Physica A 388 (2009) 2061–2071

Contents lists available at ScienceDirect

Physica A

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

Empirical analysis of the worldwide maritime transportation network

Yihong Hu [a], Daoli Zhu [a,b,*]

[a] School of Management, Fudan University, Shanghai 200433, China
[b] Shanghai Logistics Institute, Shanghai 200433, China

ARTICLE INFO

Article history:
Received 12 June 2008
Received in revised form 14 September 2008
Available online 16 December 2008

Keywords:
Scaling law
Transportation network
Complex system

ABSTRACT

In this paper we present an empirical study of the worldwide maritime transportation network (WMN) in which the nodes are ports and links are container liners connecting the ports. Using the different representations of network topology — the spaces L and P, we study the statistical properties of WMN including degree distribution, degree correlations, weight distribution, strength distribution, average shortest path length, line length distribution and centrality measures. We find that WMN is a small-world network with power law behavior. Important nodes are identified based on different centrality measures. Through analyzing weighted clustering coefficient and weighted average nearest neighbors degree, we reveal the hierarchy structure and rich-club phenomenon in the network.

1. Introduction

The recent few years have witnessed a great devotion to exploration and understanding of underlying mechanism of complex systems [1–5] as diverse as the Internet [6,7], social networks [8] and biological networks [9]. As critical infrastructure, transportation networks are widely studied. Examples include airline [10–15], ship [16], bus [17–20], subway [21] and railway [22,23] networks.

Maritime transportation plays an important role in the world merchandise trade and economics development. Most of the large volume cargo between countries like crude oil, iron ore, grain, and lumber are carried by ocean vessels. According to the statistics from United Nations [24], the international seaborne trade continuously increased to 7.4 billion tons in 2006 with a robust annual growth rate of 4.3%. And over 70% of the value of world international seaborne trade is being moved in containers.

Container liners have become the primary transportation mode in maritime transport since 1950's. Liner shipping means the container vessels travel along regular routes with fixed rates according to regular schedules. At present most of the shipping companies adopt hub-and-spoke operating structure which consists of hub ports, lateral ports, main lines and branch lines, forming a complex container transportation network system [25].

Compared with other transportation networks, the maritime container liner networks have some distinct features: (1) A great number of the routes of container liners are circular. Container ships call at a series of ports and return to the origin port without revisiting each intermediate port. It's called pendulum service in container transportation. While bus transport networks and railway networks are at the opposite with most of buses or trains running bidirectionally on routes. (2) The network is directed and asymmetric due to circular routes. (3) Lines are divided into main lines and branch lines. Main lines are long haul lines which involves a set of sequential port calls across the oceans. Sometimes long haul lines call at almost 30 ports. Branch lines are short haul lines connecting several ports in one region to serve for main lines.

* Corresponding author at: School of Management, Fudan University, Shanghai 200433, China. Tel.: +86 2165643072.
E-mail address: dlzhu@fudan.edu.cn (D. Zhu).

0378-4371/$ – see front matter © 2008 Elsevier B.V. All rights reserved.
doi:10.1016/j.physa.2008.12.016

Table 1
Number of sea ports by major geographic region.

Region	No of sea ports
Africa	96
Asia and Middle East	251
Europe	311
North America	61
Latin America	96
Oceania	63
Total	878

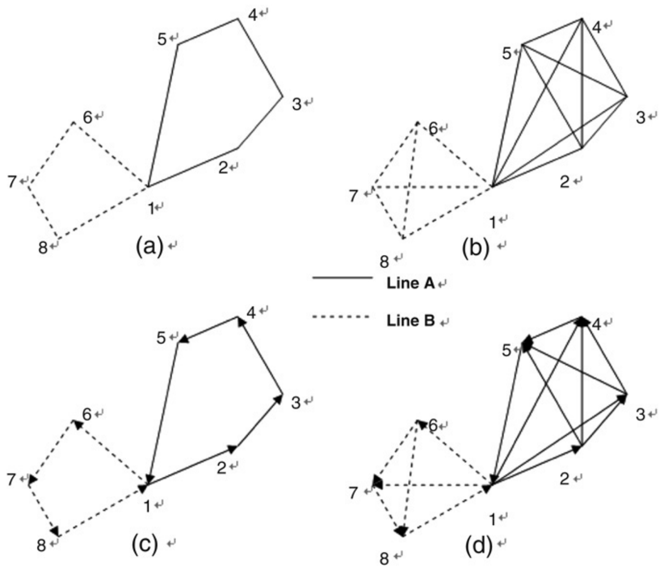

Fig. 1. Description of the space L and the space P. (a) and (b) are the undirected representations in the space L and the space P, respectively. (c) and (d) are the directed representations in the space L and the space P, respectively. In the space L, a link is created between consecutive stops in one route. In the space P all ports that belong to the same route are connected. Line A (solid line) and line B (dashed line) are two different pendulum routes sharing one common node: the port No. 1.

We construct the worldwide maritime transportation network (WMN) using two different network representations, namely the spaces L and P, and analyze basic topological properties. Our result shows that the degree distribution follows a truncated power-law distribution in the space L and an exponential decay distribution in the space P. With small average shortest path length 2.66 and high clustering coefficient 0.7 in the space P, we claim that WMN is a small world network. We also check the weighted network and find the network has hierarchy structure and "rich-club" phenomenon. Centrality measures are found to have strong correlations with each other.

The rest of the paper is organized as follows: in Section 2, we introduce the database and set up the network using two different network representations. In Section 3 various topological properties are studied including degree distribution, degree correlations, shortest path length, weight distribution and strength distribution, etc. Section 4 discloses the hierarchy structure by studying the weighted and unweighted clustering and degree correlations. Centrality measures correlations and central nodes' geographical distribution are studied in Section 5. Section 6 gives the conclusion.

2. Construction of the network

We get the original data from an authoritative container industry database named CI-online [26] which provides up-to-date statistics of all the 878 sea ports in the world and 1802 container lines served by 434 ship companies covering all the big ship companies such as Maersk, MSC, Evergreen, Cosco, etc. The ports are distributed in different regions and we list the number of ports in each region in Table 1.

