ij}^1, y_{jk}^2)$and (y_{ij}^2, y_{jk}^3).So the aggregate node dyads (j, k), (i, k) and (i, j) have respectively 0, 2 and 1 bridging paths stemming from aggregate nodes i, j and k, respectively.

It can be seen that the concept of bridging path is about indirect connections. Accordingly, the aggregate node with most bridging paths is node j, as it provides two bridging paths between aggregate nodes i and k.

We argue that the concept of bridges is of particular relevance when studying R&D and innovation processes. A high number of bridging paths implies a more open positioning in the network. In contrast to closed and dense network structures, such a bridging position between other nodes can be related to the access to a more diversified knowledge pool. It is assumed that the sources from which the individual actors draw their knowledge will have an impact on their ability to generate innovations, and knowledge flowing through bridging paths is more likely heterogeneous and non-redundant. Based on the bridging path concept, we propose a measurement approach for Bridging Centrality (BC) in the section that follows.

3. A measure for approximating Bridging Centrality

Given the parsimonious and effective formal definition of bridging paths, it could be assumed at a first glance that the definition of a formal BC measure is straightforward. Indeed, this is the case in pure mathematical terms as the *true* measure of BC for an aggregate node i would just be the number of bridging paths stemming from the actors of node i, probably normalised by all bridging paths in the network. However, de facto we are most often confronted with a well-known problem in social sciences, namely not to find empirical observations for the objects under scrutiny. This a particular critical issue in the context of science and innovation studies, where we are usually focusing on networks driven by individual researchers as main actors, sometimes large-scale networks with often ten thousands and more nodes and even more edges. Most prominent examples are project-based R&D networks, e.g. constituted under the heading of the EU Framework Programs (see, e.g., Scherngell and Barber 2011; Scherngell and Lata 2013), co-publication networks (see, e.g., Glänzel and Schubert, 2004) or co-patent networks (see, e.g., Wanzenböck, Scherngell, & Lata, 2015; Wanzenböck, Scherngell & Brenner, 2014). These networks have in common that – in contrast to multi-level networks – information on networking links at the level of the individual researchers cannot be traced; even when information is available (as e.g. for authors in publications and/or inventors in patents) the observation of large-scale empirical networks is infeasible due to immense efforts for data cleaning, in particular name standardisation over space and time.

Thus, we propose an alternative measure for BC that – both conceptually and mathematically – approximates the number of bridges of an aggregate node. From a conceptual perspective, and based on theoretical considerations from innovation research, we argue that the number of bridging actors of an aggregate node is a function of three different components. These components reflect an aggregate node's *i*) *participation intensity, ii*) *relative outward orientation* and *iii*) *diversification of network links*. Common data constraints usually impede the observation of the actual number of bridges for an aggregate node. However, we argue that these three components are both feasible and attractive to approximate its number of bridges. *First*, a higher *participation intensity* simply enhances the likelihood of getting into a bridging position by increasing the network intensity and the pool of partners accordingly. It is a general measure of how well an aggregate node is embedded in the R&D network, and can therefore be interpreted as a broad measure of its relational capacity; an aggregate node's size will amplify the probability of yielding more bridges between other nodes. *Second*, a higher *relative outward orientation* points to a positioning of an aggregate node to tap important external knowledge sources (see, e.g. Wanzenböck et al., 2015). A high number of node-internal collaborations would have a negative influence as it potentially reduces the number of actors connecting different aggregate nodes. *Third*, as concerns *diversification of network links*, we argue that a bridging position between aggregate nodes can be positively related to the access to a more diversified pool of actors, and thus, to a more diversified knowledge pool. Concentration of links reduces the possibility to build bridges among different aggregate nodes.

One central property of the measure is that it takes account of the peculiar shortcomings that may occur when networks are observed at the level of aggregated nodes. Indeed, such networks are characterised by the structure of node-internal and node-external links. In a R&D context, an aggregate node's ability to benefit from new ties in the network or exploit external knowledge sources via the links may be determined by all three components together. Outward orientation and higher diversification in particular may help actors belonging to an aggregate node to develop and renew their knowledge base faster, or prevent lock-in situations in certain technologies (see, e.g., Breschi and Lenzi, 2015).

Turning to the formal description and combining the three components introduced above in a linear-multiplicative way, we denote C_i as the basic approximated BC for aggregate node i by

$$C_i = q_i s_i (1 - h_i) \tag{1}$$

with the three components defined as follows:

q_i is the weighted degree of aggregate node i, defined as the total number of links excluding node-internal ones (loops), i.e. $q_i = g_i - g_{ii}$, with g_i denoting the total number of links of node i, referring – as discussed above – to the overall *participation intensity* in the network of aggregate node i.

s_i is the share of outer collaborations of aggregate node i with $s_i = q_i/g_i$, indicating the above mentioned *relative outward orientation* of all established network linkages, i.e. the relative degree of external linkages of aggregate node i.

h_i refers to the degree of *diversification of network links* of aggregate node i among other nodes. It is measured by the Herfindahl-Hischman (HH) index on the distribution of i 's outer collaborations defined as $\mathrm{h_i} = \sum_{\mathrm{j \neq i}} (\mathrm{g_{ij}}/\mathrm{q_i})^2$. For the calculation of the C_i we use the term 1- h_i so that it varies between 0 and 1, indicating how an aggregate node's linkages are distributed along its neighbouring nodes in the network.

