

A fine-grained perspective on the robustness of global cargo ship transportation networks

PENG Peng^{1,2}, CHENG Shifen^{1,2}, CHEN Jinhai^{3,4}, LIAO Mengdi⁵, WU Lin⁶,
LIU Xiliang^{1,7}, *LU Feng^{1,7,8}

1. State Key Laboratory of Resources and Environmental Information System, Institute of Geographic Sciences and Natural Resources Research, CAS, Beijing 100101, China;
2. University of Chinese Academy of Sciences, Beijing 100049, China;
3. Navigation Aids Technology Research Center of Jimei University, Xiamen 361021, Fujian, China;
4. National & Local Joint Engineering Research Center for Marine Navigation Aids Services, Xiamen 361021, Fujian, China;
5. College of Geomatics, Shandong University of Science and Technology, Qingdao 266590, Shandong, China;
6. Institute of Computing Technology, CAS, Beijing 100190, China;
7. Fujian Collaborative Innovation Center for Big Data Applications in Governments, Fuzhou 350003, China;
8. Jiangsu Center for Collaborative Innovation in Geographical Information Resource Development and Application, Nanjing 210023, China

Abstract: The robustness of cargo ship transportation networks is essential to the stability of the world trade system. The current research mainly focuses on the coarse-grained, holistic cargo ship transportation network while ignoring the structural diversity of different sub-networks. In this paper, we evaluate the robustness of the global cargo ship transportation network based on the most recent Automatic Identification System (AIS) data available. First, we subdivide three typical cargo ship transportation networks (i.e., oil tanker, container ship and bulk carrier) from the original cargo ship transportation network. Then, we design statistical indices based on complex network theory and employ four attack strategies, including random attack and three intentional attacks (i.e., degree-based attack, betweenness-based attack and flux-based attack) to evaluate the robustness of the three typical cargo ship transportation networks. Finally, we compare the integrity of the remaining ports of the network when a small proportion of ports lose their function. The results show that 1) compared with the holistic cargo ship transportation network, the fine-grain-based cargo ship transportation networks can fully reflect the pattern and process of global cargo transportation; 2) different cargo ship networks behave heterogeneously in terms of their robustness, with the container network being the weakest and the bulk carrier network being the strongest; and 3) small-scale intentional attacks may have significant influence on the integrity of the container network but a minor impact on the bulk carrier and oil tanker transportation networks.

Received: 2017-07-01 **Accepted:** 2017-08-24

Foundation: Key Project of the Chinese Academy of Sciences, No.ZDRW-ZS-2016-6-3; National Natural Science Foundation of China, No.41501490

Author: Peng Peng (1989–), PhD Candidate, specialized in maritime transportation GIS, complex network analysis.

E-mail: pengp@lreis.ac.cn

***Corresponding author:** Lu Feng (1970–), PhD and Professor, specialized in TGIS, complex network analysis.

E-mail: luf@lreis.ac.cn

These conclusions can help improve the decision support capabilities in maritime transportation planning and emergency response and facilitate the establishment of a more reliable maritime transportation system.

Keywords: complex network; fine-grained; cargo ship transportation network; robustness; automatic identification system

1 Introduction

Maritime transportation plays a crucial role in the world trade system. It makes up more than 90% of worldwide trade volume (Li *et al.*, 2015). Global trade stability may be significantly affected if political factors interfere with maritime transportation. For instance, the unrest in Egypt that occurred in 2011 resulted in a large-scale paralysis in the Suez port for several days, leading to a rise in the global crude oil price and causing a detour for cargo ship transportation from Europe to Asia that greatly reduced transportation efficiency¹. In addition, maritime transportation is influenced by meteorological factors, such as Hurricane Matthew in 2016, which shut down several ports in Florida². The financial crisis in 2007 also seriously affected the port system, including the handling capacity of ports and forcing adjustments to shipping routes (Goodhart, 2008). These impacts demonstrate the vulnerability of international trade systems and the urgent need to evaluate the robustness of the cargo ship transportation network.

Based on a comprehensive review of the literature, we found that previous research did not fully represent the robustness of the global cargo ship network due to the following reasons:

(1) Focusing on the coarse-grained holistic network without analyzing the diversity of cargo ship transportation (Woolleymeza *et al.*, 2011; Zong *et al.*, 2016): There are many types of cargo in maritime transportation that require different ships, including oil tankers, bulk carriers, and container ships. Each type of cargo transportation can be represented as an individual network from an operational perspective in terms of transport and cargo-handling technologies (Ducruet, 2017).

(2) Ignoring the differences in infrastructure and the functional design among different ports, which may directly affect route planning and the robustness evaluation of the cargo ship transportation network: Some ports are designed specifically for certain commodities while being unsuitable for others (Ducruet, 2013). For instance, Zhoushan port can offer container ships, bulk carriers and oil tankers for calling, but Beilun port nearby can only be used for bulk carriers. If we study the holistic cargo ship transportation network, all of the ports are roughly assumed to have the same function. When a cargo ship plans its journey, misleading information would be provided, and the real robustness status of the network may be erroneously estimated.

(3) Using statistical data offered by some liner companies, which only cover the main routes (OD information) of the container ships, and cannot fully depict the holistic cargo ship transportation network structure: For instance, the data offered by liner companies in 2016 only covers 777 ports and 7553 routes, according to the latest research (Zong *et al.*, 2016), but the data in this paper covers 1488 ports and 17135 routes just for the container ship transportation network in 2015.

¹ http://www.cnjxol.com/finance/content/2013-02/16/content_2394440.htm

² <http://news.hsdhw.com/388690>

In this paper, we employ the automatic identification system (AIS) data of global cargo ships in 2015, which can provide detailed information on every cargo ship, including the real-time location correspondence (Ristic *et al.*, 2008; Pallotta *et al.*, 2013). These data can give the most accurate reflection of the running status of a ship, and we can calculate the arrival and departure records in its calling port. In our analysis of AIS data, the holistic cargo ship transportation system constitutes many types of ships, namely, one single network made of complementary sub-networks. In these ship types, the oil tanker loads oil products, the container ship carries primarily sundry goods, and the bulk carrier loads coal, mineral. These goods represent the world's major trade products, with these three types of cargo ships making up 74% of the world's cargo transportation volume, which can greatly influence the world economy (Kaluza *et al.*, 2010).