To construct the worldwide maritime transportation network from the above data, we have to introduce the concept of spaces L and P as presented in Fig. 1. The idea of spaces L and P is first proposed in a general form in [22] and later widely used in the study of public bus transportation networks and railway networks. The space L consists of nodes being ports and links created between consecutive stops in one route. Degree k in the space L represents the number of directions passengers or cargoes can travel at a given port. The shortest path length in the space L is the number of stops one has to pass to travel between any two ports. In the space P, two arbitrary ports are connected if there is a container line traveling between both ports. Therefore, degree k in the space P is the number of nodes which can be reached without changing the line. The shortest path length between any two nodes in the space P represents the transfer number plus one from one node to another and thus is shorter than that in the space L.

Since WMN is a directed network, we extend the concept of spaces L and P to directed networks according to [16]. See Fig. 1. Line A and B are two different pendulum routes crossing at the port No. 1. (a) and (b) is the undirected network representation. (c) and (d) is the respective directed version.

Based on the above concepts we establish the network under two spaces represented by asymmetrical adjacent matrices A^L, A^P and weight matrices W^L, W^P. The element a_{ij} of the adjacent matrix A equals to 1 if there is a link from node i to j or 0 otherwise. The element w_{ij} of weight matrix W is the number of container lines traveling from port i to port j.

We need to define the quantities used in this weighed and directed network. We employ $k^L_{in}(i)$ and $k^L_{out}(i)$ to denote in-degree, out-degree of node i in the space L, and $k^L_{un}(i)$ to represent undirected degree in the space L. Similarly $k^P_{in}(i)$, $k^P_{out}(i)$ and $k^P_{un}(i)$ are employed in the space P. Hence we have

$$k^L_{in}(i) = \sum_{j \neq i} a^L_{ji} \tag{1}$$

$$k^L_{out}(i) = \sum_{j \neq i} a^L_{ij} \tag{2}$$

$$k^L_{un}(i) = \sum_{j \neq i} (a^L_{ij} + a^L_{ji}) \tag{3}$$

which also holds for the space P.

Strength is defined as the total weight of vertex connections [11]. It is also divided into in-strength and out-strength. In the space L the in-strength of node i is denoted by $S^L_{in}(i)$ and out-strength is denoted by $S^L_{out}(i)$. Undirected strength (total strength) is represented by $S^L_{un}(i)$. They can be calculated according to the following equations:

$$S^L_{in}(i) = \sum_{j \neq i} w^L_{ji} \tag{4}$$

$$S^L_{out}(i) = \sum_{j \neq i} w^L_{ij} \tag{5}$$

$$S^L_{un}(i) = \sum_{j \neq i} (w^L_{ij} + w^L_{ji}) \tag{6}$$

which also holds for $S^P_{in}(i)$, $S^P_{out}(i)$, $S^P_{un}(i)$ in the space P.

Other quantities like clustering coefficient and average nearest neighbors degree also have different versions in directed and weighted WMN. We employ c^L_i and c^P_i to denote the unweighted clustering coefficient of node i in the space L and P, respectively. Analogously, $k^L_{nn,i}$ and $k^P_{nn,i}$ are used to denote the average nearest neighbors degree of node i in the space L and P, respectively. For weighted WMN we add superscript W to the above quantities and consequently they become c^{WL}_i, c^{WP}_i, $k^{WL}_{nn,i}$ and $k^{WP}_{nn,i}$.

3. Topological properties

3.1. Degree distribution and degree correlations

First we examine the degree distributions in two spaces. Fig. 2 shows that in-degree, out-degree and undirected degree distributions in the space L all follow truncated power-law distributions with nearly the same exponents. In-degree and out-degree obey the function $P(k) \sim k^{-1.7}$ before $k = 20$. When $k > 20$ their distribution curves bend down to the function $P(k) \sim k^{-2.95}$. Unweighted degree in the space L has the same exponents of -1.7 and -2.95 but the critical point becomes $k = 30$. Truncated power-law degree distributions are often observed in other transportation networks like the worldwide air transportation network [12], China airport network [13], US airport network [14] and the Italian airport network [15]. It is explained in Ref. [10] that the connection cost prevents adding new links to large degree nodes. Analogous cost constraints also exist in the maritime transport network. Congestion in hub ports often makes ships wait outside for available berth for several days, which can cost ships extremely high expense. Consequently new links are not encouraged to connect to those busy ports.

While in the space P three degree distributions all follow exponential distributions $P(k) \sim e^{-\alpha k}$. The parameters are estimated to be $\alpha = 0.0117$ for in-degree, and $\alpha = 0.0085$ for out-degree, $\alpha = 0.0086$ for unweighted degree. The property that degrees obey truncated power-law distributions in the space L and exponential distributions in the space P is identical to public transportation networks [17,18] and railway networks [22]. Particularly the Indian railway network [22] has exponential degree distributions with the parameter 0.0085 almost the same with in-degree and out-degree distribution in WMN.