An interesting feature of the measure is that it mathematically corresponds to the *expected number of bridges* (ENB) measure first introduced by Bergé (2017) when re-specifying Eq. (1) in form of

$$\begin{aligned} C_i &= \frac{q_i^2}{g_i}\left(1 - \sum_{j \neq i} (g_{ij}/q_i)^2\right) = \frac{q_i^2}{g_i} - \frac{1}{g_i}\sum_{j \neq i} {g_{ij}}^2 \\ &= \frac{1}{g_i}\sum_{j \neq i} g_{ij}(q_i - g_{ij}) = \frac{1}{g_i}\sum_{j \neq i}(g_{ij}\sum_{k \neq i,j} g_{ik}) \\ &= \sum_{j \neq i}\sum_{k \neq i,j}\frac{g_{ij}g_{ik}}{g_i} = \sum_{j \neq i}\sum_{k \neq i,j} ENB_{jk}^i \end{aligned} \tag{2}$$

where ENB_{jk}^i corresponds to the expected number of bridges between j and k stemming from i under a simple random matching assumption.[3] The random matching process approximates the underlying micro-structure by deriving an expected number of bridging paths between two aggregate nodes. To put it simply, the ENB for an aggregate node (e.g. an organisation) increases with the observed number of links to other aggregate nodes.

Note that in Eq. (2) all links are treated equally, i.e. the basic BC measure does not differentiate between specific types of bridging paths. In some situations we might be interested to learn more about the characteristics of a 'bridging role', based on the assumption that some linkages, for instance, between science and industry, national and international, or linkages bridging scientific or technological fields, are of a different type, or value, than linkages to the same category. The identification of organisations occupying such a specific 'bridging' function may be of interest for both the scientific domain and policy makers. The same applies to the identification of cross-border 'bridging' organisations, i.e. key organisations in providing bridges between international knowledge bases. The notion of bridging path is flexible enough to address such questions.

To this end, we define the BC measure generalised to account for specific node types by departing from the ENB as derived in Eq. (2). This measure leverages the ENB definition and counts the number of bridges that an organisation provides between organisations of different types. It is defined as follows:

$$\mathrm{C}_i^{(\mathrm{CAT})} = \sum_{j/T_j \neq T_i}\sum_{k/T_k \neq T_i}\frac{g_{ij}\,g_{ik}}{g_i} = \frac{g_i^{(\mathrm{S})}\,g_i^{(\mathrm{D})}}{g_i} \tag{3}$$

where T_i is a specific category/type of node i, and $g_i^{(\mathrm{S})}$ is the number of connections that node i has with organisations of similar type (excluding loops), while $g_i^{(\mathrm{D})}$is the number of connections that i has with organisations of different type. To follow on with the example, if i is a public research organisation, then $g_i^{(\mathrm{S})}$is the total number of links between i and all other public research organisations, while $g_i^{(\mathrm{D})}$is the number of links between i and all non-public research organisation, e.g. firms.

[3] Random matching statistically assumes an equal probability for each micro-actor of an aggregate node to be connected with a micro-actor of another aggregate node. Though – as we know very well from network science – this random matching assumption is not valid at the individual node level, it provides a rather suitable baseline for a micro-network (e.g. researcher-by-researcher) that is likely to occur having information only available at the aggregate network (e.g. organisation-by-organisation). This may particularly hold for larger aggregate nodes (e.g. large organisations) given the law of large numbers (see Bergé 2017 for further details). The random matching assumption can be extended to other linking assumptions (e.g. preferential attachment).

$C_i^{(CAT)}$ as defined by Equation (3), can be applied to any context involving different types of actors. In the empirical analysis that follows, we illustrate the potential of the two presented BC measures (C_i and $C_i^{(CAT)}$) in the context of Nanotechnology co-publication networks, and restrict ourselves to two different examples of categories, namely public-private and national-international linkages.

4. An application to the European Nanotechnology co-publication network

In this section, we shift attention to an illustration of the BC measures in order to demonstrate its behaviour and interpretative power. In this example, we use data on inter-organisational *Nanotechnology* co-publications, defined as publications featuring at least two authors affiliated with two different organisations. The focus on Nanotechnology fits very well to previous research works that have increasingly appeared over the past decade, investigating its dynamics and related STI policy implications (see e.g. Larédo, Delemarle, & Kahane, 2010; Bonaccorsi, 2008). For observing the Nanotechnology co-publication networks, we draw on data recorded in the “Nano S&T Dynamics” database.[4]

In this illustrative example, our focus will be on the organisational dimension. In order to construct our Nanotechnology inter-organisational co-publication network, we extract publications from the year 2010, with authors affiliated to organisations located either in one of the EU-28 members states, the EU associated states (excluding Israel) or Russia. One major advantage of the database used are the harmonized organisation names, based on intensive standardisation procedures and cleaning of organisation names (see Kahane, Mogoutov, Cointet, Villard, & Laredo, 2015 for further details). Accordingly, an extracted list of all publication identifiers (record numbers of publications in the database), with the respective standardized organisation names, constitutes an unweighted bipartite graph, connecting publications to organisation names (based on author affiliations). The latter can be used to project an organisation-by-organisation co-publication network based on authors' affiliations,[5] with loops indicating the number of co-publications within an organisation.

