Gauging the bridging function of nodes in a network: Q-measures for networks with a finite number of subgroups

Gauging the bridging function of nodes in a network: Q-measures for networks with a finite number of subgroups

Raf Guns[1] and Ronald Rousseau[2,3,4]

[1] *raf.guns@ua.ac.be*
University of Antwerp, IBW, Venusstraat 35, City Campus, B-2000 Antwerpen, Belgium

[2] *ronald.rousseau@khbo.be*
KHBO (Association K.U.Leuven), Industrial Sciences and Technology,
Zeedijk 101, B-8400 Oostende, Belgium
[3] Hasselt University, Universitaire Campus, B-3590 Diepenbeek, Belgium
[4] K.U.Leuven, Department of Mathematics, Celestijnenlaan 200B, B-3001 Leuven (Heverlee), Belgium

Abstract

We generalize the idea of Q-measures for two groups in a network to the case that there are n groups ($n \geq 2$). The generalized Q-measure is an indicator of the bridging function of nodes, describing their role as a direct or indirect connection between groups in the network. It is shown that there are several possible generalizations, including a global and a local indicator. Two artificial examples and a small informetric case study in journal similarity are analyzed. Finally, several possible applications in informetrics, social network analysis (SNA) and beyond are outlined.

Introduction

Flom et al. (2004) introduced Q-measures as indicators for the bridging function of nodes between two subnetworks in a connected, undirected, unweighted network. The notion of a Q-measure has further been studied in (Rousseau, 2005; Chen & Rousseau, 2008; Rousseau & Zhang, 2008). In this article we indicate how Q-measures can easily be defined for undirected, unweighted networks subdivided into more than two groups. This idea has already been proposed, without giving details, by Brandes (2008).

Our approach differs from the one taken by Everett and Borgatti (1999) where entire groups of nodes are the units of study whereas the units of study in our article are the nodes themselves.

This article is composed as follows: In the following section we recall the definitions of different types of Q-measures as originally defined for a network with two groups. Subsequently we give two approaches (a global and a local one) for determining Q-measures in case there are more than two subgroups, and we discuss some theoretical and practical issues. In the next sections we provide some examples, both artificial and real, that illustrate the calculation of the global and the local Q-measure and the difference between them. The last two sections outline several ideas for applications and contain the conclusions.

Q-measures for two groups (Flom et al., 2004)

In (Flom et al., 2004) the authors point out that most network studies either study interactions within one network, where all actors (possibly) interact, or distinguish two groups and interactions only occur between groups and not within one group (so-called bipartite networks). They call attention to an additional possibility, where interactions are possible between all nodes, but the network is subdivided in two or more groups and the focus of the study is on the way different groups are connected and the role played by actors in establishing these connections. In informetrics one may for instance be interested in the collaboration network between universities. This network is divided into several groups according to countries or larger regions. One is then interested to know which universities play a bridging function facilitating collaborations between different countries. In another field, Trotter, Rothenberg and Coyle (1995) wrote that the idea of identifying bridges in a network is a key area for future research in HIV prevention and public health interventions.

Flom et al. (2004) further remark that classical centrality measures do not make a distinction between nodes of different groups, or between geodesics – i.e., the shortest path(s) between two nodes – which remain in one group and those which cross to other groups. For that reason they introduce the notion of Q-measures. Their Q-measure is defined as follows. Let us assume that there are T actors or nodes in the network and that the network is divided into two subgroups, denoted as G and H. Group G contains m nodes, while group H contains n nodes, hence $T = m + n$. If actor a belongs to group G, and assuming that actor a is g_k, then the Q-measure for this actor is defined as:

$$Q(a) = \frac{1}{(m-1)n}\left(\sum_{\substack{i=1\\ i\neq k}}^{m}\sum_{j=1}^{n}\frac{p_{g_i h_j}(a)}{p_{g_i h_j}}\right) \tag{1}$$

If actor a belongs to group H, and assuming that it is actor h_k, then its Q-measure is defined as:

$$Q(a) = \frac{1}{m(n-1)}\left(\sum_{i=1}^{m}\sum_{\substack{j=1\\ j\neq k}}^{n}\frac{p_{g_i h_j}(a)}{p_{g_i h_j}}\right) \tag{2}$$

where $p_{g_i h_j}$ denotes the number of geodesics, i.e. shortest paths, connecting $g_i \in$ G and $h_j \in$ H. The symbol $p_{g_i h_j}(a)$ represents the number of geodesics connecting g_i and h_j passing through a, where a is not one of the endpoints. Flom et al. (2004) make a further distinction between Q_1 for which only single-cross geodesics are included in the definition, and Q_2 for which all geodesics are included (also those that cross groups several times). Finally, they also define a Q-measure characterizing the whole network, but as we will not study this notion we refer the reader to the original article (Flom et al., 2004) for the definition and further details.