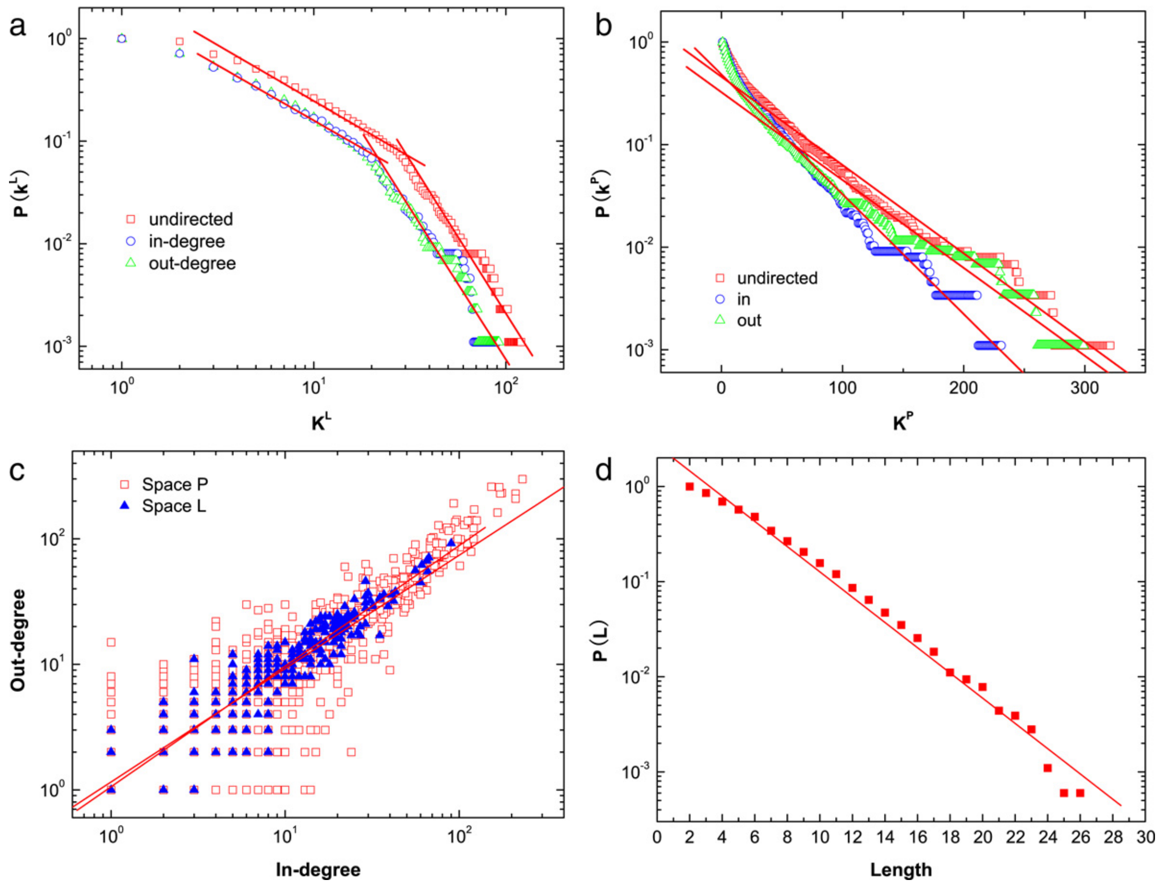

Fig. 2. (a) Cumulative degree distributions of degree in the space L obey truncated power-law distributions with almost the same exponents. The turning points are at $k = 20$ and $k = 30$, respectively. (b) Cumulative degree distributions of degree in the space P all follow exponential distributions. (c) Positive correlations between in-degree and out-degree. In two spaces they have nonlinear relations: $k_{out}^{L} \sim (k_{in}^{L})^{0.96}$ and $k_{out}^{P} \sim (k_{in}^{P})^{0.90}$. (d) Cumulative probability distribution of line length. It can be approximated by a straight line in semi-log plot indicating an exponential decay estimated to be $P(l) \sim \mathrm{e}^{-0.13}$.

Next, the relation between in-degree and out-degree is studied. Fig. 2(c) is a plot of out-degree k_{out} vs. in-degree k_{in}. They have positive and nonlinear correlations under two spaces and are estimated to be: $k_{out}^{L} \sim (k_{in}^{L})^{0.96}$ and $k_{out}^{P} \sim (k_{in}^{P})^{0.90}$. Evidently the in-out degree correlation is very strong.

3.2. Line length

Let's denote line length, i. e. the number of stops in one line, as l. In Fig. 2(d) the probability distribution of line length $P(l)$ can be approximated as a straight line in the semi-log picture representing an exponential decay distribution $P(l) \sim \mathrm{e}^{-\alpha l}$ with the parameter $\alpha = 0.13$. It indicates there are much more short haul lines than long haul lines in maritime transportation. Long haul lines use large vessels and travel long distance from one region to another region while short haul lines as branch lines travel between several neighboring ports and provide cargo to main lines. For example, the line consisting of the following ports: Shanghai–Busan–Osaka–Nagoya–Tokyo–Shimizu–Los Angeles–Charleston–Norfolk–New York–Antwerp–Bremerhaven–Thamesport–Rotterdam–Le Havre–New York–Norfolk–Charleston–Colon–Los Angeles–Oakland–Tokyo–Osaka–Shanghai, is a typical long haul line connecting main ports in Asia and Europe, calling at ports for 24 times.

3.3. Shortest path length

The frequency distributions of shortest path lengths d in the spaces L and P are plotted in Fig. 3. The distribution in the space L has a wider range than in the space P. The average shortest path length is 3.6 in the space L and 2.66 in the space P (see Table 2). This means generally in the whole world the cargo need to transfer for no more than 2 times to get to the destination. Compared with the network size $N = 878$, the shortest path length is relatively small.

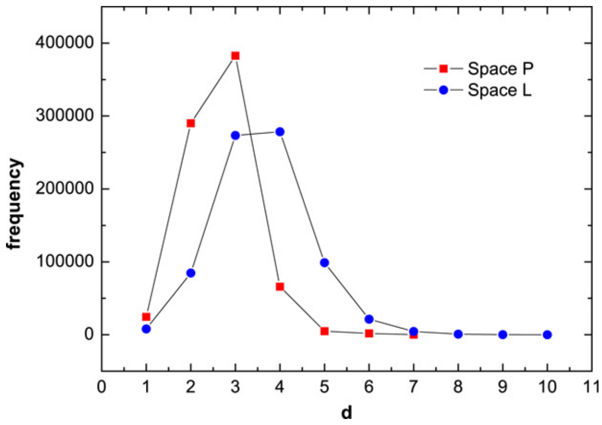

Fig. 3. Frequency distributions of shortest path length under two spaces. The distribution in the space L has a wider range than in the space P.

