Measuring Integration in the Network Structure: Some Suggestions on the Connectivity Index

1. Measures of Connectivity

The connectivity can be divided into two levels, one is domestic connectivity, in the case of the physical connectivity as covered in the report appears to focus on the domestic connectivity. The other is connectivity between the APEC economies. That is cross-board connectivity, a network measure. The network can be envisaged as a collection of nodes, each of which represents an economy.

The structure of the network is given by a link matrix, denote it as $\mathrm{L} = (l_{ij})_{n\times n}$,where a typical element $l_{ij} = 1$ if there is a direct link going from economy i to economy j; otherwise, it is equal to zero. The activity on each link of the network is measured in terms of the flow on the network, which is summarized in the matrix $\mathrm{X} = (x_{ij})_{n\times n}$. A typical element x_{ij}, is the total flow from economy i to economy j, and x_{ji} is the flow in the other direction from j to i (i.e. transport networks are bi-directional).

Matrix X:

$$\mathrm{X} = \begin{pmatrix} X_{11} & X_{12} & \cdots & X_{1n} \\ X_{21} & X_{22} & \cdots & X_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ X_{n1} & X_{n2} & \cdots & X_{nn} \end{pmatrix}_{n\times n}$$

Note:

(1) The cross-board connectivity is less involved as an indicator in the report. Whereas for most economies, the domestic connectivity is fairly advanced, the cross-board connectivity among them such as transnational roads, energy pipelines, power grids and other infrastructure is far less developed, which is a weak link of the regional connectivity and also a bottleneck for the regional

development. Therefore a more balanced system of indicators should be designed to reflect the cross-board connectivity in the index.

(2) In order to reflect specifically the connectivity between pairs of economies, the raw connectivity measures data should be bilateral level.

For example:

UNCTAD presents Liner Shipping Bilateral Connectivity Index (LSBCI), which indicates a country pair’s integration level into global liner shipping networks. The LSBCI is an extension of UNCTAD’s country-level Liner Shipping Connectivity Index (LSCI) and based on a proper bilateralization transformation.

- The current version of the LSBCI includes 5 components. For any pair of countries A and B represented in the sample, the LSBCI is based on:
  - ✓ the number of transshipments required to get from country A to country B;
  - ✓ the number of direct connections common to both country A and B;
  - ✓ the geometric mean of the number of direct connections of country A and of country B;
  - ✓ the level of competition on services that connect country A to country B;
  - ✓ the size of the largest ships on the weakest route connecting country A to country B.
- In order to establish a unit free index, all components are normalized using the standard formula: Normalized_Value = (Raw - Min(Raw)) / (Max(Raw) - Min(Raw)).
- The LSBCI is computed by taking the simple average of the five normalized components. As a consequence, the LSBCI can only take values between 0 (minimum) and 1 (maximum).

2. Methodology and Technique

A first intuitive approach to connectivity is simply to use the total outflow or inflow from each node:

$$\mathrm{X}_{i.} = \sum_j X_{ij} \quad \text{or} \quad \mathrm{X}_{.j} = \sum_i X_{ij}$$

Such simple output indicators obviously contain significant information, but by definition remain local rather than global, in the sense that they do not incorporate information from the full structure of the network. That is, they do not capture indirect connections, which are of particular importance in physical transport where connecting nodes are common. Nor do these types of measures reflect the interactions among nodes.

A more sophisticated approach, which makes use of more information from the matrix, is to use concentration indices such as the Herfindhal or Theil indices of the flows to and from a node in the network. The Theil index is the ***entropy*** of the relative weight of the outflows or inflows from node i to the neighboring node js, and is given by the following expression:

$$Theil\_Index = -\sum_j x_{ij} \ln(x_{ij}) = -\sum_j \frac{X_{ij}}{X_{i.}} \ln\left(\frac{X_{ij}}{X_{i.}}\right)$$

These kinds of concentration indicators have no dimension. However, they do not incorporate information about the structure of the network. They are essentially local measures, because they only use information from a single node. Relative entropy measures are better in this respect, since they compare the composition of flows to or from a node to an average composition. The most commonly used is the Kullback-Leibler distance, which is a modification of the Theil index. In this case, the reference composition should be the relative weight of each node.

Clustering is an important concept in network theory. The clustering coefficient of node i is an intuitive measure of how well connected the nodes in the neighborhood of i are. This number, comprised between 0 and 1. This definition applies equally to directional and non-directional networks; that is, a triangle is counted once for each direction it can be run, ijk and ikj. A variant of this definition applies to weighted networks. It gives higher weight to triangles with higher flows to the node:

$$Cluster\_Index_i = \frac{\sum_{j \neq k} \sum_k x_{ij} L_{jk} x_{ki}}{1 - \sum_j x_{ij} x_{ji}}$$

By construction, it not only includes information about interactions with neighboring nodes, but also about interactions between neighboring nodes.

A Gravity-Based Definition of Connectivity

It needs develop a measure of connectivity in the APEC economies' network, in the sense that it captures the full range of interactions among all network nodes, even when there is no direct connection between them, as in the Air Connectivity Index measuring.

The generic bi-proportional gravity model takes the following form:

$$\mathrm{X}_{ij} = \mathrm{A}_i \mathrm{B}_j \mathrm{K}_{ij}$$

where A_i is the repulsive potential of node i, and B_j is the attractive potential of node j; the flow X_{ij} is "pushed" from i and "pulled" to j. The bilateral variable K_{ij} measures the interaction between origin and destination. The essence of the gravity interpretation of spatial interaction models is a bi-proportional structure.