Rousseau and Zhang (2008) observed that, although formulae (1) and (2) are normalized sums of quotients, in directed valued networks it is preferable to define Q-measures as quotients of sums. So it seems at least reasonable to explore an alternative for formulae (1) and (2) and to discuss the case that Q-measures of two groups are defined as quotients of sums. This is done in the appendix where it is concluded that probably the original definition reflects best our intuition of playing the role of a bridge between two groups in an unweighted network.

In the next section we propose generalizations of Q-measures for the case that more than two groups can be distinguished in the network.

How to calculate a Q-measure in the general case of a network subdivided into a finite number of subgroups

Assume that a connected, undirected and unweighted network consists of S (≥ 2) disjoint groups. We will not make a distinction between single-cross geodesics and general ones, and will only consider the general case. Hence we will impose no restrictions on shortest paths. Let group G_k, $k = 1,\ldots,S$ consist of m_k members and let $\sum_{k=1} m_k = M$ be the total number of nodes or actors in the network.

We propose two types of Q-measures, each generalizing the case of two groups. The first one is the one we refer to as a global Q-measure, the second one is a local Q-measure.

Global Q-measure

The global Q-measure for actor a is defined as:

$$Q_G(a) = \frac{1}{C}\sum_{k,l}\left(\frac{1}{TP_{k,l}}\sum_{\substack{g\in G_k\\ h\in G_l}}\frac{p_{g,h}(a)}{p_{g,h}}\right) \tag{3}$$

where C = $\binom{S}{2} = \frac{S(S-1)}{2}$ is the number of ways to choose two different groups and $\frac{1}{TP_{k,l}}\sum_{\substack{g\in G_k\\ h\in G_l}}\frac{p_{g,h}(a)}{p_{g,h}}$, which we will denote as $Q_{k,l}(a)$, is the Q-measure of *a* with respect to groups G_k and G_l (as defined in the previous section). In the context of Q-measures for several groups such a Q-measure is called a partial Q-measure. Symbols $p_{g,h}$ and $p_{g,h}(a)$ are defined as in formulae (1) and (2). $\sum_{\substack{g\in G_k\\ h\in G_l}}\frac{p_{g,h}(a)}{p_{g,h}}$ is divided by $TP_{k,l}$, the total number of possible paths. This is $m_k \cdot m_l$ if *a* does not belong to G_k or G_l. It is $(m_k - 1)\cdot m_l$ if *a* belongs to G_k and $m_k\cdot(m_l - 1)$ if *a* belongs to G_l. If $m_k - 1$ or $m_l - 1$ is zero the corresponding *TP* is taken to be 1 (or any finite number). The exact value plays no role as $p_{g,h}(a) = 0$ in this case.
The global Q-measure of node *a* always satisfies the inequality $0 \le Q_G(a) \le 1$.

Local Q-measure
The local Q-measure for actor *a* belonging to group G_k with m_k nodes is defined as:

$$Q_L(a) = \frac{1}{S-1}\sum_{l\neq k}\left(\frac{1}{(m_k - 1)\cdot m_l}\sum_{\substack{g\in G_k\\ h\in G_l}}\frac{p_{g,h}(a)}{p_{g,h}}\right) \tag{4}$$

Note that formula (4) uses partial Q-measures as well, but the cases where *a* does not belong to G_k are not used here. Just as for the global Q-measure, the local Q-measure of node *a* satisfies the inequality $0 \le Q_L(a) \le 1$.
When $S = 2$, $Q_L = Q_G$ = the Q-measure as defined by Flom et al. (2004).

Discussion
There are several valid options when studying the bridging function of nodes in networks with several groups, as illustrated in Figure 1. We briefly discuss them.

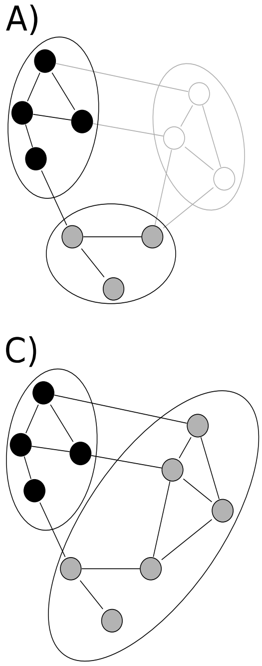

Figure 1. Possible approaches to studying bridges between groups, where light gray indicates nodes and links that are not used. A) Only look at target group and other group, B) Local Q-measures, C) Outside world view, D) Global Q-measures