Table 2
Basic parameters for spaces L and P. n is the number of nodes and m is the number of links. $\langle k\rangle$ is the average undirected degree. $\langle C\rangle$ is the average unweighted clustering coefficient. $\langle l\rangle$ is the average shortest path length.

Space	n	m	$\langle k_{un}\rangle$	$\langle C\rangle$	$\langle l\rangle$
Space L	878	7955	9.04	0.4002	3.60
Space P	878	24967	28.44	0.7061	2.66

3.4. Weight and strength distribution

Usually traffic on the transportation network is not equally distributed. Some links have more traffic flow than others and therefore play a more important role in the functioning of the whole network. Weight should be addressed especially in transportation networks. Here we study four properties of weighted WMN: weight distribution, strength distribution, in-out strength relations and the relations between strength and degree. The results are displayed in Fig. 4.

First we examine weight distributions. In Fig. 4(a) two weight distribution curves are approximately straight declining lines before $w = 40$. The power-law distributions are estimated to be $P(w) \sim w^{-0.95}$ in the space P and $P(w) \sim w^{-0.92}$ in the space L.

Next, Fig. 4(b) shows the undirected strength distributions under two spaces both obey power-law behavior with the same parameter. The functions are estimated to be $P(s) \sim s^{-1.3}$.

And we also analyze the relations between in-strength and out-strength in two spaces. As we can see from Fig. 4(c), in-out strength relations under the spaces L and P are positively correlated. The fitted straight line is estimated to be $y = \alpha x$ with $\alpha = 1.03$ for space P and $\alpha = 1.00$ for space L. The linear relations between in-out strength indicate the balance of cargo traffic in and out of ports.

Finally, an important feature of weighted WMN, the relations between strength and degree, is investigated. Under two spaces the relations between undirected strength and undirected degree are both nonlinear with the slope of the line approximately 1.3, which means the strength increase quicker than the increase of degree. This often occurs in the transportation networks and has its implication in the reality. It's easy for the port with many container lines to attract more lines to connect the port and thus to increase the traffic more quickly.

4. Hierarchy structure

In this section, we explore the network structure of WMN through studying both the weighed and unweighted versions of clustering coefficient and average nearest neighbors degree. Hierarchy structure and "rich-club" phenomena are unveiled. We conjecture this kind of structure is related to ship companies' optimal behavior to minimize the transportation cost known as the hub-and-spoke model in transportation industry.

4.1. Clustering

Clustering coefficient c_i is used to measure local cohesiveness of the network in the neighborhood of the vertex. It indicates to what extent two individuals with a common friend are likely to know each other. And $C(k)$ is defined as clustering coefficient averaged over all vertices with degree k.

We plot $C(k)$ in Fig. 5(a) in log–log scale. They don't obey power-law distributions as those observed in public transport networks [17,20] or in the ship transport network [16]. Either in the space L or in the space P, $C(k)$ exhibits a highly nontrivial

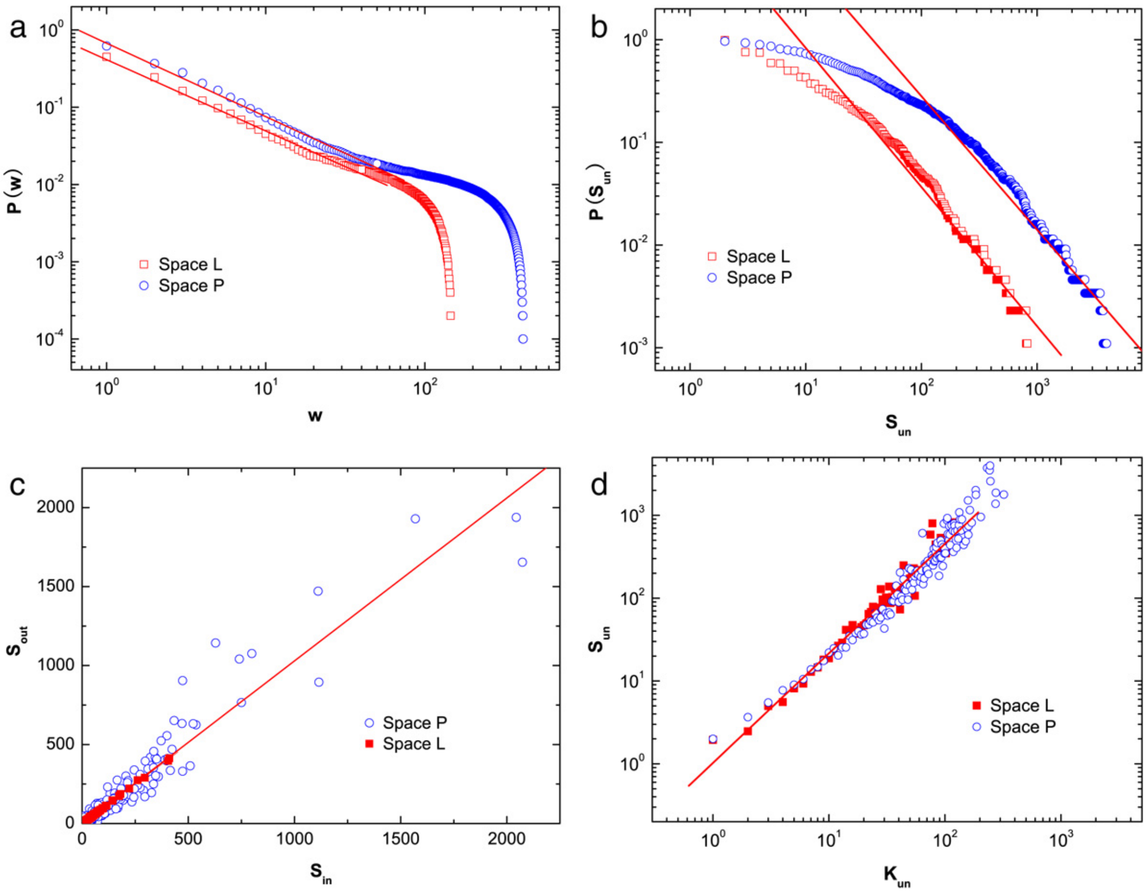

Fig. 4. (a) Cumulative probability distributions of weight in two spaces. (b) Cumulative probability distributions of undirected strength in two spaces. (c) Linear relation between in-strength and out-strength in two spaces. (d) Correlations between strength and degree in two spaces. The slopes both equal to 1.3 approximately indicating nonlinear relationships in the spaces *L* and *P*.