For the empirical illustration in this study, we extract the respective bipartite graph for the year 2010, and construct the organisation-by-organisation graph accordingly; note that we employ full counting when doing the projection, i.e. we count all co-author cross-linkages appearing for each publication as one.[6] Further note that in this example the aggregation problem described in Section 2 clearly applies as we are not able to observe co-publication activities at the level of individual researchers or research groups (since we do not have standardized author names), but have to aggregate to the organisational level (including rather large organisations with more than thousands researchers affiliated to), i.e. we cannot trace multi-level network structures. Aggregating our data to our inter-organisational co-publication network, based on a total of 20,377 Nanoscience and technology publications from 2010, produces a network of 87,865 inter-organisational collaborations linking 3,608 organisations.

We first calculate the scores for the basic BC measure as defined by Eq. (1) for the 3,608 organisations participating in the network. In order to grasp the functioning of the BC measure, and to understand its statistical and distributional properties, we compare the BC scores with the scores of three other commonly used centrality measures that is degree, eigenvector and betweenness centrality. We follow standard calculation procedures for the three other centrality measures (see Wasserman and Faust 1994 for details).[7] Before we reflect on results at the level of organisations, it is worth noting that the statistical properties of BC differ markedly from the conventional centrality measures (the rank correlation is 0.76 with degree, while only 0.63 and 0.55 with eigenvector and betweeness centrality, respectively).

Turning to the level of organisations, Table 1 represents the top 30 centralities ordered by their score achieved for the basic version of BC. The BC ranking is clearly dominated by non-university, public research organisations (PROs), with the top 10 almost exclusively containing PROs, most of them located in Germany and France.[8] Following the bundle of PROs

[4] The Nano database is maintained by Université Paris-Est Marne La-Valleé (UPEM) and is available at the RISIS platform (risis.eu). It comprises systematic information on publications focusing on Nano science and technology from 1991 to 2010, recording more than 1.1 million manuscripts published during this time period, including publications indexed in the Web of Science (WoS). The dataset is organised around three major dimensions, that is i) the organisational dimension with the affiliations of authors, ii) the geographical dimension of authors based on addresses (regions, countries, cities and clusters), and iii) a thematic dimension based on the subject categories of the WoS (see Kahane et al., 2015). The database also comprises Nano science and technology patents, recording more than 2.6 million Nano patents applied for during the time period 1991–2010. Though co-patent or patent citation networks are not subject to this study, they may be an interesting case for comparison purposes in future research.

[5] Formally, it would be more precise to use the notion of organisational co-authorship network. However, we follow the widely used convention in the literature that mostly focuses on organisations when using the notion of co-publications based on co-authorships.

[6] For example, a publication with, e.g., three authors A, B and C, and authors A and B affiliated to organisation Y, while author C to organisation Z, would produce two links between organisations Y and Z (from A to C, and B to C), and one loop (internal link from A to B) in organisation Y.

[7] Degree is simply interpreted as the degree of prestige an organisation has due its simple number of connections to other organisations. It is here calculated as the number of unique R&D interactions the agents of an organisation are involved in. Betweenness centrality of a vertex is defined as the fraction of geodesic paths between any pair of vertices on which this vertex lies. Those actors, who are located on the shortest paths between many actors, therefore hold a key position for controlling the flow of information within the network (gatekeeper function). Eigenvector centrality accords each vertex a centrality that depends both on the number and the quality of its connections by examining all vertices in parallel and assigning centrality weights that correspond to the average centrality of all neighbours. In this sense, a high Eigenvector centrality of an organisation indicates that this organisation is connected with other organisations that also show many connections, rather than to peripheral organisations (Hu, Scherngell, Man, & Wang, 2013). The three comparative centrality measures are computed using the package igraph available in the statistical software R.

[8] Next to the well-known, large scale PROs from France (CNRS), Germany (Max Planck, Helmholtz and Leibniz), Italy (CNR) and Spain (CSIC), a comparably high number of PROs from Eastern European countries, mainly the traditional academies are under the top 30, with the Russian Academy of Science showing

Table 1
Centralities of Top 30 organisations with respect to BC (ranks in brackets).

