

# A New Connectivity Index for Container Ports



**GEORGIA TECH**  
**Logistics Innovation & Research Center**  
*Panama City, Panama*  
A Unit of the Supply Chain & Logistics Institute

*August 2014*

# A New Connectivity Index for Container Ports

John J. BARTHOLDI III\*, Pisit JARUMANEEROJ, and Amar RAMUDHIN  
The Supply Chain & Logistics Institute  
Georgia Institute of Technology

June 2011; revised August 20, 2014  
(submitted for publication)

## Abstract

We propose a new index, the Container Port Connectivity Index (CPCI), to measure the trade connectivity of ports within the network of container shipping. This index is based on both economics and network topology, and a distinctive feature is that the strength of a port is based on its position within the global structure of the shipping network and not just on local information such as the number of TEUs handled or direct links to other ports. Furthermore, it produces separate scores for inbound and outbound container movement and so supports more detailed analyses. We explore the usefulness of this index by analyzing the global network of scheduled mainline container-shipping services as it existed in September 2011.

The interested reader may wish to explore the global network of container shipping in more detail at <http://www2.isye.gatech.edu/~jjb/wh/apps/transportation/container-network.html>.

Keywords: *container port, shipping network, connectivity*

---

\*Corresponding author: [john.bartholdi@gatech.edu](mailto:john.bartholdi@gatech.edu), 765 Ferst Dr NW, Atlanta, GA 30332-0205 USA; phone:+1 404 894 3036

What makes a container port attractive? From an operational point of view, one important factor is *connectivity*: Are there convenient services to and from other important ports? Several measures borrowed from graph theory have been applied to measure connectivity, but these are typically based only weakly on economics. We suggest a new measure of importance with which to compare container ports, which we call the *Container Port Connectivity Index* (CPCI). This measure is based on a richer model than used heretofore of the intensity of container trade between pairs of ports. Then we use this model to compute the importance of each port as if ranking web pages. In this computation, the importance of a port is based not just on the importance of immediate neighbors but also on the importance of neighbors of those neighbors, and so on. We use these tools to analyze a model of the global network of mainline, scheduled, container-shipping services and argue that they offer a more nuanced and accurate reflection of the relative importance of ports.

## 1 The global network of container-shipping

There are several network-based models of the world-wide movement of containers by sea. The one closest to ours is that of [5]. Both their model and ours represent each container port as a vertex. Kaluza et al. include a link (edge) directed from vertex (port)  $i$  to vertex (port)  $j$  if some container ship traveled directly from port  $i$  to port  $j$  *at any time* during 2007, as reported by [www.sea-web.com](http://www.sea-web.com). In our model the meaning of a link is slightly different: There is a link directed from port  $i$  to port  $j$  if there was mainline, scheduled container service traveling directly from port  $i$  to port  $j$ , as reported by commercial data source Compair Data on September 2011. In other words our model is a snapshot of the network as it would be engaged by a shipper, while theirs includes more ephemeral phenomena, such as seasonal feeder services. This makes sense for the problem in which Kaluza et al. is interested, which is bio-invasion by species that are spread by ships. In contrast, our concern is operational: what is the nature of the network on which a particular container might move?

Ducruet and Notteboom [3] also studied network models of container shipping that are constructed, like that of Kaluza et al., from the time-aggregated movements of all container ships over a year. In one of their models two ports are connected if a container ship has traveled directly from one to the other any time during the reference year. In their other model, two ports are connected if both appeared anywhere on the same scheduled container service during that year (so that each service is represented by a complete subgraph).

Kaluza et al. perform topological analyses of their time-aggregated shipping network. Ducruet and Notteboom go further in relating the topological analyses to geography and economics, but they model container flows by *undirected* links and so ignore the direction of container movement.

Figure 1 shows the network described by our dataset. Several large patterns are im-