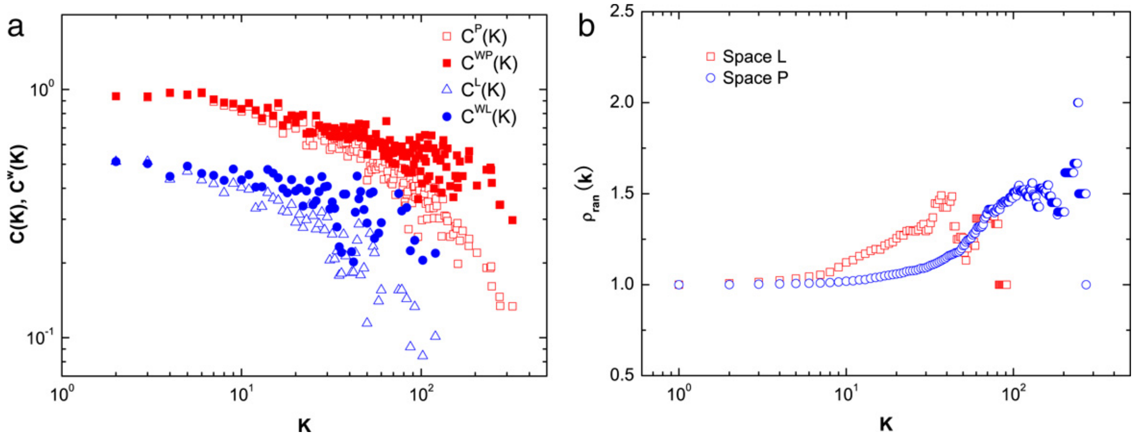

Fig. 5. (a) Clustering coefficients. $C(k)$ under two spaces both exhibit nontrivial behavior with decay curves as functions of degree *k*. And weighted versions of clustering coefficient are larger than unweighted clustering coefficient, indicating high traffic edges between interconnected vertices. (b) Rich-club coefficient. With $\rho_{ran}(k) > 1$ the worldwide maritime transportation network shows a rich-club ordering under both the space *L* and *P*.

behavior with a decay curve as a function of degree *k*, signaling a hierarchy structure in which low degrees belong generally to well interconnected communities (high clustering coefficient), while hubs connect many vertices that are not directly connected (small clustering coefficient).

$C^P(k)$ lies above $C^L(k)$ and the average clustering coefficient of the network in the space P is 0.7 larger than 0.4 in the space L. This can be explained by the fact that in the space P each route gives rise to a fully connected subgraph. With high clustering coefficient 0.7 and small average shortest path length 2.66 in the space P, we conclude that the WMN, as expected, has the small-world property.

Weighed quantities for clustering and assortativity measures are first proposed in Ref. [11]. Through the case study of WAN and SCN, [11] demonstrates that the inclusion of weight and their correlations can provide deeper understanding of the hierarchical organization of complex networks. The weighted clustering coefficient is defined as:

$$c_i^w = \frac{1}{s_i(k_i - 1)} \sum_{j,h} \frac{w_{ij} + w_{ih}}{2} a_{ij} a_{ih} a_{jh} \tag{7}$$

which takes into account the importance of the traffic or interaction intensity on the local triplets. And we define $C^w(k)$ as the weighted clustering coefficient averaged over all vertices with degree k. In real weighted network we may have two opposite cases of the relation between $C^w(k)$ and $C(k)$. If $C^w(k) > C(k)$ in the network, interconnected triplets are more likely formed by the edges with larger weights. If $C^w(k) < C(k)$ the largest interactions or traffic is occurring on edges not belonging to interconnected triplets.

In Fig. 5(a) we report weighted clustering coefficient under two spaces. Evidently the weighted clustering coefficient $C^{wP}(k)$ and $C^{wL}(k)$ are both above the corresponding unweighted clustering coefficient, i. e. $C^{wP}(k) \geq C^P(k)$, and $C^{wL}(k) \geq C^L(k)$. This indicates some closely interconnected nodes with large degrees have the edges with larger weights among themselves. In other words, high-degree ports have the tendency to form interconnected groups with high-traffic's links, thus balancing the reduced clustering. This is so called "rich-club" phenomenon.

Rich-club phenomenon is quantified by rich-club coefficient. It is first defined in Ref. [27] as the fraction of actual edges among potential edges in given vertices:

$$\phi(k) = \frac{2E_{>k}}{N_{>k}(N_{>k} - 1)} \tag{8}$$

where $N_{>k}$ refers to the number of nodes with degrees higher than a given value k and $E_{>k}$ stands for the number of edges among the $N_{>k}$ nodes. Then it is normalized to $\rho_{ran}(k) = \phi(k)/\phi_{ran}(k)$ in Ref. [28] where $\phi_{ran}(k)$ is the rich-club coefficient of the maximally random network with the same degree distribution [29]. $\rho_{ran}(k) > 1$ denotes an actual rich-club ordering in the network.

We calculate the normalized rich-club coefficient $\rho_{ran}(k)$ according to the method described in Ref. [28]. Fig. 5(b) shows $\rho_{ran}(k) > 1$ and the increasing dependence of $\rho_{ran}(k)$ on k under both the space L and P, clearly indicating rich-club phenomenon in the worldwide maritime transportation network.