In this paper, we subdivide three typical cargo ship transportation networks (i.e., oil tanker, container ship and bulk carrier) from the original cargo ship transportation networks. Then, by carefully choosing observation parameters based on complex network theory and employ four attack strategies including the random attack and three intentional attacks (i.e., degree-based attack, betweenness-based attack and flux-based attack), we quantitatively evaluate the robustness of the cargo network thoroughly, from the granularity of a port to the holistic structure using different types of cargo networks. Based on this view model, we can accurately measure the robustness of different types of networks.

The rest parts of this paper are organized in the following ways. In section 2, the related researches are introduced. The experimental data and the cargo transportation network structure are analyzed afterwards. In section 4, we comparatively evaluate the robustness of the three derived selected networks. In section 5, the impacts of small-scale malfunction under different attacks are depicted. In section 6, we discuss the robustness of these three cargo ship networks and compare ports' importance based on different ranking strategies. In section 7, we give a conclusion to our work.

2 Related works

The complex network theory offers us a new perspective on researching the cargo ship transportation network's robustness. Barabási put forward the concept of a scale-free network, which has recently become an important focus of multidisciplinary development (Barabási and Albert, 1999; Barabási, 2009). We find that the networks selected in this paper are scale-free networks, and thus complex network methods can be used to analyze them.

Research on maritime network analysis also exists (Notteboom, 2007; Tsiotas *et al.*, 2015). Ducruet *et al.* (2010, 2011) established its foundation, including the definition of key concepts, the correct procedures for analysis, and the levels of rigor and precision of the quantitative techniques themselves. Hu *et al.* (2009) found that the world maritime network is a small-world network with power law behavior, and hierarchical structure property and rich-club phenomenon are revealed by analyzing the weighted clustering coefficient and the weighted average nearest neighbors' degree. Due to a lack of data, the recently maritime network analyses have focused on container ship network (Notteboom, 2004; Meng *et al.*, 2011; Ducruet *et al.*, 2012). In fact, the maritime network can be defined as a multi-layer network structure composed of a series of sub-networks (Ducruet, 2017). Kaluza and Kolzsch considered that most maritime cargo ships can be divided into three categories: oil

tankers, bulk dry carriers and container ships. The three categories of ships differ greatly in their mobility patterns and network structures (Kaluza *et al.*, 2010). Ducruet (2013) investigated the interdependence of five disparate subnetworks and demonstrated the strong influence of multiplexity on the centrality of ports and the network structure.

Originating from studies of complex networks, the robustness of a network refers to its ability to maintain the functionality under attacks or failures (Holme *et al.*, 2002). A robustness analysis on the maritime network can help identify the regions sensitive to regional and large-scale failure of the network (Woolleymeza *et al.*, 2011). In addition, complex network theory already provides a powerful theoretical tool in robustness analysis (Holme *et al.*, 2002; Wang *et al.*, 2008; Ferber *et al.*, 2009), as robustness evaluation is widely applied to transportation networks such as road networks (Duan *et al.*, 2013, 2014; Yin *et al.*, 2009) and airport transportation networks (Wang *et al.*, 2011; Colizza *et al.*, 2006). Some research has been carried out on maritime network robustness evaluation (Deng *et al.*, 2008; Wu *et al.*, 2008; Woolleymeza *et al.*, 2011; Wang *et al.*, 2016). However, one issue is that recent research focused mainly on the coarse-grained holistic network robustness evaluation without looking into the diverse types of sub-networks. Woolleymeza *et al.* (2011) present a comparative network-theoretic analysis of the worldwide air transportation network and the global cargo ship network and find that the global cargo ship network is more robust to the failure of nodes. For degree-based attack, 44.3% of nodes are removed, and the network begins to split, whereas for betweenness-based attack, the threshold is 39.5%.

3 Analysis of cargo transportation network structure

3.1 Data description

In this paper, we apply the AIS data to calculate the arrival and departure records in the calling port of global cargo ships in 2015. The traffic density of all cargo ships is shown in Figure 1. We limit the cargo ships included to those heavier than 10,000 gross tonnages, which

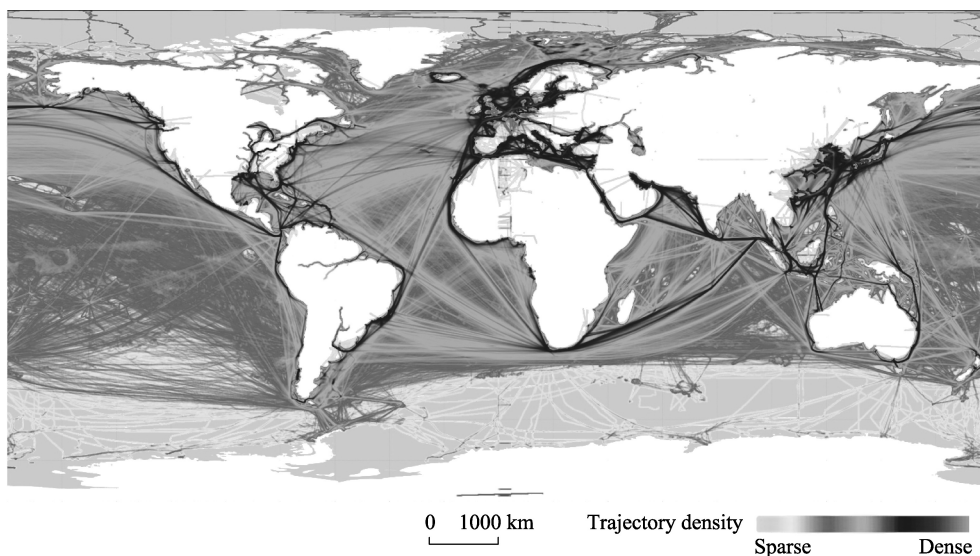


Figure 1 Traffic density of global cargo transportation in 2015

make up 89% of the total volume of the three types of cargo ship transportation. In addition, we set a threshold to exclude the journeys that ran between the two ports less than 5 times. Finally, we use the arrival and departure records to construct three cargo ship transportation networks: oil tanker, container ship and bulk carrier.