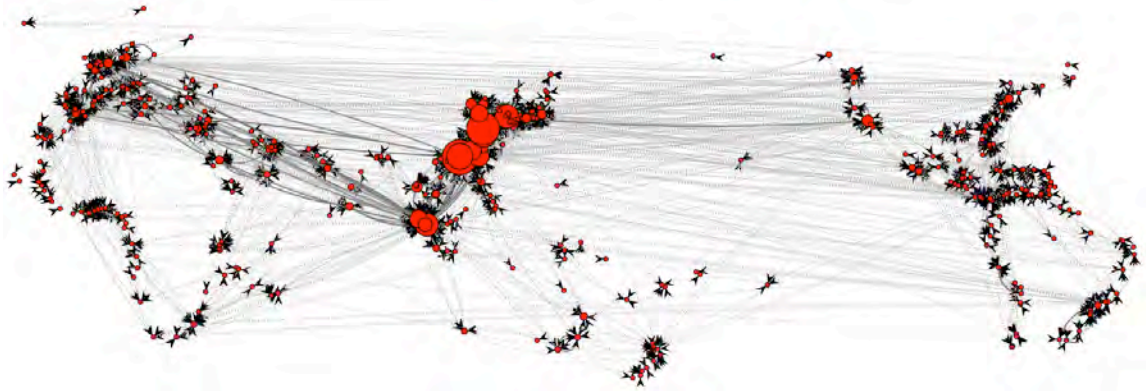


Figure 1: The network of scheduled container services among 457 ports of the world. Each arrow indicates scheduled container service from origin to destination port (but not the actual geography of the shipping route). Darker links are of greater trade intensity according to a computation based on the Liner Shipping Connectivity Index. Ports represented by larger disks scored proportionally higher according to the new measure of port connectivity described herein.

mediately evident, including the importance of the East Asian ports and the intensity of trade between East Asia and Europe, through the great transshipment ports of Southeast Asia and through the Suez Canal. Similarly, it is clear that services along the west coast of Africa or the east coast of South America are primarily local connections, from port to nearby port.

This network has 457 ports and 2,479 links and is strongly connected (that is, for any two ports, each is reachable from the other by some directed path). The mean degree of ports is 10.85, and link-diameter of the network is 11 links<sup>1</sup> (The mean degree is smaller in our network and the diameter larger than those of the time-aggregated networks, probably because they contain additional links such as seasonal and unplanned changes to shipping routes.) Because our dataset includes transit times, we can also report that the travel-time diameter of our network is 56 days (ignoring time in port)<sup>2</sup>.

The mean link-shortest path in our network is 4.02 links, with a median of 4.0. We also find the mean time-shortest path to be 18.6 days, with a median of 19 days.

<sup>1</sup>A container traveling from Maizuru (Japan) to Fortaleza (Brazil) must pass through the container ports of Niigata (Japan), Tomakomai (Japan), Hachinohe (Japan), Busan (South Korea), Savannah GA, (United States), Kingston (Jamaica), Port of Spain (Trinidad & Tobago), Degrad des Cannes (French Guiana), and Belem (Brazil).

<sup>2</sup>To ship from Honiara, Solomon Islands to Sortland, Norway requires 56 days and traverses 9 links. Any container must pass, en route, through Shanghai, Busan (South Korea), Cristobal (Panama), Manzanillo (Panama), New York, Halifax, Argentia (Newfoundland), and Reykjavik.

## 2 A link-weight based on economics

It is natural to enrich a network model to reflect the intensity of trade moving along each link. However, it is hard to get trade data at the level of containers and ports and so trade intensity is generally approximated by transport capacity. Kaluza et al. defined the *weight* of a link to be the sum of gross tonnage of all shipping traversing that link in 2007. Since we have more information regarding each link, we suggest defining the weight of a link by adapting the *Liner Shipping Connectivity Index* (LSCI). The LSCI was developed by the United Nations Conference on Trade and Development (UNCTAD) to compare the trade competitiveness of countries with respect to logistics and transport. UNCTAD computes the LSCI for a country as an aggregation of five statistics: number of liner services calling, number of liner companies, number of ships, combined container capacity of the ships (in TEUs), and capacity of the largest ship calling [9]. Despite the narrowness of focus and somewhat arbitrary method of aggregating component statistics, the LSCI is based on hard numbers and is felt to accurately reflect trade competitiveness. Indeed, the LSCI has been observed to be strongly correlated with the *Logistics Performance Index* (LPI), a comprehensive survey of perceptions that is reported annually by the World Bank [1, 9].

The LSCI implicitly treats each country of concern as if it were a single location and the entire rest of the world a single trading partner. In effect the world container network is reduced to two vertices, as in Figure 2. The five statistics on which the LSCI is based describe the container capacity connecting the country to the rest of the world and so we may interpret the LSCI as a measure of the strength of the link between two vertices.

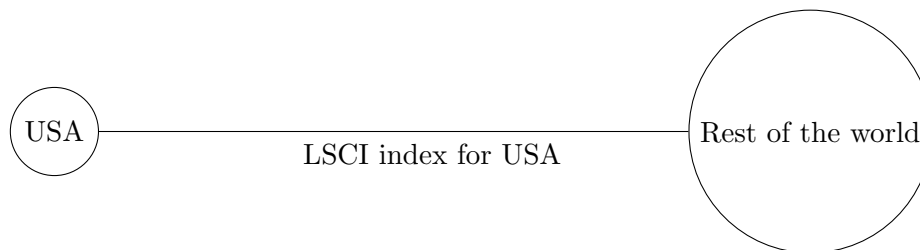


Figure 2: The Liner Shipping Connectivity Index is based implicitly on a model that aggregates all the ports of that country into one vertex and all the rest of the world into a single trading partner. The LSCI then describes the container-shipping capacity between them.