4.2. Assortativity

There is another important quantity to probe the networks' architecture: the average degree of nearest neighbors, $k_{nn}(k)$, for vertices of degree k. Average nearest neighbors degree of a node i is defined as:

$$k_{nn,i} = \frac{1}{k_i} \sum_j a_{ij} k_j. \tag{9}$$

Using $k_{nn,i}$, one can calculate the average degree of the nearest neighbors of nodes with degree k, denoted as $k_{nn}(k)$. The networks are called assortative if $k_{nn}(k)$ is an increasing function of k, whereas they are referred to as disassortative when $k_{nn}(k)$ is a decreasing function of k. As suggested in Ref. [11], weighted version of average degree of nearest neighbors is calculated by:

$$k_{nn,i}^w = \frac{1}{s_i} \sum_j a_{ij} w_{ij} k_j. \tag{10}$$

From this definition we can infer that $k_{nn,i}^w > k_{nn,i}$ if the links with the larger weights are pointing to the neighbors with larger degrees and $k_{nn,i}^w < k_{nn,i}$ in the opposite case.

Both the weighted and unweighted average degree of nearest neighbors under two spaces are plotted in Fig. 6. The curve of $k_{nn}^P(k)$ lies above the curve of $k_{nn}^L(k)$. The $k_{nn}^L(k)$ and $k_{nn}^P(k)$ grow with the increase of degrees at small degrees but decline when degrees are large. The unweighted network exhibits assortative behavior in the small degree range but disassortative behavior in large range.

When we turn to the $k_{nn}^{wL}(k)$ and $k_{nn}^{wP}(k)$, the weighted analysis provides us a different picture. We can see that under two spaces the weighted average degree of nearest neighbors exhibits a pronounced assortative behavior in the whole k spectrum. Since the number of WMN nodes is 878, this conforms with the empirical finding in Ref. [17] that public transport networks are assortative when the number of nodes in the network $N > 500$ and disassortative when $N < 500$.

From the above discussion we can see that the inclusion of weight changes the behavior of clustering coefficient and average degree of nearest neighbors. This property is identical to the worldwide airport network [11] and North America

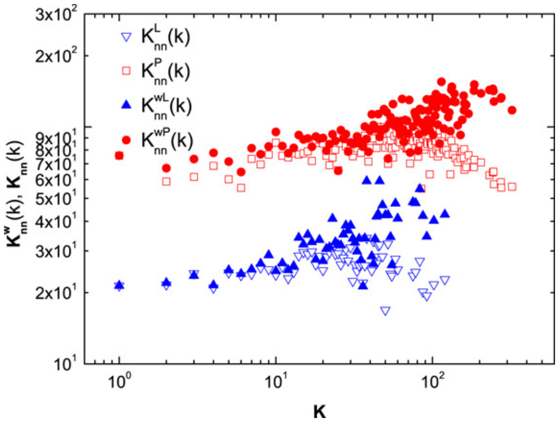

Fig. 6. Average degree of the nearest neighbors as functions of k. Inclusion of weight changes the behavior of $k_{nn}(k)$: assortative behavior in the small degree range but disassortative behavior in large degree range, to definite assortative behavior of $k^w_{nn}(k)$ in the whole k spectrum.

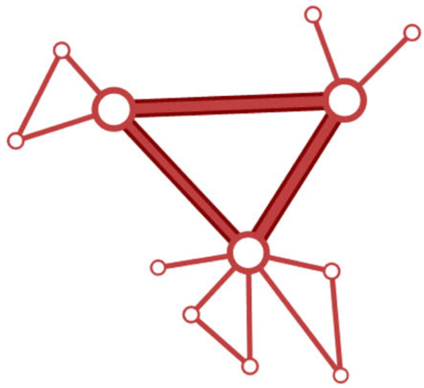

Fig. 7. A classical hub-and-spoke structure in maritime transportation. There are three central vertices having very strong links (high traffic) with each other and several nodes having weak links (low traffic) with hubs. It has the same property of $C^w(k) > C(k)$, $k^w_{nn}(k) > k_{nn}(k)$ with that of WMN.

airport network [30]. In both the airline transportation and maritime transportation networks, high traffic is associated to hubs and high-degree ports (airports) tend to form cliques with other large ports (airports). Their similar organization structure may have a similar underlying mechanisms. We conjecture that this is related to the hub-and-spoke structure which is widely adopted in practice by airline companies or ship companies to achieve the objective of minimizing the total transportation cost [31–34].

Fig. 7 describes a typical hub-and-spoke structure which consists of three interconnected hubs and other nodes allocated to a single hub. In maritime transportation main liners travel between hubs handling large traffic while branch liners visit the hub's neighboring ports to provide cargo for the main lines. This structure allows the carriers to consolidate the cargo in larger vessels to lower the transportation cost. This simple structure has the similar property of $C^w(k) > C(k)$, $k^w_{nn}(k) > k_{nn}(k)$ with that of WMN. And it also displays rich-club phenomenon. We think it worths investigating the relations between ship companies' optimal behavior and the real transportation network's hierarchy structure and rich-club property.

5. Centrality measures

In this section we analyze two centrality measures in social network analysis [35]: degree and betweenness. The most intuitive topological measure of centrality is given by the degree: more connected nodes are more important. The distribution and correlations of degree has been discussed in Section IV.

Betweenness centrality is defined as the proportion of the shortest paths between every pair of vertices that pass through the given vertex v towards all the shortest paths. It is based on the idea that a vertex is central if it lies between many other vertices, in the sense that it is traversed by many of the shortest paths connection couples of vertices. Hence we have

$$C_i^B = \frac{1}{(N-1)(N-2)} \sum_{j,k} \frac{n_{jk}(i)}{n_{jk}} \tag{11}$$

where n_{jk} is the number of shortest paths between j and k, and $n_{jk}(i)$ is the number of shortest paths between j and k that contain node i.