3.2 Characterization of cargo ship transportation network

The three selected cargo ship transportation networks have different network structures. Figure 2 shows the network structure of different cargo ship networks. Table 1 shows basic characteristics of the three selected cargo ship networks.

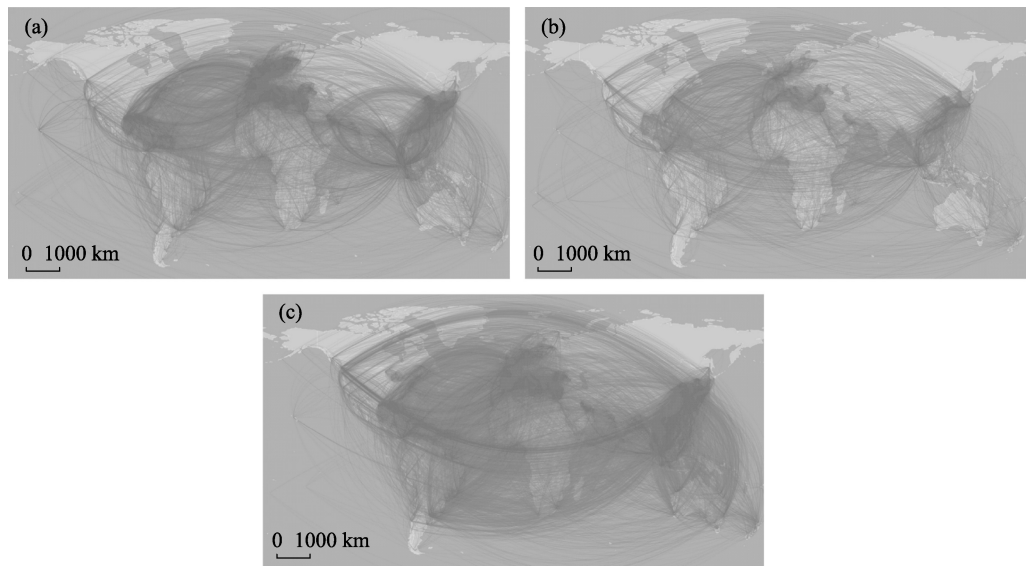


Figure 2 Network structure of oil tanker transportation (a); network structure of container ship transportation (b) and network structure of bulk carrier transportation (c) in 2015

Table 1 Characterization of different cargo ship transportation networks

Network types	No. of ships	N	E	δ	J	$\langle j \rangle$	$\langle k \rangle$	c	$\langle l \rangle$
Oil tanker transportation network	8913	2042	44219	2.1×10^{-2}	521628	255.4	43.31	0.56	2.75
Container ship transportation network	4936	1488	17135	1.5×10^{-2}	396833	266.7	23.03	0.55	2.99
Bulk carrier transportation network	10189	1969	45850	2.4×10^{-2}	274435	139.4	46.57	0.46	2.67

Note: N =Number of ports; E =Number of routes; $\delta \approx 2E/N^2$ =Network connectivity; J =Total journeys; $\langle k \rangle$ =Port mean degree; c =Network clustering coefficient; $\langle l \rangle$ =Mean shortest path length; $\langle j \rangle$ =Mean journeys per port)

On network connectivity, the container ship journeys yield a less-connected network ($\delta_c = 1.5 \times 10^{-2}$ compared to oil tanker $\delta_t = 2.1 \times 10^{-2}$ and bulk carrier $\delta_b = 2.4 \times 10^{-2}$). The bulk carrier transportation network has high connectivity ($\langle k \rangle = 49.57$), while the degree value (23.03) of the container ship transportation network confirms its sparsely connected character. Container ships always follow set schedules; they do not change routes and provide fixed services. However, for oil tankers and bulk carriers, their routes depend on the current supply and demand, and are frequently changed in a short time. Thus, these two types of networks

have more links than the container ship network (34,270 routes for container ship network compared with 44,219 routes for the oil tanker network and 45,850 for the bulk carrier network) and more visited ports than the container ship transportation network (1488 ports for the container ship network compared with 2042 ports for the oil tanker transportation network and 1969 ports for the bulk carrier transportation network), leading to a low mean degree value for the container ship transportation network. On the mean shortest path length, the bulk carrier transportation network has the smallest value ($\langle l \rangle = 2.67$), which means that the goods in this network can be transmitted more rapidly than the other two networks ($\langle l \rangle = 2.75$ and $\langle l \rangle = 2.99$). The clustering coefficients of the container ship transportation network ($C=0.55$) and oil tanker transportation network ($C=0.56$) are densely clustered, whereas the bulk carrier transportation network is less clustered, with a coefficient value of $C=0.46$.

The difference also exists in the betweenness centralities of the three types of networks. Some ports rank highly in all the three categories (e.g., Singapore ranks in the top 10 in all categories), whereas others depend on certain types of networks. For example, the Zhoushan port in China ranks 5th in the container ship transportation network, 18th in the bulk carrier transportation network and only 38th in the oil tanker transportation network.

3.3 Analysis of port importance metrics

Figure 3 shows the degree distribution, under double logarithmic coordinates, of these three cargo ship transportation networks, indicating that they are heterogeneous. Its x-axis represents the degree of ports, and the y-axis indicates the proportion of degree distribution $P(k)$. The result in Figure 3 demonstrates that most ports have few port connections, however, there exist some important ports, such as Singapore port, that are visited by many ships from different ports and that link hundreds of ports. We carry out the power law function fitting with all of the R-square values larger than 0.7 for the three types of cargo ship transportation networks and find that their degree distribution follows power-law distribution. This result reveals that these three networks are scale-free network (Barabási and Albert, 1999). In section IV, we conduct a comprehensively quantitative robustness analysis on them.