We follow the idea of the LSCI to compute, for each pair of ports  $i$  and  $j$ , a weight reflecting the intensity of container capacity moving from  $i$  directly to  $j$ . The computation is exactly that of the LSCI, except for ports rather than for countries, and for directed transit (that is *from* port  $i$  *to* port  $j$ ). Figure 3 shows the resultant distribution of weights for all direct links in our network, and Table 1 lists the twenty links of greatest weight. The most distinctive pattern is that all but one of these links are intra-Asian. Notably,

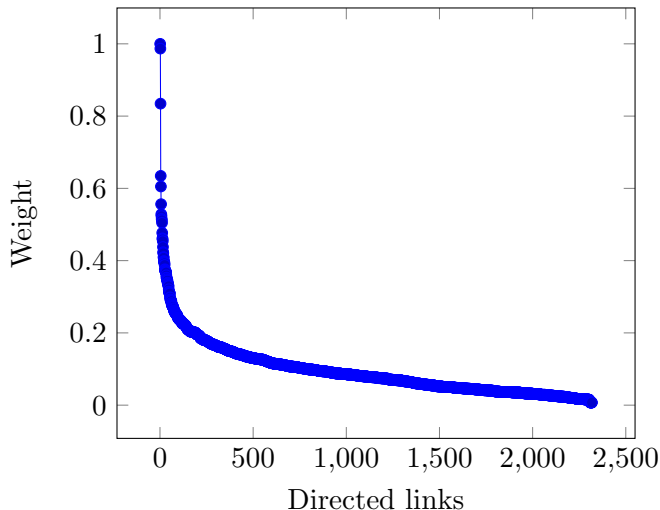


Figure 3: Distribution of LSCI weights of directed links between ports in SC network

Shanghai figures in six of these links, three times as an origin and three times as a destination. Hong Kong appears seven times and always as a destination, reflecting its role as a marshaling point for exports. This dominance of East Asian container flows is consistent with the statistics reported by Global Insight, quoted in [4], which also observes “Particularly striking is the fact that in 2010, the volume of trade in the Intra-Asia market is four times higher than the volume of trade in the transatlantic”.

## 2.1 Clustering and communities

A *community* within a network is a collection of vertices with dense and strong connections among themselves but sparser and weaker connections to other vertices. [2] identified communities among countries trading several important commodities. Here we take a more granular look and identify natural trading communities among container ports, as revealed by LSCI-weighted links.

To recognize communities, we rely on an objective function termed *modularity*. The idea is that the modularity  $Q$  of a group of communities  $\{c_i\}$  is large when there is more total weight contributed by edges within the communities than might be expected by chance [8]. More formally,

$$Q = \frac{1}{m} \sum_{ij} \left( A_{ij} - \frac{(\sum_k A_{ik})(\sum_k A_{kj})}{m} \right) \delta_{c(i),c(j)}$$

where  $A_{ij}$  has value  $w_{ij}$  if there is a link of weight  $w_{ij}$  from vertex (port)  $i$  to vertex (port)

Table 1: The twenty links of greatest weight, as determined by an LSCI-like computation. Only one is outside East Asia.

<b>From</b>	<b>To</b>	<b>Weight</b>
Shanghai	Ningbo	1.000
Ningbo	Shanghai	0.987
Hong Kong	Yantian, Shenzhen	0.834
Port Klang	Singapore	0.635
Busan	Shanghai	0.605
Singapore	Hong Kong	0.556
Yantian, Shenzhen	Hong Kong	0.528
Shanghai	Busan	0.523
Hong Kong	Shekou, Shenzhen	0.515
Singapore	Port Klang	0.519
Qingdao	Shanghai	0.505
Ningbo	Hong Kong	0.477
Shanghai	Hong Kong	0.462
Kaohsiung	Hong Kong	0.459
Rotterdam	Hamburg	0.454
Yantian, Shenzhen	Tanjung Pelepas	0.438
Chiwan, Shenzhen	Hong Kong	0.423
Shekou, Shenzhen	Hong Kong	0.421

$j, m = \sum_{ij} A_{ij}$ ,  $\delta_{ij}$  is the Kronecker delta symbol, and  $c_{(i)}$  is the index of the community to which vertex (port)  $i$  is assigned.