Organisation	BC (C_i)		Degree		Eigenvector		Betweenness	
CNRS	1.000	(1)	0.764	(2)	0.024	(4)	1.000	(1)
CNR	0.566	(2)	0.512	(3)	0.005	(21)	0.497	(4)
Max Planck	0.529	(3)	0.502	(4)	0.009	(10)	0.503	(3)
Helmholtz	0.408	(4)	0.396	(6)	0.010	(9)	0.478	(5)
CSIC	0.384	(5)	0.482	(5)	0.001	(120)	0.458	(6)
Russian AS	0.353	(6)	1.000	(1)	1.000	(1)	0.980	(2)
UPMC	0.336	(7)	0.242	(13)	0.006	(16)	0.177	(12)
CEA	0.305	(8)	0.349	(8)	0.006	(18)	0.160	(15)
Leibniz Gesell	0.264	(9)	0.243	(12)	0.008	(12)	0.233	(8)
ETHZ	0.229	(10)	0.353	(7)	0.003	(47)	0.179	(11)
Uni Paris sud	0.211	(11)	0.196	(21)	0.004	(35)	0.112	(23)
INSERM	0.205	(12)	0.093	(84)	0.001	(347)	0.075	(41)
Uni Cambridge	0.203	(13)	0.277	(10)	0.000	(349)	0.214	(10)
UJF	0.194	(14)	0.142	(38)	0.003	(40)	0.051	(77)
Czech AS	0.190	(15)	0.266	(11)	0.005	(22)	0.229	(9)
Polish AS	0.188	(16)	0.184	(23)	0.007	(15)	0.243	(7)
Uni Vienna	0.184	(17)	0.108	(59)	0.000	(351)	0.100	(28)
INFN	0.180	(18)	0.093	(86)	0.000	(354)	0.034	(109)
UCL	0.177	(19)	0.212	(19)	0.000	(350)	0.170	(13)
ESRF	0.167	(20)	0.097	(78)	0.005	(24)	0.057	(72)
Uni Strasbourg	0.160	(21)	0.147	(35)	0.001	(180)	0.066	(54)
Uni Toulouse 3	0.157	(22)	0.182	(24)	0.005	(25)	0.076	(40)
KIT	0.153	(23)	0.311	(9)	0.016	(5)	0.124	(21)
Uni Paris Diderot	0.146	(24)	0.101	(69)	0.001	(186)	0.033	(112)
Uni Munchen	0.145	(25)	0.220	(18)	0.001	(145)	0.147	(17)
Uni Oxford	0.143	(26)	0.196	(22)	0.004	(34)	0.135	(18)
Uni Tech Vienna	0.139	(27)	0.113	(57)	0.000	(363)	0.110	(24)
Uni Padua	0.128	(28)	0.145	(37)	0.000	(357)	0.069	(49)
Uni Bologna	0.124	(29)	0.141	(39)	0.002	(67)	0.105	(27)
Uni Helsinki	0.124	(30)	0.145	(36)	0.000	(397)	0.079	(37)

Notes: AS denotes Academy of Science; CNRS Centre National de la Recherche Scientifique; CNR Consiglio Nazionale delle Ricerche; CSIC Consejo Superior de Investigaciones Cientificas/Higher Council for Scientific Research; ETHZ Eidgenössische Technische Hochschule; ESRF European Synchrotron Radiation Facility;INSERM Institut National de la Sante Et de la Recherche Medicale; INFN Istituto Nazionale di Fisica Nucleare; KIT Karlsruhe Institute of Technology. UCL University College London; UJF University of Grenoble; UPMC Universität Pierre und Marie Curie.

under the Top 10, and a mix of PROs and universities between ranks 11–20, ranks 21–30 is dominated by universities.[9] The differences in the BC ranking and the ranking produced for the other centrality measures are remarkable, even under the Top-30 organisations. Most interestingly, the Russian Academy of Sciences shows the highest degree centrality, even higher than CNRS, while it is just ranked 6th for the BC score. In other words, the Russian Academy of Sciences is able to take a central position in inter-linking with a high number of other organisations, but its role to 'bridge' knowledge between other, not directly connected, organisations is less pronounced. In contrast, organisations like INSERM or the University of Vienna show a very high BC score, but a rather low ranking in degree centrality. INSERM is at rank 12th for BC, but just at rank 84th for degree centrality, while the University of Vienna comes at 17th for BC and 59th for degree. Obviously these organisations – though belonging to the top organisations concerning the pure number of links – feature very peculiar characteristics related to the ability in 'bridging' different knowledge regimes between a large number of not directly connected organisations (Table 2).

With the definition of the basic BC measure in terms of the three components, our centrality measure is able to give us – next to an assessment of the overall bridging role of organisations in networks – some indication on the reasons behind the 'bridging ability' of an organisation. In this context, Fig. 2 visualizes the backbone of the Nanotechnology co-publication network, extracting the Top30 organisations according to BC and plotting their network linkages (organisations that are strongly interconnected are positioned nearer to each other). We plot the networks three times; one time node size is proportional to the participation intensity q_i of an organisation i (box A), one time with its relative outward orientation s_i (box B), and finally with its diversifications intensity h_i (box C). This gives an insightful impression on the prevalent characteristics of an organisation to reach a certain BC score. Table A1 and Fig. A1 in the Appendix A complement this analysis, providing a snapshot on the top organisations and their scores as well as rankings across the three components. For instance, the Russian Academy of Science – that has the highest degree centrality of all organisations – does not come

up at rank 6. The high ranking of the Russian Academy is not surprising, given its high Nanotechnology publication intensity and immense growth in recent years, meanwhile belonging to the most important producers of knowledge related to Nanotechnology (see Karaulova, Shackleton, Gok, Kotsemir, & Shapira, 2014).

[9] Note that firms do not appear under the Top-100 for the BC measure which is not that surprising since we are looking at scientific publications. The highest ranked firm is Philips N. V., showing up at rank 176th.

Table 2
Top 30 organisations in Bridging Centrality between public and private organisations (BC Public-Private).