A) Consider a target group and each other group separately. One considers only shortest paths that connect nodes of the target group to the other group. Paths passing through other nodes may not be used. This is a special case of the theory of Q-measures for two groups.
B) Consider a target group and each other group separately. One considers only shortest paths that connect nodes of the target group to the other group. Paths passing through other nodes may be used in this case, but the roles played by those other nodes are not taken into account. For two groups this is another special case of the theory of Q-measures for two groups. For three or more groups this Q-measure indicates a node's bridging function between its own group and the other groups. Averaging leads to the local Q-measure shown in formula (4).
C) Consider a target group and the union of all other groups. This union can be considered as the 'outside world'. This situation is just a special case of the theory of Q-measures for two groups.
D) Finally one can consider the global approach as described by formula (3). Here the total topological structure plays a role in determining a node's bridging function between all different groups in the network.

Of course one may also average case A, but it seems counterintuitive to first ignore all other groups and then take them into account through an averaging process.
We note that a singleton group always has a local Q-measure equal to 0, but this is not necessarily the case for the global measure. On the contrary, such a group can have the largest value in the network: consider, for instance, the case of two countries situated on opposite sides of a high mountain range, where a city situated on the mountain pass is an independent city state. Depending on the network topology this single city state can have the city with the largest global Q-value (examples are given in the next section).
If each node in a network is considered as a singleton group then $S = M$ and $Q_G(a)$ of node a in this network is equal to the betweenness centrality of a divided by $C = \binom{S}{2} = \frac{S(S-1)}{2}$.

Finally, one may imagine circumstances, e.g. real-world transportation networks, where normalization, leading to a relative measure, is not optimal. Then no division by C (in the global case), by $S-1$ (in the local case) or even TP is performed. This yields an absolute measure of 'bridgeness'. In particular, if no normalization is performed $Q_G(a)$ becomes equal to the betweenness centrality if each node is considered as a singleton group.

Note

In some cases, one group may be considered more important than another. In a transportation network for example, it is more important to be a bridge between two capital regions than between two peripheral regions. Q-measures can easily be adapted in case the different subgroups have unequal weights. This observation leads to the generalized global Q-formula:

$$Q_G(a) = \sum_{k,l} w_{k,l} \cdot Q_{k,l}(a) \text{ with } \sum_{k,l} w_{k,l} = 1 \tag{5}$$

And the generalized local Q-formula:

$$Q_L(a) = \sum_{l \neq k} \frac{w_l}{(m_k - 1)\cdot m_l} \sum_{\substack{g \in G_k \\ h \in G_l}} \frac{p_{g,h}(a)}{p_{g,h}} \text{ with } \sum_{l \neq k} w_l = 1 \tag{6}$$

Two artificial examples

Example 1

Consider the 8-node network depicted in Figure 2, consisting of three groups: $G_1 = \{a_1, a_2, a_3\}$, $G_2 = \{b_1, b_2\}$ and $G_3 = \{c_1, c_2, c_3\}$.

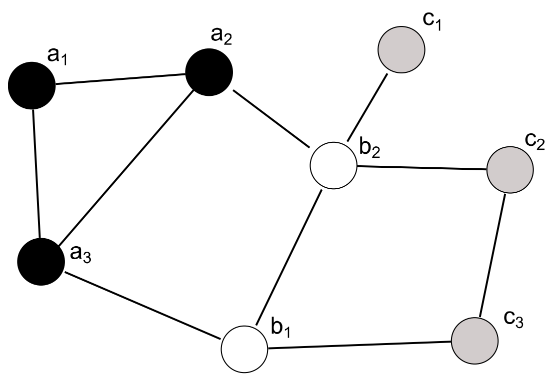

Figure 2. Example network with nodes from three groups

Following Flom et al.'s suggestion (Flom et al., 2004), we describe the relation between any two subgroups in a table, where the nodes in the cells are nodes situated on a shortest path connecting the column and the row nodes (but not one of these endpoints). If there are several shortest paths, these are grouped by accolades. If the column and row nodes are directly connected a '–' is placed in the corresponding cell.

Table 1. Shortest paths between G_1 and G_2

	a_1	a_2	a_3
b_1	a_3	$\{a_3, b_2\}$	–
b_2	a_2	–	$\{a_2, b_1\}$

From Table 1 we derive: $Q_{1,2}(a_1) = 0$; $Q_{1,2}(a_2) = 3/8$; $Q_{1,2}(a_3) = 3/8$;
$Q_{1,2}(b_1) = 1/6$; $Q_{1,2}(b_2) = 1/6$;
$Q_{1,2}(c_1) = 0$; $Q_{1,2}(c_2) = 0$; $Q_{1,2}(c_3) = 0$.