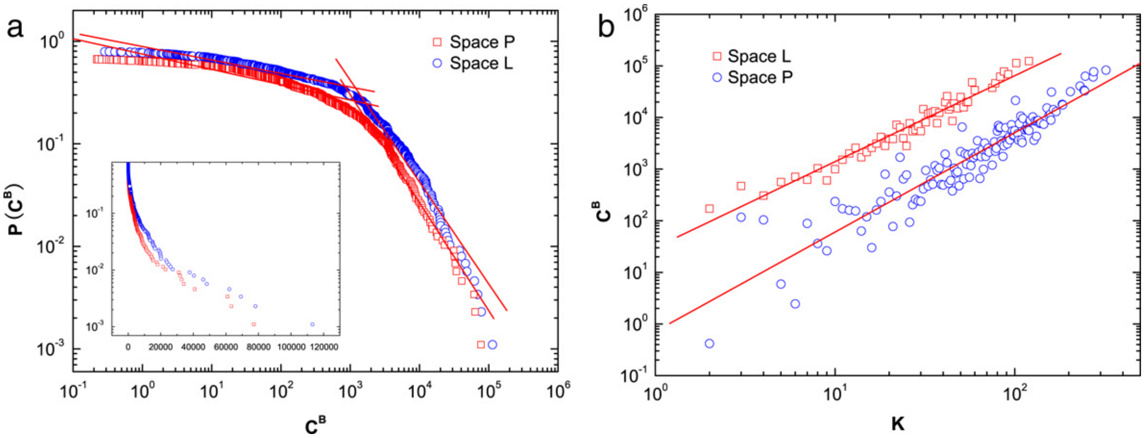

Fig. 8. (a) Cumulative probability distribution of betweenness obeys truncated power-law distribution. The inset plots betweenness distribution in the semilog scale which clearly is not an exponential decay. (b) Betweenness is a straight line in log–log picture indicating a power-law relation with degree. The exponent in the space L is estimated to be 1.66 and that in the space P is estimated to be 1.93.

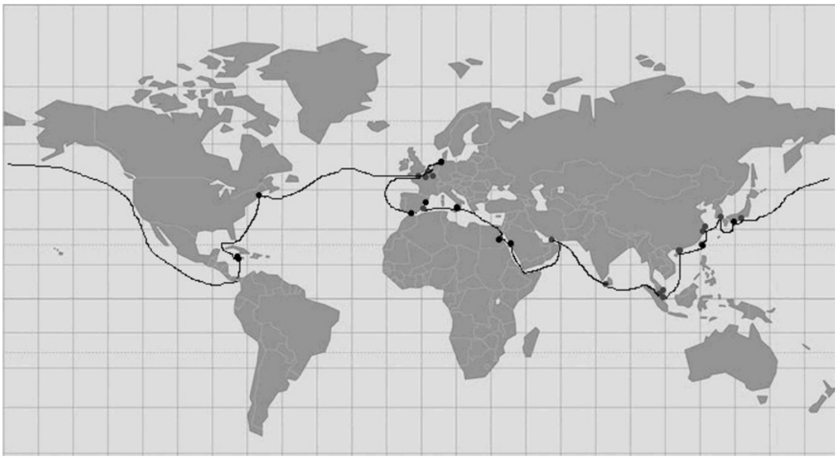

Fig. 9. The geographical distribution of 25 most connected ports. They are located along east–west lines, including 13 ports in Asia and Middle East, 9 in Europe, 1 in Africa, 1 in North America and 1 in Latin America.

Correlations between two centrality measures are presented in Fig. 8(b). In both the spaces there is a clear tendency to a power-law relation with degree k: $C^B(k) \sim k^\alpha$ with $\alpha = 1.66$ in the space L and $\alpha = 1.93$ in the space P. The power-law correlations between degree and betweenness is also found in bus transportation networks [17,20] and ship transport network [16]. It's worth noting that this power-law relations together with the truncated scale-free behavior of the degree distribution implies that betweenness distribution should follow a truncated power law. This behavior is clearly identified in Fig. 8(a). We find the betweenness centrality has two-regime power-law behavior $P(C^B) \sim C^{-\alpha}$. For the two spaces, exponents are almost the same: $\alpha = 0.14$ at small degree regime and $\alpha = 1.0$ at large degree regime.

The power-law relations between degree and betweenness suggest that they are consistent with each other. It is proved in the comparison of each port's degree and betweenness. The 25 most connected ports are listed in Table 3. Singapore is the most busy ports in the world with the largest degree and betweenness. Antwerp and Bushan are the second and third either in degree or in betweenness measures. Only 5 ports in these ports are not listed in the 25 most central ports in betweenness measure. WMN is not like the case of the worldwide airline network [12] which has anomalous centrality due to its multicommunity structure. The difference may due to the fact that there are less geographical and political constraints in maritime transportation than in air transportation. Ships can travel longer distance than airplanes. And airports are usually classified into international and domestic airports and international airlines are limited to connect international airports instead of domestic airports. So there are distinct geographically constrained communities in WAN. In the maritime transportation there are no such constraints. Sea ports basically are all international ports with the possibility to connect to any other sea ports in the world.

In Fig. 9 we plot the 25 most connected ports on the world map. They show unbalanced geographical distribution mainly located in Asia and Europe, including 13 ports in Asia and Middle East, 1 in Africa, 9 in Europe, 1 in North America and 1 in Latin America. Particularly they are located along the east–west lines. Lines in maritime transportation are usually

Table 3
The 25 most connected ports in the worldwide maritime transportation network.