To identify communities in a network one must search over all partitions  $\{c_i\}$  of the vertices to find one that maximizes modularity  $Q$ . We used the heuristic search method of [7], which is known to work well, under which the SC network resolved into eight communities based on links weighted by LSCI. The results, shown in Figure 4, are suggestive. For example, the computation clearly recognized important global patterns, including trans-Pacific trade (Figure 4a), as well as trans-Atlantic (Figure 4b), and intra-American trade (Figure 4e). Other observations:

- Figure 4a: This community is the most strongly-defined in the sense that it includes the ports that contribute most to the total modularity, such as the giants Shanghai, Ningbo, and Hong Kong, and these Asian ports are the anchors of this community.

It may seem surprising that this Pacific-spanning community also includes the Caribbean port of Colon, Panama (all the other Panamanian ports are, as would be expected, in the Caribbean community of Figure 4e). But this makes sense because many services from Asia to the US East Coast find it convenient to transship at Colon for subsequent disbursement throughout the Caribbean.

- Figure 4b: Rotterdam and Hamburg are the core ports of this community.
- Figure 4c: The Mideast community is based on trade through the Suez Canal. It includes East Africa above the ports of Tanzania and the Comoros and Seychelles Islands.
- Figure 4d: The East African ports below Tanzania, including the large ports of South Africa, are better connected to the West African trading community than to others. Tanjung Pelepas is the easternmost member, reflecting its role as point of distribution of manufactured goods from East Asia to Africa. The few European members are connected primarily through the ports of Tanger or Algeciras.
- Figure 4e: The Caribbean community includes two outliers inviting comment. Wilmington, Delaware, in the US, has strong ties to Central America because of its specialization in the handling of tropical fruits and fruit juices. On the west coast, San Diego is more strongly connected to Latin America than to East Asia because the Asian services prefer to call at Los Angeles or Long Beach for their larger regional market and superior hinterland storage and transportation infrastructure.
- Figure 4f: Port-of-Spain (Trinidad) is the northernmost member. Services travel from it into this community.
- Figure 4g: This community is an artifact of the isolation of New Zealand. It consists of the regional ports of Lyttelton, Napier, Port Chalmers, and Wellington, which



have very few direct international connections. They are better connected amongst themselves than to the rest of the world. The international connections to New Zealand call mainly at Auckland and Tauranga, which are members of the Asia-Pac and trans-Pacific trading community.

- Figure 4h: This is another community determined by geography. These ports are locally connected but all significant connections to the outside world are mainly through a few ports near the Straits of Gibraltar, through which ships must pass to enter the Mediterranean Sea.

The ports that contribute most to the modularity score of a community are, in a sense, the anchors of those communities. Those of highest modularity are overwhelmingly Asian and especially Chinese, with the top ten being Shanghai, Ningbo, Hong Kong, Busan, Rotterdam, Yantian, Hamburg, Port Klang (Malaysia), and Qingdao. The ports that contribute most within the Trans-Atlantic community are Rotterdam, Hamburg, and Savannah; within the South Asia/Mideast community: Port Klang, Jeddah, and Dubai; within the West/South Africa: Tanjung Pelepas, Cape Town, and Durban; and within the southeastern US, Caribbean, and Pacific South American community: Callao (Peru), Manzanillo (Panama), and Balboa (Panama).

It is worth noting that Singapore is not among the ten largest contributors to modularity. It is a member of the powerful Asia-Pac and trans-Pacific community, but it does not have dense local connections as do the big China ports. Instead, it serves more more as a transshipment hub, with services to and from other ports that may not be directly connected themselves. This is reflected in that the *clustering coefficient* of Singapore, which measures how connected to each other are its immediate neighbors [11], is the very *lowest* among all container ports, followed by other important transshipment hubs such as Port Klang, Algeiras, Kingston, and Cartagena. These ports send and receive containers to many other ports, but their immediate neighbors do not ship much directly to each other.

### 3 A new index of strength for container ports

We have defined the weight of a link to be the value of its LSCI; now we use these weights to compute a new index of port connectivity. We compute the *Container Port Connectivity Index* (CPCI) by the “HITS” algorithm (“Hyperlink-Induced Topic Search”), which is an eigenvector-based method to rank web pages [6]. (See Appendix A for details.) The HITS algorithm computes two scores for each vertex of a network of directed edges. In the context of container shipping, we refer to these as inbound and outbound scores. Roughly speaking, a port with a high inbound score has greater power to aggregate goods; and a port with a high outbound score has greater power to distribute goods.

The CPCI will differentiate between the port of Figure 5 and one with the directions of freight flow reversed. A port will be assigned a high inbound score if container capacity



Figure 4: Trading communities identified by maximizing total LSCI weight of links within groups (rather than between groups)

flows to it from ports with high outbound score, or if it is not too far downstream from a such a port. Similarly, a port will be assigned a high outbound score if container capacity flows from it to many ports with high inbound score, or if it is not too far upstream from such a port.