Table 2. Shortest paths between G_1 and G_3

	a_1	a_2	a_3
c_1	$b_2 - a_2$	b_2	$\begin{cases} b_2 - a_2 \\ b_2 - b_1 \end{cases}$
c_2	$b_2 - a_2$	b_2	$\begin{cases} b_2 - a_2 \\ c_3 - b_1 \\ b_2 - b_1 \end{cases}$
c_3	$b_1 - a_3$	$\begin{cases} b_1 - a_3 \\ b_1 - b_2 \\ c_2 - b_2 \end{cases}$	b_1

From Table 2 we derive: $Q_{1,3}(a_1) = 0$; $Q_{1,3}(a_2) = 17/36$; $Q_{1,3}(a_3) = 4/18$;
$Q_{1,3}(b_1) = 23/54$; $Q_{1,3}(b_2) = 19/27$;
$Q_{1,3}(c_1) = 0$; $Q_{1,3}(c_2) = 1/18$; $Q_{1,3}(c_3) = 1/18$.

Table 3. Shortest paths between G_2 and G_3

	c_1	c_2	c_3
b_1	b_2	$\begin{cases} c_3 \\ b_2 \end{cases}$	–
b_2	–	–	$\begin{cases} c_2 \\ b_1 \end{cases}$

From Table 3 we derive: $Q_{2,3}(a_1) = 0$; $Q_{2,3}(a_2) = 0$; $Q_{2,3}(a_3) = 0$;
$Q_{2,3}(b_1) = 1/6$; $Q_{2,3}(b_2) = 3/6$;
$Q_{2,3}(c_1) = 0$; $Q_{2,3}(c_2) = 1/8$; $Q_{2,3}(c_3) = 1/8$.

These three tables and the derived partial Q-measures can be used to determine the global as well as the local Q-measures. For the calculation of the local Q-measure we use the partial Q-measures $Q_{k,l}(s)$ only if $s \in G_k$ or $s \in G_l$ (see formula (4)).

Table 4. Calculation of the Q-measures for the artificial example

Global Q	Local Q
$Q_G(a_1) = \frac{1}{3}(0+0+0) = 0$	$Q_L(a_1) = \frac{1}{2}(0+0) = 0$
$Q_G(a_2) = \frac{1}{3}\left(\frac{3}{8} + \frac{17}{36} + 0\right) = \frac{61}{216} \approx 0.28$	$Q_L(a_2) = \frac{1}{2}\left(\frac{3}{8} + \frac{17}{36}\right) = \frac{61}{144} \approx 0.42$

$Q_G(a_3)=\frac{1}{3}\left(\frac{3}{8}+\frac{4}{18}+0\right)=\frac{43}{216}\approx 0.20$	$Q_L(a_3)=\frac{1}{2}\left(\frac{3}{8}+\frac{4}{18}\right)=\frac{43}{144}\approx 0.30$
$Q_G(b_1)=\frac{1}{3}\left(\frac{1}{6}+\frac{23}{54}+\frac{1}{6}\right)=\frac{41}{216}\approx 0.25$	$Q_L(b_1)=\frac{1}{2}\left(\frac{1}{6}+\frac{1}{6}\right)=\frac{1}{6}\approx 0.17$
$Q_G(b_2)=\frac{1}{3}\left(\frac{1}{6}+\frac{19}{27}+\frac{3}{6}\right)=\frac{37}{81}\approx 0.46$	$Q_L(b_2)=\frac{1}{2}\left(\frac{1}{6}+\frac{3}{6}\right)=\frac{1}{3}\approx 0.33$
$Q_G(c_1)=\frac{1}{3}(0+0+0)=0$	$Q_L(c_1)=\frac{1}{2}(0+0)=0$
$Q_G(c_2)=\frac{1}{3}\left(0+\frac{1}{18}+\frac{1}{8}\right)=\frac{13}{216}\approx 0.06$	$Q_L(c_2)=\frac{1}{2}\left(\frac{1}{18}+\frac{1}{8}\right)=\frac{13}{144}\approx 0.09$
$Q_G(c_3)=\frac{1}{3}\left(0+\frac{1}{18}+\frac{1}{8}\right)=\frac{13}{216}\approx 0.06$	$Q_L(c_3)=\frac{1}{2}\left(\frac{1}{18}+\frac{1}{8}\right)=\frac{13}{144}\approx 0.09$

We can see in Table 4 that the relative importance of nodes b_1 and b_2 is lower in the local case than in the global case, where their prominence was derived to a large extent from their bridging position between G_1 and G_3.
Node a_2 becomes the most important node, considered from the point of view of the local group, i.e. using Q_L. Also node a_3 has a higher Q_L value than node b_1 (while it is the other way around for Q_G).