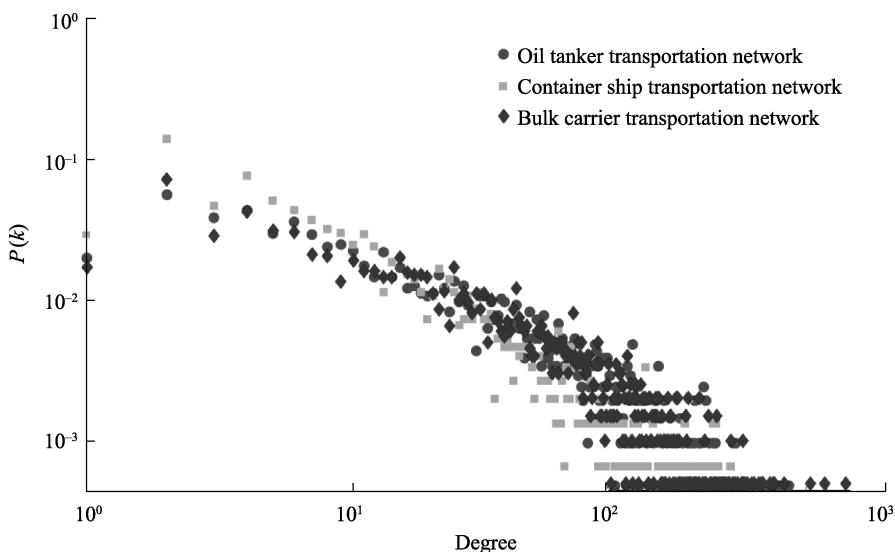


Figure 3 The degree distribution of cargo ship transportation networks

Figure 4 shows the betweenness centrality distribution of the three cargo ship transportation networks under double-log plots. The x-axis represents the betweenness centrality value, and the y-axis represents the proportion of ports with a given betweenness centrality depicted by $P(bc)$. As shown in Figure 4, almost all of the ports have different betweenness centrality values.

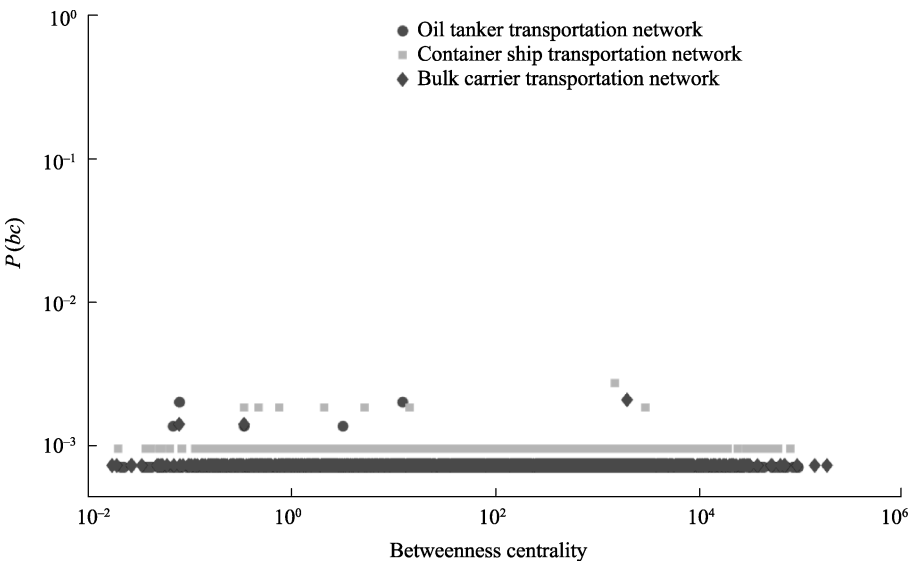


Figure 4 The betweenness centrality distribution of cargo ship transportation networks

Figure 5 shows the flux distribution of the three cargo ship transportation networks under double logarithmic coordinates. The x-axis represents the flux value, and the y-axis represents the proportion of ports with a given flux depicted by $P(f)$. Figure 5 demonstrates that the three networks have a significant diversity of flux distributions. Most of the ports have few cargo ships passing through, but some ports are visited by many cargo ships.

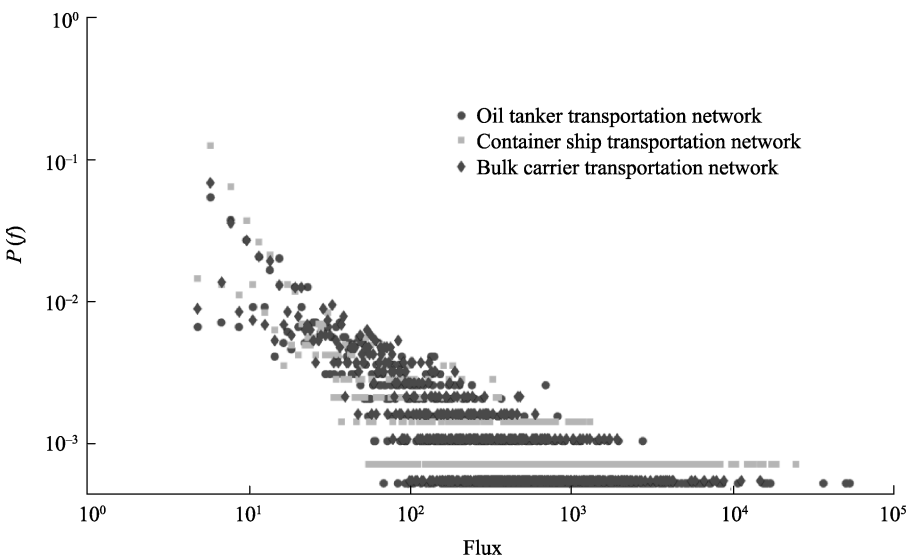


Figure 5 The flux distribution of cargo ship transportation networks

Figures 3–5 indicate that great differences exist in degree, betweenness centrality and flux. As indicated in Figure 6, we calculated the correlation coefficient between each of the two metrics in our empirical data set to further confirm that the three metrics are not always correlated, particularly in the oil tanker transportation network, which shows a