The potentials are estimated from the requirement that row and column totals in the gravity model estimates must equal the total outflows or inflows of the nodes. Thus:

$$\mathrm{X}_{i.} = \sum_j A_i B_j K_{ij} \quad \text{or} \quad \mathrm{X}_{.j} = \sum_i A_i B_j K_{ij}$$

As a consequence of the non-linear nature of the model, the potential of a node does not depend upon its own variables, but on every other interaction in the network. In the trade literature, a question is addressed using the concept of multilateral resistance (MR), which corrects for origin and destination interactions with the rest of the world. Let D_i be the pull exercised by destinations in the rest of the world on origin node i, and let O_j be the push exercised by origins in the rest of the world on destination node j. We can then define:

$$X_{i.} = A_i D_i \qquad D_i = \sum_j B_j K_{ij}$$

$$X_{.j} = B_j O_j \qquad O_j = \sum_i A_i K_{ij}$$

$$X_{ij} = X_{i.} D_i^{-1} X_{.j} O_j^{-1} K_{ij}$$

Total outflows are, thus, as expected roughly proportional to total outflows/inflows of the origin/destination multiplied by an impedance factor. This conclusion is maintained notwithstanding the correction for the pull and push from the rest of the world, or adjustment for multilateral resistance.

A first candidate for connectivity:

$$C_i = \frac{D_i}{\sum_j B_j - B_i} = \frac{\sum_{j \neq i} B_j K_{ij}}{\sum_{j \neq i} B_j}$$

The expression applies to the connectivity of outflows, but permuting A, B, D, and O gives the corresponding value for inflows. This expression can be interpreted in two ways. First, connectivity is equal to the average impedance K_{ij}, weighted by the potential of each partner. An alternative interpretation is that the numerator summarizes the pull or push of all partners, and the denominator represents the maximum possible pull or push.

The denominator in equation excludes an economy's own potential from the sum $\sum_{j \neq i} B_j$. This choice leads to some inconsistency. Hence a consistent definition of connectivity should also include an economy's own contribution to push and pull.

$$C_i = \frac{D_i + B_i}{\sum_j B_j} = \frac{X_{i.}/A_i + B_i}{\sum_j B_j}$$

This improved definition amounts to including in the flow matrix X a diagonal term, which corresponds to the effective flow between each economy and itself with an impedance of one.

For a symmetric definition of connectivity, we can take the geometric average of the connectivity of i as origin and destination. This approach gives:

$$\overline{C}_i = \left(\frac{X_{i.}/A_i + B_i}{\sum_j B_j} \right)^{\frac{1}{2}} \left(\frac{X_{.i}/B_i + A_i}{\sum_j A_j} \right)^{\frac{1}{2}}$$

The construction produces a consistent definition of connectivity, which is rooted not only in the topology of the network, but also in a fundamental understanding of spatial interactions among the nodes. Under this definition, connectivity is a non-dimensional number between zero and one. An economy's connectivity depends not only on its neighbors, but also on all of the interactions among the other economies in the network (just as multilateral resistance depends on trade costs across all potential trading partners).

Note:

To implement the model empirically, we need estimates of the potentials (the A and B terms in equation). They can be consistently obtained by using origin and destination fixed effects, as in much of the trade literature.

Jean-François Arvis and Ben Shepherd estimate an ***Air Connectivity Index*** (ACI) for 2007. They derive the potential terms (A and B) as fixed effects using a Poisson estimator. In the context of air transport, it posits that the bilateral impedance is a function of distance, or equivalently time of flight, with a modification of the shifted logarithmic function:

$$\tilde{X}_{ij} = A_i B_j \exp\left(-\beta f(d_{ij})\right)$$

where the dependence of distance is given by a shifted log:

$$f(d) = a \times \left[\log(a+d) - \log(a)\right]$$

In this expression, the constant a represents the "natural scale" of the network. The intuitive interpretation of this scale is that there is a fixed minimum cost in the interaction between nodes, or in the context of air transport the time to take off and land when moving from terminal to terminal.

The primary question for implementing the shifted log estimator in distance is the choice of scale parameter a. To do this, they adopt a grid search approach and run a series of fixed effects Poisson regressions. The value of the log-likelihood is maximized at approximately $a = 3900$.

3. Introduce more candidate indicators

For instance, physical connectivity relates to infrastructure assets such as power generation, transmission, and distribution systems, transportation or communications networks, and water and sanitation systems. However, the hard or visible component must be combined with some soft infrastructures, which include policies and regulations. By some definition, social, organization or regulatory institutions can also be considered as infrastructure (often referred to as “soft” infrastructure to differentiate it from “hard” physical infrastructure). Connectivity indices as a systemic framework consist of hard and soft components. The soft infrastructure must also support the hard infrastructure to ensure that the right mix and synergy of the two can provide completely connectivity index (even in sub-level indices).

Note:

- The energy index system can be expanded and refined by adding data of power supply grid, oil and gas pipeline. The port index can be included for evaluating the connectivity in the physical connectivity pillar.
- The border management and the degree of integration of the customs, such as the degree of mutual recognition of two economy’s customs documents, proof materials etc., can be used for measuring institutional connectivity.
- The education and training indicators, such as the exchange of the number of foreign students, can be included in the people-to-people connectivity pillar.