Organisation	BC Public-Private		BC Nat.-Int.		Degree	
CNRS	1	1.00	1	1.00	2	696
Russian AS	2	0.99	3	0.60	1	1040
CNR	3	0.94	5	0.49	5	453
CEA	4	0.72	11	0.27	10	291
Leibniz	5	0.56	8	0.30	8	319
Uni Tech Vienna	6	0.53	20	0.16	57	104
Helmholtz	7	0.51	6	0.43	6	392
ETHZ	8	0.46	12	0.26	7	346
Fraunhofer	9	0.46	64	0.09	24	161
Uni Tech Denmark	10	0.43	67	0.09	21	173
Uni Lille	11	0.42	43	0.11	43	114
Uni Tech Eindhoven	12	0.42	55	0.10	38	123
KIT	13	0.39	14	0.20	16	222
Uni Ghent	14	0.39	48	0.11	45	113
Max Planck	15	0.39	2	0.67	3	622
Uni Cambridge	16	0.36	21	0.15	9	298
Uni Paris Sud	17	0.36	16	0.19	20	188
Imperial College	18	0.36	26	0.14	18	198
Uni Lyon Claude Bernard	19	0.33	38	0.12	19	190
CSIC	20	0.33	4	0.54	4	512
EPFL	21	0.32	35	0.12	13	249
KU Leuven	22	0.31	17	0.18	22	167
Uni Tech Delft	23	0.30	44	0.11	31	142
Uni Belgrade	24	0.29	39	0.11	61	100
Uni Tech Aachen	25	0.29	83	0.08	40	120
Uni Twente	26	0.28	174	0.05	42	116
Uni Tech Chalmers	27	0.28	115	0.07	105	76
Uni Oxford	28	0.27	29	0.13	17	207
Uni Liverpool	29	0.27	147	0.06	138	59
Uni Freiburg	30	0.27	41	0.11	108	73

Notes: AS denotes Academy of Science; CNRS Centre National de la Recherche Scientifique; CNR Consiglio Nazionale delle Ricerche; CEA Centre d'études nucléaires de Grenoble, ETHZ Eidgenössische Technische Hochschule, KIT Karlsruhe Institute of Technology, CSIC Consejo Superior de Investigaciones Cientificas, EPFL École polytechnique fédérale de Lausanne.

on the top for BC because its participation intensity is lower than that of the other top organisations; moreover, its outward orientation is very small, just ranking 3,079th out of 3,608 organisations (see also Table A1).

While the results for the basic BC provide us some interesting insights into the general 'bridging' capability of an organisation, we have defined and motivated a generalised version of the measure in Section 3 that accounts for different types/categories of nodes. For our focus on inter-organisational Nanotechnology networks, at least two interesting applications come to mind: *First*, we intend to get some indication of an organisation's 'briding role' between public research organisations and firms in the network, and *second*, between national and international actors. In doing so, we calculate the generalised BC measure as defined by Eq. (3), discriminating between the respective node categories.

Table 2 reports the results with a focus on public-private categorization of nodes, featuring the Top 30 organisations in Bridging Centrality between public and private organisations (BC Public-Private), according with their Bridging Centrality between national and international organisations and their degree for comparison purposes. The results are quite interesting, most importantly showing that

i) the top organisations in the ranking for the basic BC (Table 1) do differ remarkably from the top organisations 'bridging' public research organisations and firms (Table 2), with a much stronger role of technical universities acting as important 'bridges' between science and industry.

ii) organisations taking a significant public-private 'briding role' do not necessarily take the same role between national and international organisations (e.g. Technical University of Vienna), and vice-versa (Max Planck, University of Oxford).

Such results may be a promising starting point for digging into further research questions going beyond this study. One may for instance study the role of universities versus non-university research organisation in their efficiency in providing bridges between science and industry (normalised by their publication intensity). For the purpose of this study, the results show the remarkable potential of the generalised BC version to shed a more complete light on the positioning of actors in R&D networks. They allow to capture specific concepts that would otherwise be difficult to grasp empirically and which we believe are of crucial relevance to studies focusing on networks and innovation.

5. Summary and conclusions

In science and innovation studies, we can observe a lively debate on the notion of positioning indicators. Such indicators are intended to account for recent changes in research and innovation processes, shifting emphasis to actors in the innovation