Example 2: single-city state

We now turn to another example, which was already briefly mentioned in the preceding section: the single-city state.

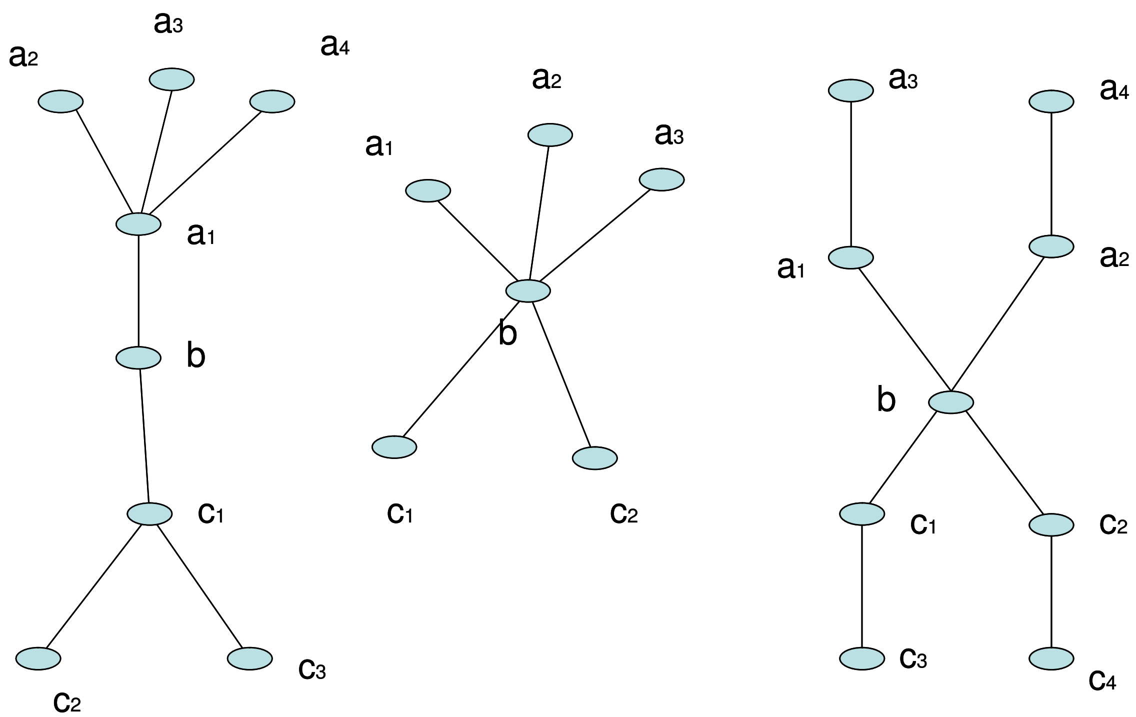

Figure 3. The single-city state. Cases I, II and III

We consider three cases. In case I the single-city state is connected to one city in country A, which is connected to all other A cities (m−1 other cities), and one city in country C, which is

connected to all C cities ($n-1$ cities). Here $Q_G(a_1) = Q_G(c_1) = 1$; $Q_G(b) = 1/3$. All other nodes have $Q_G = 0$. As for the local Q-measures, we find that $Q_L(a_1) = Q_L(c_1) = 1$. All other nodes, including b, have $Q_L = 0$.
In case II the single-city state is connected to all (m) A cities and all (n) C cities. Here $Q_G(b) = 1/3$, and all other nodes have $Q_G = 0$. Since the only node in a one-member group always has $Q_L = 0$ (as mentioned earlier), in this case all nodes have $Q_L = 0$.
Finally in case III, the single-city state is connected to half of the A cities ($a_1, \ldots, a_{m/2}$), (we assume m to be even), which are each connected to one other A city, and to half of the C cities ($c_1, \ldots, c_{n/2}$), (assuming also n to be even), each connected to one other C city.
Here $Q_G(b) = 1/3$; $Q_G(a_1) = \ldots = Q_G(a_{m/2}) = \frac{1}{3}\left(\frac{1}{m-1} + \frac{n}{(m-1)n}\right) = \frac{2}{3(m-1)}$; $Q_G(c_1) = \ldots = Q_G(c_{n/2}) = \frac{1}{3}\left(\frac{1}{n-1} + \frac{m}{(n-1)m}\right) = \frac{2}{3(n-1)}$; all other nodes have Q_G equal to zero. As for the local Q-measures, we find that $Q_L(a_1) = \ldots = Q_L(a_{m/2}) = \frac{1}{2}\left(\frac{1}{m-1} + \frac{1}{m-1}\right) = \frac{1}{m-1}$; $Q_L(c_1) = \ldots = Q_L(c_{n/2}) = \frac{1}{2}\left(\frac{1}{n-1} + \frac{1}{n-1}\right) = \frac{1}{n-1}$. All other nodes, including b, have $Q_L = 0$.
Depending on the situation the single-city state has (case II, and case III with $m, n > 3$) or has not (case I, and case III with m or $n \leq 3$) the largest Q_G value. Its Q_L value on the other hand is always equal to 0.