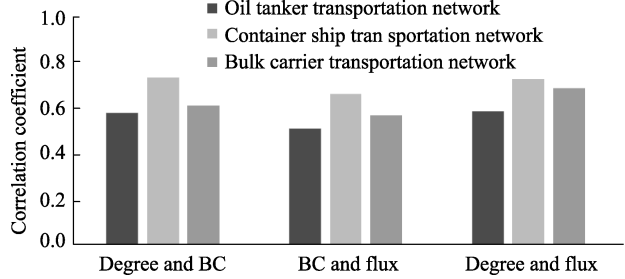


Figure 6 The correlation coefficient of the three metrics

weak correlation. These findings suggest that the destruction process caused by different types of attack, including attacks on high degree ports, on ports with high betweenness centrality and on high flux ports, are quite different. Thus, the degree, betweenness centrality and flux should all be taken into consideration when selecting the target port in the simulations.

4 Robustness evaluation of cargo transportation network

In this paper, we adopt an undirected graph $G = (N, E)$ to model the cargo ship transportation network, where N represents ports in the cargo ship transportation system, and E is the set of all routes linking two ports in N . The G represents the global cargo exchange relations.

4.1 Metrics

4.1.1 Metrics of port importance

In this paper, we choose three metrics, namely, degree, betweenness centrality and flux, to evaluate the port importance of the cargo ship transportation network.

According to the complex network theory, a higher degree port means that the port connects to more ports and contributes to increased accessibility in the local area (Kay, 1977). Its definition is as follows:

$$d_v = \sum_{j \in N, j \neq v} e_{vj} \quad (1)$$

where d_v is the degree value of port v and e denotes the element of the adjacent matrix. If there exists a direct route between port v and port j , $e_{vj} = 1$, else $e_{vj} = 0$.

A further indication of the port importance is its betweenness centrality value (Barthélemy, 2004). Generally, the betweenness centrality value of a port is the number of topologically shortest paths passing through the port of the network, i.e., the Suez Canal is the shortcut from the Indian Ocean to Europe. However, in some cases, other factors can also cause high betweenness centrality values. Shanghai port, for instance, has a high betweenness centrality because it is connected to many other ports. The definition is as follows:

$$Bc(v) = \sum_{i \neq j \in N} \frac{\sigma_{ij}(v)}{\sigma_{ij}} \quad (2)$$

where σ_{ij} represents the number of topologically shortest paths between any two port i and port j and $\sigma_{ij}(v)$ represents the path that passes through port v .

Port flux is also a vital metric for measuring port importance. A higher flux port means that more cargo ships are passing through it. The definition is as follows:

$$F(v) = \sum_{j \neq v \in N} w_{vj} \quad (3)$$

where $F(v)$ indicates the number of cargo ships which pass through the port v , and w_{vj} means the number of cargo ships trading between port v and port j .

4.1.2 Metrics of network complexity

For cargo ships, the most vital factor is its running efficiency, including the local and global efficiency. In the complex network, efficiency is usually described by the average length of the shortest paths l , which stands for global efficiency, and the local efficiency measurement – clustering coefficient C (Albert and Barabási, 2002). l measures the topologically distance between two ports, and a smaller value of l implies that a shorter topologically path length between the ports and the goods on the network can be transmitted rapidly. C measures the level of connectivity of a typical neighborhood, and a higher C indicates that there is a more clustered ports around, thus the goods can be transported in high-efficiency within the local area. The definition of l is as follows:

$$l = \frac{1}{n(n-1)} \sum_{v \in V} \sum_{w \neq v \in V} d(v, w) \quad (4)$$

where n is the number of ports of the network and $d(v, w)$ denotes the length of the topologically shortest path between port v and port w by counting the number of ports in the shortest path. If the graph is not connected, l equal to the average path length of the largest connected component of the graph.

To compare the different types of networks, we adopt the normalized average path length L to stand for the global efficiency:

$$L = \frac{l}{D} \quad (5)$$

where D is the network diameter, which is the largest topologically shortest path length between any two ports and can be calculated as follows:

$$D = \text{Max}_{v, w \in N, v \neq w} (d(v, w)) \quad (6)$$

where n is the number of ports of the network and $d(v, w)$ has the same definition as depicted in formula (4).

C is the mean clustering coefficient of the network. Its definition is as follows:

$$C = \frac{1}{N} \sum_{v \in V} C_v \quad (7)$$

N is the number of ports. C_v denotes the clustering coefficient of port v :

$$C_v = \frac{|\{e_{ij} : i, j \in N_v, e_{ij} \in E\}|}{k_v(k_v - 1) / 2} \quad (8)$$

where N_v is the set that connected port v , k_v is the total number of set N_v , $k_v(k_v-1)/2$ is the maximum possible routes between k_v ports, and e_{ij} is the routes existing in the network between all k_v connected ports with port v .

Cargo ship transportation networks may suffer severe damage but will not collapse. We thus introduce S , which stands for the relative size of the largest connected component, to reflect the structural variations of cargo ship transportation networks when under attacks, and to investigate the impact of attacks on the network structure. Its definition is as follows:

$$S = \frac{n_s}{n} \quad (9)$$

where n_s is the ports in the largest connected component of the remaining network and n is the ports in the original network.

4.2 Experiment design

The network robustness can be expressed as a function of how the network reacts to attacks (Callaway *et al.*, 2000). According to the cause of formation, attacks can be grouped into two types: non-recurrent incidents such as typhoons or hurricanes, which occur in different ports at random, and other failures like the terrorist attacks, which target the most important ports. To quantitatively evaluate the network robustness, we designed a whole train of experiments to simulate the reactions of the networks under real attacks (as described in section 3.2), including four attack strategies: one random attack strategy and three intentional attack strategies based on degree, betweenness centrality and flux. To quantitatively measure overall robustness of the networks, we conducted successive attacks to continuously destroy the network structure until it collapses (the network is broken into single ports only). Meanwhile, we recorded the structural changing status during the whole attack process.