Rank	Ports	Degree	Betweenness	Region
1	Singapore	120	124110.1258	Asia
2	Antwerp	102	113368.6161	Europe
3	Bushan	92	69094.7490	Asia
4	Rotterdam	87	78097.8754	Europe
5	Port Klang	83	62111.6226	Asia
6	Hongkong	78	46072.9799	Asia
7	Shanghai	75	37748.4316	Asia
8	Hamburg	60	40362.4625	Europe
9	Valencia	60	27346.0956	Europe
10	Le Havre	58	48231.1636	Europe
11	Gioia Tauro	55	20148.4667	Europe
12	Yokohama	54	19716.7287	Asia
13	Kaohsiung	52	15363.5781	Asia
14	Port Said[a]	52	16121.8476	Africa
15	Bremerhaven	50	24998.9088	Europe
16	Colombo[a]	48	15128.362	Asia
17	Tanjung Pelepas	46	18735.1267	Asia
18	Jeddah[a]	45	8570.7724	Middle East
19	Jebel Ali	44	20137.5255	Middle East
20	Ningbo[a]	44	12315.9515	Asia
21	Algeciras	43	18701.2619	Europe
22	Barcelona[a]	43	15256.3903	Europe
23	Kobe	43	8869.7313	Asia
24	New York	42	25935.6692	North America
25	Kingston	41	23281.3223	Latin America

[a] These ports are not among the 25 most central ports.

divided into east–west lines, north–south lines and south–south lines [25,36]. The fact that 25 most connected ports in the world are in east–west trade routes represents rapid growth and large trade volume in Europe–America, Asia–America and Asia–Europe trade [24].

6. Conclusion

In this paper we have presented an empirical study of the worldwide maritime transportation network (WMN) under different representations of network topology. We study the statistical properties of WMN and find that WMN is a small world network with power law behavior. There are strong correlations in degree–degree, strength–degree and betweenness-degree relations. Central nodes are identified based on different centrality measures. Based on the analysis of weighted clustering coefficient and weighted average nearest neighbors degree, we find that WMN has the same hierarchy structure and "rich-club" phenomenon with WAN. We conjecture that this structure is related to optimal behavior both existing in air transportation and maritime transportation. So our future research direction is the evolution modeling of WMN using optimal behavior to reproduce real properties in WMN.

Acknowledgments

This work is supported by Natural Science Foundation of China (NSFC 70432001) and Shanghai Leading Academic Discipline Project (B210). And the author would like to thank Ms. Zui Xu's help in collecting the original data.

References

[1] R. Albert, A.-L. Barabasi, Rev. Modern Phys. 74 (2002) 4797.
[2] M.E.J. Newman, SIAM Rev. 45 (2003) 167256.
[3] S.N. Dorogotsev, J.F.F. Mendes, Evolution of Networks: From Biological Nets to the Internet and WWW, Oxford Univ. Press, Oxford, 2003.
[4] S. Boccaletti, V. Latora, Y. Moreno, M. Chavez, D.-U. Hwang, Phys. Rep. 424 (2006) 175308.
[5] G Caldarelli, Scale-Free Networks: Complex Webs in Nature and Technology, Oxford Univ. Press, Oxford, 2007.
[6] A.-L. Barabasi, R. Albert, Science 286 (1999) 509.
[7] R. Pastor-Satorras, A. Vespignani, Evolution and Structure of the Internet: A Statistical Physics Approach, Cambridge Univ. Press, Cambridge, 2004.
[8] M.E.J. Newman, Proc. Natl. Acad. Sci. USA 98 (2001) 404.
[9] H. Jeong, B. Tombor, R. Albert, Z.N. Oltvai, A.-L. Barabasi, Nature 407 (2000) 651.
[10] L.A.N. Amaral, A. Scala, M. Barthelemy, H.E. Stanley, Proc. Natl. Acad. Sci. USA 97 (2000) 11149.
[11] A. Barrat, et al., Proc. Natl. Acad. Sci. USA 101 (11) (2004) 3747.
[12] R. Guimera, S. Mossa, A. Turtschi, L.A.N. Amaral, Proc. Natl. Acad. Sci. USA 102 (2005) 7794.
[13] W. Li, X. Cai, Phys. Rev. E 69 (2004) 046106.
[14] L.P. Chi, R. Wang, H. Su, X.P. Xu, J.S. Zhao, W. Li, X. Cai, Chin. Phys. Lett. 20 (8) (2003) 1393.
[15] M. Guida, F. Maria, Chaos Solitons Fractals 31 (2007) 527.
[16] Xinping Xu, Junhui Hu, Feng Liu, Chaos 17 (2007) 023129.

[17] Julian Sienkiewicz, Janusz Holyst, Phys. Rev. E 72 (2005) 046127.
[18] Xinping Xu, Junhui Hu, Feng Liu, Lianshou Liu, Physica A 374 (2007) 441.
[19] Liping, et al., Chin. Phys. Lett. 23 (12) (2006) 3384.
[20] C. von Ferber, T. Holovatch, Yu. Holovatch, V. Palchykov, 2008.
[21] V. Latora, M. Marchiori, Physica A 314 (2002) 109.
[22] P. Sen, S. Dasgupta, A. Chatterjee, P.A. Sreeram, G. Mukherjee, S.S. Manna, Phys. Rev. E 67 (2003) 036106.
[23] W. Li, X. Cai, Physica A 382 (2007) 693.
[24] United Nation conference on trade and development, Review of maritime transport 2007 http://www.unctad.org/en/docs/rmt2007en.pdf.
[25] Jean-Paul Rodrigue, Claude Comtois, The Geography of Transport Systems, Routledge, New York, 2006.
[26] http://www.ci-online.co.uk/.
[27] Shi Zhou, Ral J. Mondragn, IEEE Commun. Lett. 8 (3) (2004).
[28] Colizza, et al., Nature Phys. 2 (2006) 110–115.
[29] S. Maslov, K. Sneppen, Science 296 (2002) 910–913.
[30] A. Barrat, M. Barthélémy, A. Vespignani, J. Stat. Mech. (2005) P05003.
[31] M. O'Kelly, J. Transport Geography 6 (3) (1998) 171.
[32] Ross Robinson, Maritime Policy Management 25 (1) (1998) 21.
[33] M.C. Moura, M.V. Pato, A.C. Paixa, Maritime Policy Management 29 (2) (2002) 135.
[34] M. Bryan, O'Kelly, Journal of Regional Science 39 (2) (1999) 275.
[35] S. Wasserman, K. Faust, Social Networks Analysis, Cambridge University Press, Cambridge, 1994.
[36] A. Fremont, J. Transportation Geography 15 (2007) 431.