A real example: the citation environment of *Nano Letters*

Leydesdorff (2007b) studies social network measures in the local citation environment of a number of journals. The local citation environment of a journal J consists of all journals that cite (or are cited by) J up to one percent of its citation rate (Leydesdorff, 2007a). The similarity between journals in a citation environment is calculated with the well-known cosine measure (Salton & McGill, 1983). Similarity measures exceeding a threshold (≥ 0.2) indicate the most important connections. Thus, the local citation environment can be treated as an undirected network, where nodes are journals and links indicate a fair amount of similarity.
As a practical example, we study Q-measures in the local citation environment of the journal *Nano Letters*, an important journal in the field of nanotechnology. Since Leydesdorff (2007b) used this journal's citation environment in 2004 as an example, we will also determine Q-measures for the 2004 data, taken from http://www.leydesdorff.net/jcr04. In this year, 17 journals have cited *Nano Letters* more than 1% of its total number of citations.
Next, we look up each journal's category in the Journal Citation Reports (JCR). If a journal is assigned to more than one JCR category, it is assumed to belong to the largest of its categories (i.e., the category with the largest number of journals). The rationale for this is that we wanted to avoid a large number of singleton groups. Still, as shown in Table 5, three of the seven categories contain only one journal.
In the corresponding network, there are thus 17 nodes belonging to 7 different groups. Global and local Q-measures were calculated using formulae (3) and (4), summarized in Table 5. In general, we find that there is a clear correlation between Q_G and Q_L ($r = 0.73$), between Q_G and betweenness centrality ($r = 0.97$), and between Q_L and betweenness centrality ($r = 0.80$).
Figure 4 illustrates these correlations: the horizontal size of nodes corresponds to their global Q-measure, while the vertical size corresponds to their local Q-measure. From the figure, it can be readily seen that the three 'nano' journals primarily function as bridges between other groups: their global Q-measure is much higher than their local one. Although *Chemical Physics Letters* is important both globally and locally, its vertical size suggests that this

journal's primary bridging role is in connecting other journals in physical chemistry with journals in other disciplines. Most of the other chemistry and physics journals are much less important, as was already evidenced by their low betweenness centrality (Leydesdorff, 2007b).

Table 5. Overview of JCR categories, journals and their Q-measures

JCR category	Journal	Q_G	Q_L
CHEMISTRY, MULTIDISCIPLINARY	AngewChemIntEdit	0.000	0.000
	ChemCommun	0.005	0.009
	JAmChemSoc	0.005	0.009
CHEMISTRY, PHYSICAL	ChemPhysLett	0.085	0.126
	JPhysChemB	0.005	0.009
	Langmuir	0.005	0.009
MATERIALS SCIENCE, MULTIDISCIPLINARY	AdvMater	0.032	0.009
	ChemMater	0.015	0.004
	JMaterChem	0.015	0.004
	JNanosciNanotechno	0.164	0.057
	NanoLett	0.082	0.028
	Nanotechnology	0.075	0.053
PHYSICS, APPLIED	ApplPhysLett	0.014	0.009
	JApplPhys	0.004	0.000
PHYSICS, CONDENSED MATTER	PhysRevB	0.000	0.000
PHYSICS, MULTIDISCIPLINARY	PhysRevLett	0.001	0.000
POLYMER SCIENCE	Macromolecules	0.000	0.000