For the random attack strategy, we randomly deleted one port and its linked routes at each step from the network. Then, we loop the procedure until the whole network collapses, namely, $S \approx 0$, which means that the network collapses into single port only. In addition, the performance (efficiency and fragmentation) of a network under attack is quantified.

In degree-based intentional attack, the highest degree port is removed at each loop, and then the degree of each port are recalculated. We repeated the removal procedure until the whole network collapses. The betweenness centrality-based attack and flux-based attack are similar to the degree-based attack, except that we removed the port that had the highest betweenness centrality value or the highest flux value.

4.3 Experiment results

In this section, we evaluate the robustness of the three types of cargo ship transportation networks by using the evaluation indices described in section 3.2. Figures 7–9 and 11 depict the results of the four attack strategies; all of their x-axes represent the proportion of the total ports that have been removed (denoted as f), and the y-axes represent the variations of L , C , S , which record the mean shortest path, clustering coefficient and fragmentation of the observed network, separately. A red line represents the variation of oil tanker transportation networks, an orange line represents the variation of container ship transportation networks, and a blue line represents the variation of bulk carrier transportation networks. Figure 7 shows the results of the random attack strategy.

At the beginning of the attack procedure, the values of L and C show tiny changes, which means that the network efficiency is the highest (including both global efficiency and local efficiency), and decreases gradually. Then, L increases, C decreases, and f continuously increases, which means that the network efficiency begins to decrease quickly. At the end, we can observe that L decreases and C decreases, and although the network is not completely collapsed, it will likely not function well.

By removing the ports gradually, barely any large changes of S occur. In other words, random attack almost avoids making fatal errors to the networks. This reveals that the three

types of cargo ship networks are closely linked, and random attack barely harms them. Figure 8 shows the results of the degree-based attack strategy.

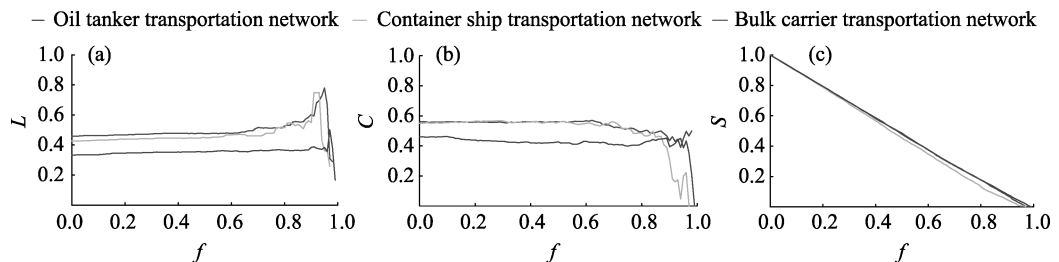


Figure 7 Network structural changes under random attack

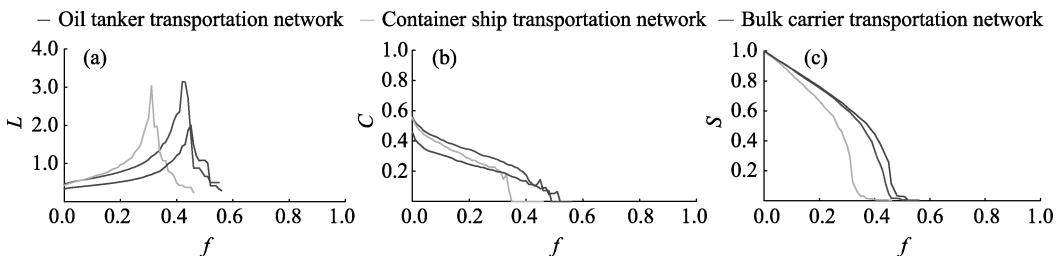


Figure 8 Network structural changes under degree-based attack

By comparing the variation of S in Figures 7 and 8, it can be concluded that the degree-based attack is more harmful than the random attack for the three networks. Under degree-based attack, only 60% of ports removed are sufficient to cause the network to collapse. Compared with the random attack, where the network nearly remained intact during the entire attack procedure, under degree-based attack the whole network begins to split into small components when only a few percent of the ports are removed.

Unlike for the random attack, where ports are removed at the beginning of the attack procedure, the average distance between ports increases as the clustering coefficient decrease for the degree-based attack. A gradually increasing L indicates a decrease in global efficiency. The removal of ports with a high degree, which are vital in local area, leads to a decrease of C . Then, L begins to decrease and C continuously decreases. Since the measurement L is highly impacted by the initial network scale and the current largest component scale, L can only provide the efficiency of the largest component when the network splits into some components, which may be useless for the overall network. By contrast, C depends on the local connection of ports, and fragmentation may not have a great impact on C . Figure 9 shows the results of the betweenness-based attack strategy.

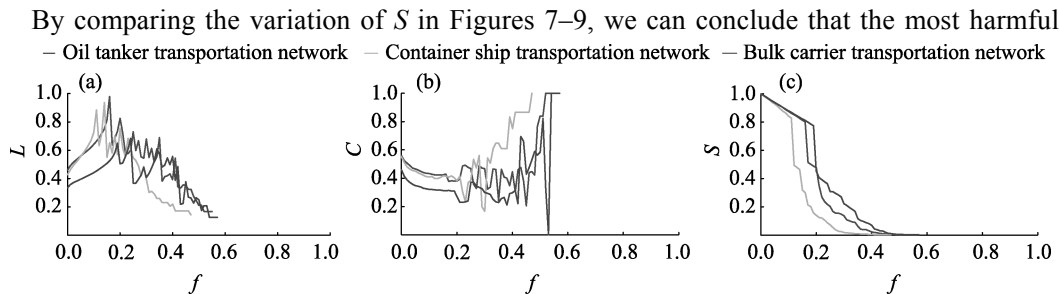


Figure 9 Network structural changes under betweenness-based attack

attack strategy for the three networks is betweenness-based attack; if only approximately 50% of ports are removed, the network collapses. Compared with degree-based attack, fewer ports need to be removed to split the network into small components.