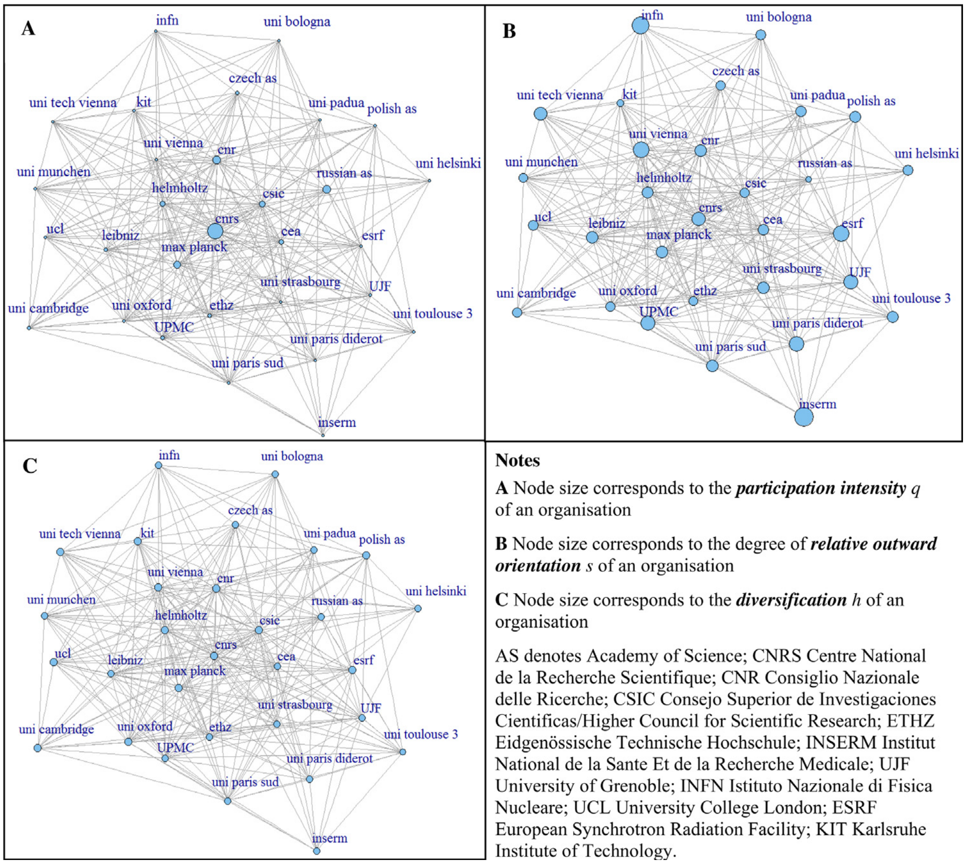

Fig. 2. Bridging centrality of Top30 organisations in the European Nanotechnology co-publication network decomposed by three components.

process and their linkages and R&D interactions with other actors. It is assumed that the innovation performance of an actor is highly dependent on its positioning in the wider web of R&D interactions. However, up to now only few indicators exist which are able to provide a rather comprehensive estimation of an actor's positioning in the innovation system, reflecting structural characteristics of its internal and external R&D linkages.

In this study, we have therefore shifted attention to the notion of Bridging Centrality (BC), and suggested a new approach for assessing the positioning of actors in research and innovation systems. We use the centrality concept derived from Social Network Analysis (SNA) to capture positioning, and use the notion of R&D networks for the conceptualisation of research and innovation systems. Based on the concept of bridging paths, i.e. a set of two links connecting three actors across three different aggregate nodes, e.g. organisations, regions or countries, we argue that triangulation in networks is a key issue for knowledge recombinations and the extension of an actor's knowledge base. As bridges are most often not empirically observable at the individual level of research teams, we propose an approximated BC measure that provides a flexible framework for dealing with the aggregation problem in positioning actors. Hereby, BC is viewed as a function of a node's *i*) participation intensity in the network, *ii*) its openness to other nodes (i.e. the relative outward orientation of network links), and *iii*) the diversification of links to other nodes. With these three components – which are both intuitive and computationally simple – we provide an integrative perspective that enables us to achieve a better understanding of the role of certain actors in R&D networks. Further, we have presented a generalised version of the BC measure that is able to account for different node categories. Such a measure broadens the range of possibilities for R&D studies by enabling to spot the critical agents/organisations that play a bridging role between different types of actors (e.g. science vs industry, national vs international, scientific or technology fields, etc.).

An illustrative example investigating the 2010 co-publication network of organisations in Nanoscience and -technology across Europe demonstrates the performance and usefulness of the BC measures. Not only do we observe differing patterns in basic statistics like correlations of the centralities or the distributional characteristics of BC as compared to other SNA measures, we were also able to show striking and interesting differences in the structure of the inter-organisational Nanotechnology co-publication linkages. Most importantly, the BC measure as such picks up a ranking reflecting on the 'bridging role' of organisations in networks that cannot be identified when looking at the conventional

SNA measures. The results reveal that thinking only of the degree of participation is not enough in this respect, showing that organisations having just a high degree of participation do not necessarily take a pronounced 'bridging position' for knowledge diffusion between different groups of organisations in the network. When we look at the generalised BC version, specifically accounting for different node types, we can see that organisations with a high basic Bridging Centrality are not necessarily also important bridges between science and industry, or between the national and the international level. The presented measures are able to provide deeper insights into empirically relevant phenomena for science and innovation studies, e.g by shedding light on the stronger role of technical universities for science-industry interactions.

Of course, there is the necessity for further investigations of the measure, and room for improvements of the measurement concept. Indeed, a *first* crucial point for future research is to dig further into the analysis of the statistical properties of the proposed BC measure. An approach using simulated networks would be a promising route to get a deeper impression on how the measure reacts based on differing underlying network structures, on the one hand, and based on different properties of the three components, on the other hand. *Second*, the measure is not limited to the context of R&D but may prove to be useful also for the application in other types of network structures, such as trade flows or economic value chains.

Author contributions

Thomas Scherngell: Conceived and designed the analysis, Collected the data, Contributed data or analysis tools, Performed the analysis, Wrote the paper.

Laurent Berge: Conceived and designed the analysis, Contributed data or analysis tools, Wrote the paper.

Iris Wanzenböck: Conceived and designed the analysis, Contributed data or analysis tools, Wrote the paper.