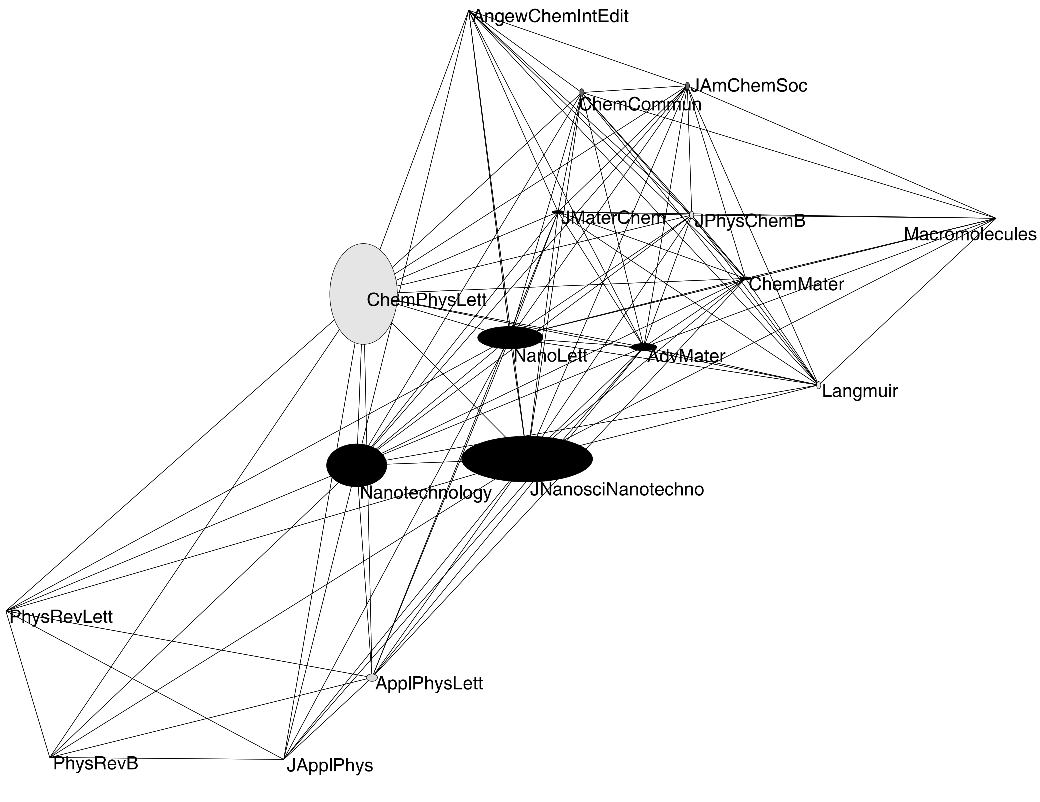

Figure 4. Q-measures of 17 journals in the citation environment of *Nano Letters* in 2004

Ideas for applications

In most network studies a lot of attention has gone to the centrality of nodes, measured in different ways. Yet, locating nodes that play a bridging function and putting a numerical value to this role, opening the possibility of ranking nodes with respect to their bridging function, is equally important.

Research areas where the theory of Q-measures can be applied are manifold; some examples are listed in Table 6. The work of Chen and Rousseau (2008) is an example of the third case, restricted to the *Journal of Fluid Mechanics* and just two countries (Germany and England).

Note that examples 5 and 11 usually imply a directed network. Indeed, we should point out that all formulae for Q-measures – the original Q-measures for two groups as well as Q-measures for any finite number of groups – can be applied to directed networks as well. In fact, Brandes' (2008) algorithm for Q-measures treats undirected networks as symmetric directed ones. While not explicitly studied here, this generalization is straightforward.

Table 6. Example applications of Q-measures for *n* groups

	Network	Nodes	Groups	Types of linkages
1	University	Researchers	Departments	Co-authorship
2	Several countries (as e.g. the European Union)	Universities	Countries	Co-authorship
3	Several countries (as e.g. the European Union)	Universities	Countries	Co-authorship, restricted to one particular field; or in one particular journal (Chen & Rousseau, 2008)
4	Company	Personnel	Divisions	"has formal meetings with" "has informal meetings with"
5	University Web	Departments	Faculties or departments studying similar topics	Weblinks (Björneborn, 2006)
6	World-wide airport network (WAN)	Airports	Countries	"have a direct airline connection" (Barrat et al., 2004)
7	Navigation routes	Harbours	Countries	Trade routes
8	Primate behaviour	Monkeys	Age groups	Interactions = joint presence at a river (Everett & Borgatti, 1999)
9	Human Disease Network	Diseases	Disorder classes	Diseases share at least one gene in which mutations are associated with both disorders (Goh et al., 2007)
10	Schools	Pupils	Class groups	Friendship
11	Football match	Players	Position (4 groups: goalkeeper, defenders, midfielders, forwards)	"passes to"

Conclusion

As attention of informetricians and policy makers for collaboration and interdisciplinarity increases, there exists a need for quantitative indicators in these and related areas. While well-known measures like betweenness centrality can be an indicator of interdisciplinarity (Leydesdorff, 2007b), Q-measures have the obvious advantage of taking subgroups into account. The original definition of Q-measures is an indicator of 'bridgeness' between only two groups, limiting its potential use. However, it forms the basis for our generalized Q-measures for n groups ($n \geq 2$).

A global as well as a local Q-measure can be defined. While the former indicates the extent to which a node bridges between all pairs of groups in the network, the latter only indicates a node's bridging function for its own group. These measures were illustrated with artificial examples and a small informetric application. Finally, several possible application examples were given where these measures could provide a deeper insight in an existing network.