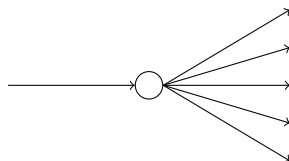


Figure 5: This port has direct connections *to* many ports, and so will receive a higher outbound score than would a port with the direction of freight flows reversed.

Praia, a container port in Cape Verde, is an example. It is unusual in the region in having a relatively high inbound score, which arises because it receives service directly from Algeciras, a regional hub with a relatively high outbound score. That service continues on to St. Vicente, which has a lower inbound score because it is further removed from Algeciras. St. Vicente, in turn, has a higher inbound score than the next few regional ports farther down-service.

The CPCI thus combines economics with network topology. Economics is reflected in the weight of the links, which are scored by an adaptation of the LSCI. And network topology is reflected in the recursive ranking of the HITS algorithm. A port scores well under the CPCI if it has strong trade connections; but it also inherits some of the importance of its neighbors, and — with diminishing effect — their neighbors, and so on.

### 3.1 Ranking ports by CPCI

As measured by the CPCI the best-connected ports are not necessarily those with the most links. For example, Cartagena receives services from 20 different ports, which is more than the 18 received by Yantian. Nevertheless, Yantian ranks much higher as scored by the CPCI with regard to inbound (0.286 versus 0.0036). This is a reflection of the fact that the CPCI depends not just on the number of links but also on the weights of those links and the scores of the ports to which they connect.

Similarly, the best-connected ports are not necessarily the busiest. Table 2 shows the CPCI scores of the twenty ports that scored highest with respect to our measure of *inbound* connectivity (where, for comparison, we have included ranking by TEUs handled in 2010). Similarly, Table 3 gives the highest ranking ports by outbound score. The ports of East Asia dominate with respect to either measure, inbound or outbound. Even though Shanghai handled more TEUs, Hong Kong ranks higher by CPCI, presumably because it is better connected within the global container-shipping network<sup>3</sup>.

<sup>3</sup>The ranking by volume combines several of the Shenzhen ports, including Yantian, Chiwan, Shekou, and

Table 2: All but two of the twenty highest-scoring scoring ports by CPCI (inbound) are in Asia. (Volume rankings are based on the number of TEUs transported through the port in 2010 [12].)

Rank	Port	<b>inbound</b>	<b>Outbound</b>	<b>Country</b>	<b>By volume</b>
1	Hong Kong	0.4080	0.4035	China	3
2	Shanghai	0.3726	0.3475	China	1
3	Ningbo	0.3040	0.3419	China	6
4	Yantian	0.2861	0.2646	China	
5	Busan	0.2515	0.2415	South Korea	5
6	Singapore	0.2456	0.3420	Singapore	2
7	Kaohsiung	0.2054	0.2009	Taiwan	12
8	Chiwan	0.1973	0.1905	China	
9	Xiamen	0.1933	0.1841	China	19
10	Shekou	0.1859	0.1614	China	
11	Port Klang	0.1748	0.1935	Malaysia	13
12	Qingdao	0.1725	0.1593	China	8
13	Nansha	0.1369	0.1109	China	
14	Tanjung Pelepas	0.1289	0.1134	Malaysia	16
15	Gwangyang	0.1289	0.1182	South Korea	
16	Los Angeles CA	0.0927	0.0286	USA	17
17	Long Beach CA	0.0917	0.0187	USA	18
18	Xingang/Tianjin	0.0890	0.0802	China	11
19	Da Chan Bay	0.0829	0.0809	China	
20	Laem Chabang	0.0818	0.0633	Thailand	22

Table 3: All but one of the twenty highest-scoring ports by CPCI (outbound) are in Asia. Oakland appears on this list because it is frequently the last US port visited before a service returns to one of the giant East Asian ports.

Rank	Port	inbound	Outbound	Country	By volume
1	Hong Kong	0.4080	0.4035	China	3
2	Shanghai	0.3726	0.3475	China	1
3	Singapore	0.2456	0.3420	Singapore	2
4	Ningbo	0.3040	0.3419	China	6
5	Yantian	0.2861	0.2646	China	
6	Busan	0.2515	0.2415	South Korea	5
7	Kaohsiung	0.2054	0.2009	Taiwan	12
8	Port Klang	0.1748	0.1935	Malaysia	13
9	Chiwan	0.1973	0.1905	China	
10	Xiamen	0.1933	0.1841	China	19
11	Shekou	0.1859	0.1614	China	
12	Qingdao	0.1725	0.1593	China	8
13	Gwangyang	0.1289	0.1182	South Korea	
14	Yokohama	0.0619	0.1147	Japan	36
15	Tanjung Pelepas	0.1289	0.1134	Malaysia	16
16	Nansha	0.1369	0.1109	China	
17	Oakland	0.0132	0.0883	USA	
18	Da Chan Bay	0.0829	0.0809	China	
19	Xingang/Tianjin	0.0890	0.0802	China	11
20	Cai Mep	0.0517	0.0750	Viet Nam	28

On the other hand our ranking appears to neglect the high-volume European ports such as Rotterdam, Antwerp, and Hamburg, as well as the busy Mideast port of Dubai, but this is because they are more isolated from other big ports. In contrast, the big East Asian ports are well-connected with the rest of the world — and with each other, which further increases their scores.