Acknowledgments

The work in this study has benefited significantly from the provision of original data via the RISIS (Research Infrastructure for Research and Innovation Policy Studies) infrastructure (risis.eu), drawing on data from the Nano S&T Dynamics database part of RISIS, owned and maintained by Université Paris-Est Marne-la-Vallée (UPEM). Specifically we are grateful to Lionel Villard and Philippe Laredo from UPEM in terms of data provision, and to two anonymous reviewers for useful comments to improve the manuscript. Partly this work has been funded by the Austrian National Bank, Jubiläumsfondsprojekt No. 16301.

Appendix A.

.

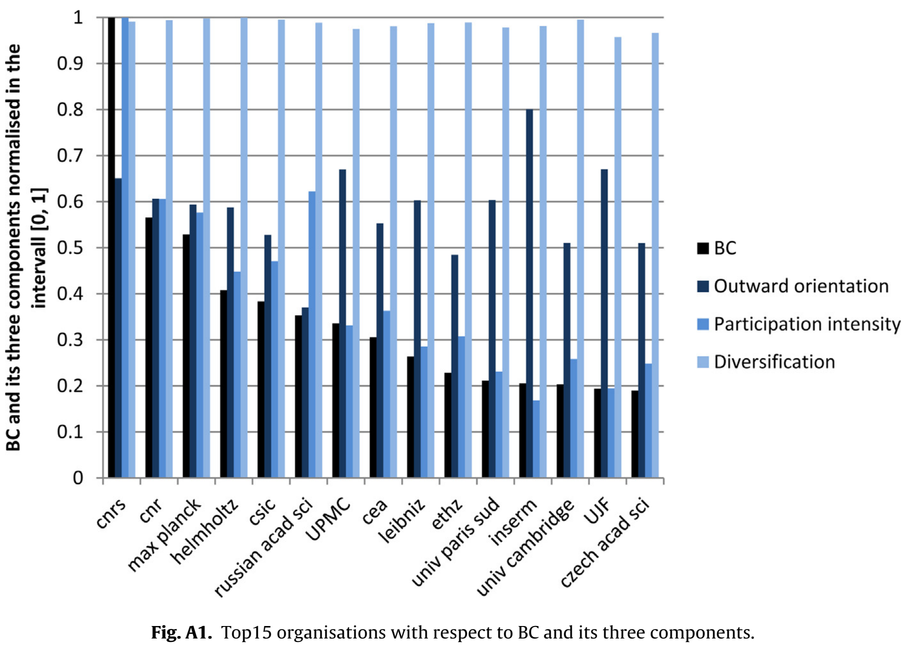

Fig. A1. Top15 organisations with respect to BC and its three components.

Table A1
Ranking of top10 organisations by three components of BC (rank in brackets).

Organisation	BC	Participation Intensity	Outward Orientation	Diversification
CNRS	1.000 (1)	1699 (1)	0.630 (1732)	0.984 (18)
CNR	0.566 (2)	1029 (3)	0.587 (1926)	0.987 (7)
Max Planck	0.529 (3)	979 (4)	0.575 (1967)	0.990 (2)
Helmholtz	0.408 (4)	761 (6)	0.569 (2007)	0.992 (1)
CSIC	0.384 (5)	800 (5)	0.511 (2190)	0.988 (4)
Russian AS	0.353 (6)	1057 (2)	0.359 (3079)	0.981 (30)
UPMC	0.336 (7)	563 (8)	0.649 (1684)	0.968 (120)
CEA	0.305 (8)	617 (7)	0.535 (2120)	0.973 (78)
Leibniz Gesell	0.264 (9)	485 (10)	0.584 (1933)	0.980 (34)
ETHZ	0.229 (10)	523 (9)	0.469 (2884)	0.981 (27)

Note: AS denotes Academy of Science; CNRS Centre National de la Recherche Scientifique; CNR Consiglio Nazionale delle Ricerche; CSIC Consejo Superior de Investigaciones Cientificas/Higher Council for Scientific Research; ETHZ Eidgenössische Technische Hochschule.
Note: AS denotes Academy of Science; CNRS Centre National de la Recherche Scientifique; CNR Consiglio Nazionale delle Ricerche; CSIC Consejo Superior de Investigaciones Cientificas/Higher Council for Scientific Research; ETHZ Eidgenössische Technische Hochschule; INSERM Institut National de la Sante Et de la Recherche Medicale; UJF University of Grenoble.