Similar to the degree-based attack, the average distance between ports increases when ports are removed, as the clustering coefficient decreases at the beginning of the attack procedure. Then, L begins to decrease. Unlike the variation under random attack and degree-based attack, the local efficiency C under a betweenness-based attack increases with the removal of high betweenness ports. This occurs mainly due to the ports with high betweenness value connect the different components in the network. Thus, the adjacent ports of these ports more likely belong to different components, which lowers their mean clustering coefficient. The removal of these ports cause a higher average clustering coefficient C of the overall network. Figure 10 shows the correlation coefficient results of the three cargo ship transportation networks. We calculated the correlation coefficient between the betweenness centrality and the clustering coefficient of ports with high betweenness centrality to affirm our explanation (we calculate the betweenness centrality and clustering coefficient of each removal port). The results show that ports possessed high betweenness centrality are apt to own a lower clustering coefficient in this procedure. Thus, the removal of these ports will cause an increase in the clustering coefficient of the overall cargo ship network. Figure 11 shows the results of the flux-based attack strategy.

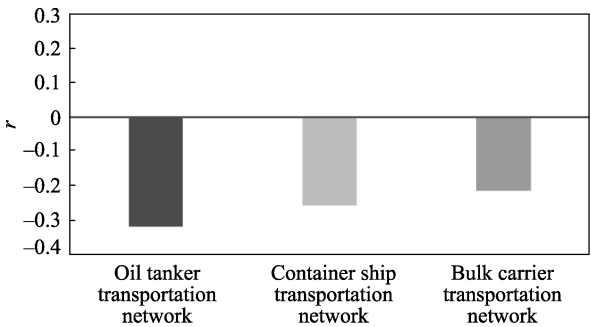


Figure 10 Correlation coefficient between betweenness centrality and clustering coefficient.

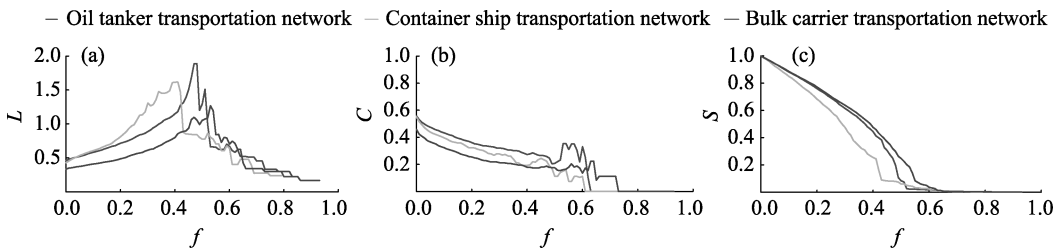


Figure 11 Network structural changes under flux-based attack

It can be concluded that the variation of the network structure under flux-based attack strategy is similar to the variation of the degree-based attack, by comparing the variation of network structure in Figures 7 and 11. When ports are removed, L increases gradually together with the decrease of C ; then, L begins to decrease and C continuously decreases. However, the main difference between these two attack strategies is that the degree-based attack is more destructive than the flux-based attack, which means that more ports should be removed to make the change comparable with degree-based attack.

For all four of the attacks, the variation of the three selected cargo ship transportation networks under successive attacks can be categorized into three phases. In the first phase,

the network efficiency decreases with an increase of ports removal. At this phase, only a few ports are removed or separated from the network, and the largest component comprises most ports yet. In the second phase, networks begin to break into several components. This fragmentation causes the size of the largest component to decrease quickly. In the final phase, the networks split apart into tiny fragments that cannot function well. This malfunction is considered to be the collapse of a network.

4.4 Splitting threshold of cargo transportation networks

From Figures 7–9 and 11, we know that the cargo ship network has a certain ability to resist different types of attacks. Despite the decrease in transportation efficiency, the network still has the ability to function before it splits; however, once the network begins to split, it will be severely damaged and quickly collapse. Identifying the splitting threshold of different kinds of networks will help them avoid total collapse. It is necessary to be vigilant regarding the potential for a total collapse when a certain number of ports becomes damaged. Thus, to identify the splitting threshold is of great importance. Figure 12 shows the splitting threshold of the three types of cargo ship networks.

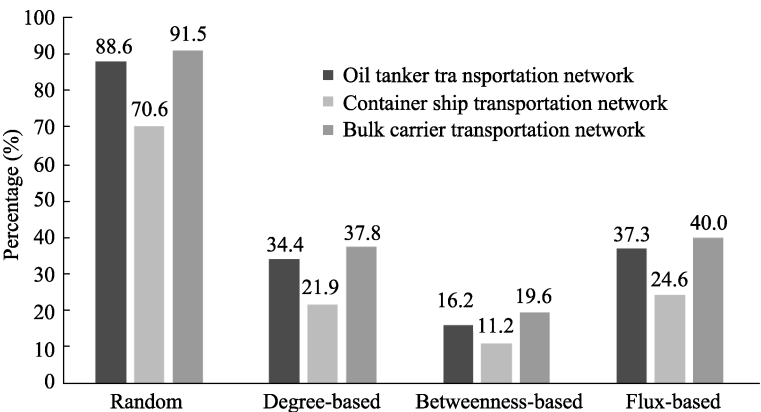


Figure 12 The splitting threshold of different types of cargo ship transportation network

From Figure 12, we know that a limited number of attacks cannot cause the entire network to split. For random attack, the container ship transportation network begins to malfunction when more than 70.6% of the ports are removed (this threshold is 88.6% for the oil tanker transportation network and 91.5% for the bulk carrier transportation network). However, the degree-based attack is more devastating, as the removal of 21.9% of the ports will cause the container ship transportation network fragmentation (this threshold is 34.4% for the oil tanker transportation network and 37.8% for the bulk carrier transportation network), and the local connection of the attacked area will more rapidly become paralyzed. Attacks based on betweenness cause the network to split after important ports, like Port Said, are removed. These ports hold the overall network together, and when they are harmed, the network will separate into several components. In addition, for betweenness-based attack, the splitting threshold of the container ship transportation network is only 11.2% (this threshold is 16.2% for the oil tanker transportation network and 19.6% for the bulk carrier transportation network). For flux-based attack, the container ship transportation network begins to malfunction when 24.6% of the ports are removed (this threshold is 37.3% for the

oil tanker transportation network and 40.0% for the bulk carrier transportation network). These quantitative splitting thresholds can be regarded as an early warning sign to maintain network safety.