Figure 6 plots scores of all 457 ports. Several stand out for the significant differences between inbound and outbound scores, and these differences illustrate how the CPCI can make structural distinctions about the position of a port in the network.

Los Angeles and Long Beach have inbound scores that are relatively high in comparison to outbound scores. This reflects the fact that these are the two main ports of entry for product manufactured in East Asia. To reduce in-transit inventory, powerful retailers in North America insist that their freight be the last loaded out of Asia and the first unloaded in North America, and so there are many direct links from big Asian ports into Los Angeles and Long Beach. Services that have traversed the Pacific Ocean to call at Los Angeles or Long Beach then typically call at Oakland before returning to the large ports of Asia. Consequently, Oakland has a high outbound score in comparison to its inbound score. This is a general pattern that may be observed along many service loops: ports that are immediately downstream from important ports tend to have higher inbound scores, while ports toward the end of the loop tend to have higher outbound scores.

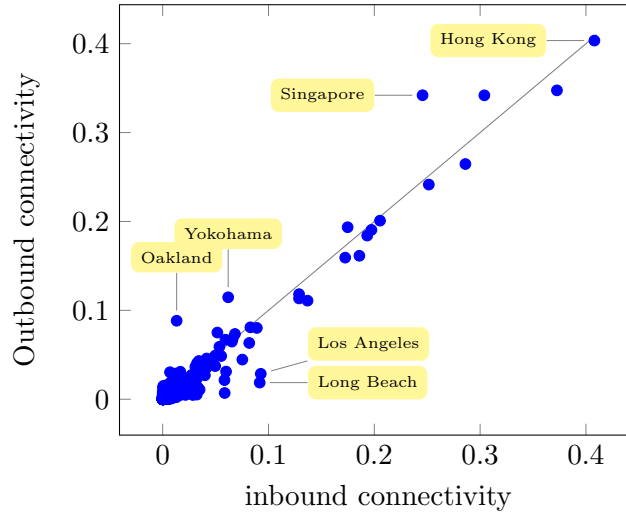


Figure 6: As computed by the Container Port Connectivity Index, those container ports for which inbound and outbound scores fall in the lower right are inbound-dominant and those in the upper left are outbound-dominant.

---

Da Chan Bay, into one, ranked fourth in volume.

### 3.1.1 North American ports

Table 4 shows that, among the ports of North America, the west coast ports, led by Los Angeles and Long Beach, dominate by the measure of inbound connectivity, reflecting the many services that come directly from the great manufacturing centers of East Asia. Moreover, many of the west coast ports score much higher with respect to inbound connectivity than to outbound.

New York is the only port on the east coast to score highly with respect to inbound scores. But Table 5 shows that east coast ports such as Savannah are more competitive with respect to outbound scores.

Table 4: All but one of the top ten North American ports by CPCI (inbound) are on the west coast.

Rank	Port	Inbound	Outbound
1	Los Angeles CA	0.0927	0.0286
2	Long Beach CA	0.0917	0.0187
3	New York NY/NJ	0.0586	0.0070
4	Lazaro Cardenas, MEX	0.0349	0.0108
5	Manzanillo, MEX	0.0325	0.0406
6	Tacoma WA	0.0305	0.0084
7	Prince Rupert BC	0.0285	0.0048
8	Vancouver BC	0.0260	0.0249
9	Seattle WA	0.0231	0.0207
10	Oakland CA	0.0132	0.0883

It will be interesting to see how these rankings change after the widening of the Panama Canal is completed in 2014.

### 3.1.2 Comparison of Container Port Connectivity Index with LSCI

The LSCI is defined for countries, while the CPCI is defined for ports. Nevertheless, we can directly compare the rankings by these indices of those countries with a single dominant port. We identified 64 container ports that were, within our data source, unique within their country, and then compared rankings by each of the 2011 LSCI and by each of the inbound and the outbound versions of Container Port Connectivity Index. The results appear in Table 6 and are generally consonant: those ranked among the top ten by LSCI are among the top twenty by CPCI, either inbound or outbound.

The differences in ranking between Gothenburg and Gdansk again illustrate how our suggested index captures structure of the network. Gdansk ranks relatively high in inbound

Table 5: Four of the top ten North American ports ranked by CPCI (outbound) are on the east coast.