Acknowledgement

The authors thank two anonymous reviewers for their suggestion to add a 'real' informetric example.

References

Barrat, A., Barthélemy, M., Pastor-Satorras, R. and Vespignani, A. (2004). The architecture of complex weighted networks. *Proceedings of the National Academy of Sciences of the USA,* 101 (11), 3747-3752.

Björneborn, L. (2006). 'Mini small worlds' of shortest link paths crossing domain boundaries in an academic Web space. *Scientometrics*, 68, 395-414.

Brandes, U. (2008). On variants of shortest-path betweenness centrality and their generic computation. *Social Networks*, 30, 136-145.

Chen, LX. and Rousseau, R. (2008). Q-measures for binary divided networks: bridges between German and English institutes in publications of the *Journal of Fluid Mechanics. Scientometrics,* 74, 57-69.

Everett, M.G. and Borgatti, S.P. (1999). The centrality of groups and classes. *Journal of Mathematical Sociology*, 23, 181-201.

Flom, P.L., Friedman, S.R., Strauss, S. and Neaigus, A. (2004). A new measure of linkage between two sub-networks. *Connections*, 26 (1), 62-70.

Goh, K-I., Cusick, M.E., Valle, D., Childs, B., Vidal, M. and Barabási, A.L. (2007). The human disease network. *Proceedings of the National Academy of Sciences of the USA,* 104 (21), 8685-8690.

Leydesdorff, L. (2007a). Visualization of the citation impact environment of scientific journals: an online mapping exercise. *Journal of the American Society for Information Science and Technology*, 58 (1), 25-38.

Leydesdorff, L. (2007b). Betweenness centrality as an indicator of the interdisciplinarity of scientific journals. *Journal of the American Society for Information Science and Technology*, 58 (9), 1303-1319.

Rousseau, R. (2005). Q-measures for binary divided networks: an investigation within the field of informetrics. *Proceedings of the 68th ASIST Conference*, vol. 42, 675-696.

Rousseau, R. and Zhang, L. (2008). Betweenness centrality and Q-measures in directed valued networks. *Scientometrics*, 75, 575-590.

Salton, G. and McGill, M.J. (1983). *Introduction to Modern Information Retrieval*. Auckland: McGraw-Hill.

Trotter, R.T., Rothenberg, R.B. and Coyle, S. (1995). Drug abuse and HIV prevention research: expanding paradigms and network contributions to risk reduction. *Connections*, 18, 29-45.

Appendix: Another definition for a Q-measure for two groups

Rousseau & Zhang (2008) noted already that formulae (1) and (2) are normalized sums of quotients, and they defined Q-measures in directed valued networks as quotients of sums. Now we will reflect upon their proposal and redefine the Q-measure of two groups as a quotient of sums. Next we discuss the difference.

As in the original definition we assume that there are T actors or nodes in the network and that the network is subdivided into two subgroups, G and H. Group G contains m nodes, while H contains n nodes. If actor x belongs to group G, and assuming that actor x is g_k, then the alternative Q-measure for this actor, denoted as $Q_{alt}(x)$ is defined as:

$$Q_{alt}(x) = \frac{\sum_{\substack{i=1 \\ i \neq k}}^{m} \sum_{j=1}^{n} c_{g_i h_j}(x)}{\sum_{\substack{i=1 \\ i \neq k}}^{m} \sum_{j=1}^{n} c_{g_i h_j}} \tag{7}$$

If actor x belongs to group H, and assuming that it is actor h_k, then its Q_{alt} measure is defined as:

$$Q_{alt}(x) = \frac{\sum_{i=1}^{m} \sum_{\substack{j=1 \\ j \neq k}}^{n} c_{g_i h_j}(x)}{\sum_{i=1}^{m} \sum_{\substack{j=1 \\ j \neq k}}^{n} c_{g_i h_j}} \tag{8}$$

where $c_{g_i h_j}$ denotes the number of geodesics connecting $g_i \in$ G and $h_j \in$ H. The symbol $c_{g_i h_j}(x)$ represents the number of geodesics connecting g_i and h_j passing through x, where x is not one of the endpoints.

We first note that there is no difference between the two types of formulae when there is exactly one shortest path between any two nodes in the network.

The original formula favours nodes that lie on shortest paths between many pairs of nodes, while the alternative one favours nodes that lie on many shortest paths (but, maybe connecting relatively few pairs of nodes). It seems to us that when there are no weights involved the original version corresponds best to our intuition of the idea of bridging function. Yet the proposals shown in equations (7) and (8) are mathematically valid alternatives.