References

Abbasi, A., Hossain, L., & Leydesdorff, L. (2012). Betweenness centrality as a driver of preferential attachment in the evolution of research collaboration networks. *Journal of Informetrics*, *6*(3), 403–412.
Bergé, L., Wanzenböck, I., & Scherngell, T. (2017). Centrality of regions in R&D networks: A new measurement approach using the concept of bridging paths. *Regional Studies*, *51*(8), 1165–1178.
Bergé, L. (2017). Network proximity in the geography of research collaboration. *Papers in Regional Science* [in press].
Bonaccorsi, A. (2008). Search regimes and the industrial dynamics of science. *Minerva*, *46*(3), 285–315.
Borgatti, S. P., & Halgin, D. S. (2011). On network theory. *Organization Science*, *22*(5), 1168–1181.
Borgatti, S. P. (2005). Centrality and network flow. *Social Networks*, *27*(1), 55–71.
Breschi, S., & Lenzi, C. (2015). The role of external linkages and gatekeepers for the renewal and expansion of US cities knowledge base, 1990–2004. *Regional Studies*, *49*(5), 782–797.
Burt, R. S. (1992). *Structural holes: the social structure of competition*. Cambridge, MA: Harvard University Press.
Edler, J., Fier, H., & Grimpe, C. (2011). International scientist mobility and the locus of knowledge and technology transfer. *Research Policy*, *40*(6), 791–805.
Fischer, M. M., Scherngell, T., & Jansenberger, E. M. (2006). The geography of knowledge spillovers between high-technology firms in Europe. Evidence from a spatial interaction modelling perspective. *Geographical Analysis*, *38*(3), 288–309.
Glänzel, W., & Schubert, A. (2004). Analysing scientific networks through co-authorship. In *Handbook of quantitative science and technology research, pages*. pp. 257–276. Springer.
Granovetter, M. S. (1973). The strength of weak ties. *American Journal of Sociology*, *78*(6), 1360–1380.
Grossman, G. M., & Helpman, E. (1991). Trade, knowledge spillovers, and growth. *European Economic Review*, *35*(2-3), 517–526.
Guns, R., & Rousseau, R. (2015). Unnormalized and normalized forms of gefura measures in directed and undirected networks. *Frontiers of Information Technology & Electronic Engineering*, *16*(4), 311–320.
Hu, Y., Scherngell, T., Man, S. N., & Wang, Y. (2013). Is the United States still dominant in the global pharmaceutical innovation network? *Public Library of Science*, *8*(11), e77247.
Hwang, W., Cho, Y. R., Zhang, A., & Ramanathan, M. (2006). Bridging centrality: identifying bridging nodes in scale-free networks. In *Proceedings of the 12th ACM SIGKDD international conference on Knowledge discovery and data mining* (pp. 20–23).
Kahane, B., Mogoutov, A., Cointet, J. P., Villard, L., & Laredo, P. (2015). *A dynamic query to delineate emergent science and technology: the case of nano science and technology. Content and technical structure of the Nano S&T Dynamics Infrastructure*. www.sciences-technologies.eu/images/stories/cv/Dynamic_query%20to_delineate_ emergent_science_ technology_pub.pdf
Karaulova, M., Shackleton, O., Gok, A., Kotsemir, M. N., & Shapira, P. (2014). *Nanotechnology research and innovation in Russia: a bibliometric analysis*. [Available at SSRN 2521012.].
Larédo, P., Delemarle, A., & Kahane, B. (2010). Dynamics of nanosciences and technologies: policy implications. *STI Policy Review*, *1*, 43–62.
Lazega, E., & Snijders, T. (Eds.). (2015). Multilevel network analysis for the social sciences. In. Springer.
Lepori, B., Barré, R., & Filliatreau, G. (2008). New perspectives and challenges for the design and production of s&t indicators? *Research Evaluation*, *17*(1), 33–44.
Newman, M. E. (2001). Scientific collaboration networks. I. Network construction and fundamental results. *Physical Review E*, *64*(1), 016131.
OECD. (2005). *Oslo manual: guidelines for collecting and interpreting innovation data*. Paris: Organisation for Economic Cooperation and Development (OECD).
Robinson, W. S. (1950). Ecological correlations and the behavior of individuals. *American Sociological Review*, *15*(3), 351–357.
Rodriguez, M. A., & Pepe, A. (2008). On the relationship between the structural and socioacademic communities of a coauthorship network. *Journal of Informetrics*, *2*, 195–201.
Scherngell, T., & Barber, M. (2011). Distinct spatial characteristics of industrial and public research collaborations: Evidence from the 5th EU Framework Programme. *The Annals of Regional Science*, *46*(2), 247–266.
Scherngell, T., & Lata, R. (2013). Towards an integrated European Research Area? Findings from Eigenvector spatially filtered spatial interaction models using European Framework Programme data. *Papers in Regional Science*, *92*, 555–577.
Scherngell, T. (Ed.). (2013). *The geography of networks and R&D collaborations*. Berlin-Heidelberg-New York: Springer- Physica Verlag.
Snijders, T. (2016). The multiple flavours of multilevel issues for networks. In *Multilevel Network Analysis for the Social SciencesMultilevel network analysis for the social sciences*. pp. 15–46. Springer.
Wanzenböck, I., Scherngell, T., & Brenner, T. (2014). Embeddedness of regions in European knowledge networks. A comparative analysis of inter-regional R&D collaborations, co-patents and co- publications. *The Annals of Regional Science*, *53*, 337–368.

Wanzenböck, I., Scherngell, T., & Lata, R. (2015). Embeddedness of European regions in EU funded R&D networks: A spatial econometric perspective. *Regional Studies*, *49*(10), 1685–1705.
Wasserman, S., & Faust, K. (1994). . *Social network analysis: Methods and applications* (vol. 8) Cambridge University Press.
Yan, X., Zhai, L., & Fan, W. (2013). C-index: A weighted network node centrality measure for collaboration competence. *Journal of Informetrics*, *7*(1), 223–239.