Rank	Port	Inbound	Outbound
1	Oakland CA	0.0132	0.0883
2	Manzanillo, MEX	0.0325	0.0406
3	Savannah GA	0.0122	0.0289
4	Los Angeles CA	0.0927	0.0286
5	Vancouver BC	0.0260	0.0249
6	Seattle WA	0.0231	0.0207
7	Long Beach CA	0.0917	0.0187
8	Boston MA	0.0007	0.0150
9	Wilmington NC	0.0008	0.0133
10	Miami FL	0.0030	0.0125

strength because it receives shipments directly from Hamburg but ships only to the lesser port of Aarhus, which accounts for its relatively lower ranking in outbound strength. On the other hand, Gothenburg receives freight only from Aarhus, but it ships to the more significant port of Bremerhaven, from which it derives a higher outbound score.

### 3.2 Comparison with other measures of centrality

The CPCI is based on both economics and network theory and so, we believe, provides a better measure of trade-connectivity than alternative measures.

One measure of the centrality of a vertex within a network is *degree centrality*, which in our context tells from how many other ports a port receives direct shipments (in-degree) or to how many others it sends direct shipments (out-degree). While interesting, this measure neglects economic issues as the volume of trade along each link. It merely records the fact of trade.

Some measures of centrality incorporate distance. This can be useful because shipping cost is roughly proportional to distance. One such measure is *closeness*. The sum of shortest distances from a vertex to all other vertices is its *farness*. The closeness of a port is the reciprocal of its farness. We can compute the closeness centrality *to* a port and also *from* a port. Again, this ignores intensity of trade. When distance from every port is equally important, the five ports of greatest inbound closeness are all in Spain; and the most important ports with regard to outbound closeness are all in and around Panama.

*Betweenness* is the number of shortest paths within the network on which the vertex (port) lies. We find the twenty ports of greatest betweenness centrality to be quite different



Table 6: Ranks among those countries represented by a single port. The LSCI and the Container Port Connectivity Index are in general agreement, especially regarding the intensity of inbound freight.

Port	by LSCI	by CPCI (inbound)	by CPCI (outbound)
Singapore	1	1	1
Rotterdam	2	2	2
Colombo	3	3	3
Malta	4	6	11
Beirut	5	4	10
Piraeus	6	7	6
Buenos Aires	7	18	26
Karachi	8	8	5
Gothenburg	9	35	20
Gdansk	10	15	57

from those identified by Kaluza et al. A systematic difference seems to be that some among the twenty ports of highest betweenness in the network of Kaluza et al. are ranked much lower in our network. We speculate that these ports had seasonal services, which added artificially to their betweenness by suggesting paths that did not actually exist. Nevertheless betweenness did produce a list that better corresponded to what most would agree were important ports. In any case, ranking ports by degree, closeness, or betweenness share the disadvantage of treating all ports as equally important and so produce results that can be misleading.

## 4 Conclusions

The Container Port Connectivity Index is a *descriptive* index. It summarizes in two numbers something about how each port is connected to others within the larger network. Importantly, the Container Port Connectivity Index expresses more than local connectivity to immediate neighbors but also neighbors-of-neighbors, and so on, with all links weighted by an estimate of the intensity of trade (currently, a specialization of the Liner Shipping Connectivity Index). Furthermore, the Container Port Connectivity Index allows inbound and outbound strengths to be studied independently, and this gives a more detailed look at the economic roles played by each port. Finally, the Container Port Connectivity Index supports what-if analysis in a way that survey-based indices cannot.

Any index of logistics performance is an attempt to summarize a complex environment.

The LSCI may be criticized for the rather arbitrary way that data is agglomerated; and the LPI for its reliance on perception rather than measurement. The Container Port Connectivity Index has weaknesses as well. In particular, because it uses an LSCI-like computation, it inherits any criticism of that. In addition, while the CPCI scores are based on connectivity, they are not based on geography, and so do not explicitly account for travel time between ports.

Nevertheless, the Container Port Connectivity Index has many useful properties. In particular, it is based on link weights that are computed just like the Liner Shipping Connectivity Index; and because the LSCI has been vetted by economists as capturing intensities of trade, our index inherits that descriptive power and exercises it at a more granular level<sup>4</sup>.

We expect the Container Port Connectivity Index to be useful in some of the same ways as the LSCI. This may include explaining how the container-shipping network changes over time or using the edge weights and port scores as explanatory variables for economic phenomena. We believe these finer-grained statistics will be easier to understand and to explain because they directly reflect immediate decisions of primary actors such as shipping companies.

It should be remarked that none of the network models discussed herein captures anything about transshipment. Even though there may be direct links from port A to port B and from port B to port C, to transport a container from A to C may require transshipment. In this case ports A and C are further apart in both time and cost than they might appear in these models. Unfortunately there is insufficient such information available to piece together a useful view; but if that information were available, it could be incorporated into a model that explicitly represents the structure of scheduled liner services, along the lines of [3], but with directed links.

---

<sup>4</sup>It is worth noting that edge weights could be plausibly generated in other ways than the LSCI-type computation suggested herein. The LSCI represents shipping *capacity* but ideally one would like to assign weights to the links in some way that reflects the actual number of TEUs transported, rather than TEU capacity. Unfortunately, data at this level of detail is not generally available. Another alternative is to assign to each edge a weight that is the reciprocal of the typical travel time so that the strength of a port would be determined by how close its neighbors are. This might make sense in any study that focused on speed of freight flow rather than volume.

## **Acknowledgements**

We thank Nadia Viljoen of CSIRO, South Africa, and Pablo Achurra of the Georgia Tech Panama Logistics Innovation & Research Center for help in gathering and preparing the data, and Stuart Sandlin of Hapag-Lloyd and H. D. Ratliff for many helpful suggestions regarding analysis.

John Bartholdi gratefully acknowledges the support of IDI through the Leaders in Logistics program at the Supply Chain & Logistics Institute at Georgia Tech.

## References

- [1] J.-F. Arvis, M.A. Mustra, J. Panzer, L. Ojala, and T. Naula. Connecting to compete: Trade logistics in the global economy. Technical report, The World Bank, 2007.
- [2] M. Barigozzi, B. Gagliolo, and G. Mangioni. Identifying the community structure of the international-trade multi-network. *Physica A*, 390:2051–2066, 2011.
- [3] C. Ducruet and T. Notteboom. The worldwide maritime network of container shipping: spatial structure and regional dynamics. *Global Networks*, 12(3):395–423, July 2012.
- [4] Y. Hayuth. Shifting world trade patterns: The intra-asia case. Technical report, Wydra Shipping and Aviation Institute, University of Haifa, Israel, 2012.
- [5] P. Kaluza, A. Kölzsch, M.T. Gastner, and B. Blasius. The complex network of global cargo ship movement. *J. R. Soc. Interface*, 7:1093–1103, 2010.
- [6] J. M. Kleinberg. Authoritative sources in a hyperlinked environment. *Journal of the Association for Computing Machinery*, 46(5):604–632, 1999.
- [7] M. E. J. Newman and M. Girvan. Finding and evaluating community structures in networks. *Physical Review E*, 69, 2004.
- [8] M.E.J. Newman. Analysis of weighted networks. *Physical Review E*, 70(056131), 2004.
- [9] L. Ojala and J. Hoffman. A comparison of the LPI and the LSCI, 2010.
- [10] N. Perra and S. Fortunato. Spectral centrality measures in complex networks. *Physical Review E*, 78, 2008.
- [11] D. J. Watts and S. Strogatz. Collective dynamics of 'small-world' networks. *Nature*, 393(6684):440–442, June 1998.
- [12] World Shipping Council. Top 50 world container ports, 2012.

## A The HITS algorithm

The HITS algorithm was originally developed to rank the web pages for a search engine [6]. It computes two scores for each web page, a *hub* score and an *authority* score, where a good authority page is a page with many incoming links, while a good hub page is a page with many outgoing links. The idea of the HITS algorithm is that any page that is cited by important hub pages should be considered an authority. Similarly, any page that cites important authority pages should be considered a hub.

In the context of container shipping, we interpret an authority as a port that receives shipments from many ports and so is good at aggregating shipments, and so worthy of a high CPCI score for inbound. Similarly, a hub is for us a port that sends shipments to many other ports, and so is good at distribution, which results in a high CPCI score for outbound.

More generally, the HITS algorithm can be exercised on any directed network. Let  $E$  be the set of directed edges of a network and  $\lambda$  a constant. Then the authority and hub scores of vertex  $i$  are the solutions  $x_i$  and  $y_i$  to

$$\lambda x_i = \sum_{j:e_{ji} \in E} y_j \quad (1)$$

$$\lambda y_i = \sum_{j:e_{ij} \in E} x_j \quad (2)$$

If  $A$  is the adjacency matrix of the network, the equations for vertices  $i = 1, \dots, n$  can be written in matrix form as

$$\lambda x = A^T y \Rightarrow \lambda^2 x = A^T A x \quad (3)$$

$$\lambda y = A x \Rightarrow \lambda^2 y = A A^T y \quad (4)$$

Each of the above systems of equations is equivalent to the problem of finding eigenpairs satisfying constraints defined by the system of equations itself, and the importance scores are the principal eigenvectors corresponding to each of the system of equations. Such measures of centrality are known as *spectral centrality measures* [10].