4.5 Comparison of the three types of networks

The comparison results can be assessed by analyzing Figures 7–9 and 11. From Figure 7, we can see that the three types of networks possess the same change procedure, without apparent differences, under random attack strategy. The small differences can be observed by analyzing the splitting threshold in Figure 12; the container ship transportation network is easier to split compared with the oil tanker transportation network and the bulk carrier transportation network.

Under intentional attacks, as shown in Figure 8, the container ship transportation network first splits into small parts when applying degree-based attack by removing ports from the network. Although the oil tanker transportation network and bulk carrier transportation network have a similar number of ports and routes, differences exist between these two networks, and it appears that the bulk carrier network is more difficult to split than the oil tanker network. The collapse order of the three types of network is the same as the splitting order – the container ship transportation network collapses first, then the oil tanker transportation network, and finally the bulk carrier transportation network. In conclusion, the robustness of the three types of networks is ranked in the following order, from weak to strong: container ship network, oil tanker network, and bulk carrier network. We produced the same results for the betweenness-based and flux-based attack strategies.

5 The impacts of small-scale malfunction under different attacks

In section 4, we simulated successive attacks to continuously destroy the network structure until it collapses, whereas in reality, the maritime transportation network will never become completely collapsed. Thus, we focus on the integrity of the remaining ports of the network when small-scale ports lose their function. As the experimental results show in section 4.4, for the most harmful attack strategy – betweenness-based attack – more than 11.2% of simultaneously removed ports (>100 ports) are needed to begin to split the network. Therefore, in this paper, we set the ratio at less than 10%. We compare the number of isolated ports disjointed from the largest component (*r*) and the size of the largest component(*s*) at different removal ratios (from 1% to 10%, with an interval of 1%). The results are shown in Tables 2–5.

Table 2 Remaining ports and the size of the largest component under random attack with different ratios

Ratio		1%	2%	3%	4%	5%	6%	7%	8%	9%	10%
Oil tanker	r	1	1	1	1	1	2	2	5	5	5
	s	2020	2000	1979	1959	1938	1917	1897	1873	1853	1832
Container ship	r	1	2	2	5	5	5	5	7	8	8
	s	1472	1457	1441	1423	1408	1393	1378	1361	1346	1331
Bulk carrier	r	1	1	1	3	4	6	6	7	8	8
	s	1948	1928	1908	1887	1866	1844	1825	1804	1783	1764

For random attack results shown in Table 2, even when 10% of ports are removed, only 8 ports are isolated from the largest component for the container ship transportation network (along with 5 for the oil tanker transportation network and 8 for the bulk carrier network). Compared with the size of the largest component and the isolated ports disjointed from the largest component, we can conclude that the random attack barely threatens the integrity of the network.

Table 3 Remaining ports and the size of the largest component under degree-based attack with different ratios

Ratio		1%	2%	3%	4%	5%	6%	7%	8%	9%	10%
Oil tanker	r	5	6	6	13	16	20	22	24	25	25
	s	2016	1995	1974	1947	1923	1899	1877	1854	1833	1812
Container ship	r	9	15	18	26	33	41	53	60	73	88
	s	1464	1443	1425	1402	1380	1357	1330	1308	1281	1251
Bulk carrier	r	5	9	13	13	15	18	18	22	24	28
	s	1944	1920	1896	1877	1855	1832	1813	1789	1767	1744

Table 4 Remaining ports and the size of the largest component under betweenness-based attack with different ratio

Ratio		1%	2%	3%	4%	5%	6%	7%	8%	9%	10%
Oil tanker	r	6	8	15	22	23	25	26	27	27	31
	s	2015	1993	1965	1938	1916	1894	1873	1851	1831	1806
Container ship	r	8	15	18	30	33	43	55	66	70	92
	s	1465	1443	1425	1398	1380	1355	1328	1302	1284	1247
Bulk carrier	r	4	10	18	20	22	24	25	25	29	32
	s	1945	1919	1891	1870	1848	1826	1806	1786	1762	1740

Table 5 Remaining ports and the size of the largest component under flux-based attack with different ratios

Ratio		1%	2%	3%	4%	5%	6%	7%	8%	9%	10%
Oil tanker	r	4	6	7	10	15	17	23	25	27	28
	s	2017	1995	1973	1950	1924	1902	1876	1853	1831	1809
Container ship	r	7	10	17	29	39	44	52	61	73	86
	s	1466	1448	1426	1399	1374	1354	1331	1307	1281	1253
Bulk carrier	r	6	8	9	12	15	15	18	20	24	29
	s	1943	1921	1900	1878	1855	1835	1813	1791	1767	1743

Combining the results shown in Tables 3–5 with the variation of S in Figures 8, 9 and 11, we conclude that the three intentional attacks, which include degree-based attack, betweenness-based attack and flux-based attack, can be regarded as having a similar changing process when the removal ports are less than 10%. In addition, the intentional attacks cause a higher fluctuation than the random attack, especially for the container ship transportation network. Using the degree-based attack as an example for the container ship transportation network, with the proportion of removal ports continuously growing, more ports become isolated from the largest component compared with the other two networks. When 10% of ports are removed, the number of remaining ports, except the largest component (88 ports), is 3 times greater than the other two networks (25 ports for the oil tanker transportation network and 28 ports for the bulk carrier transportation network.). In addition, the oil tanker

transportation network and the bulk carrier transportation network change without showing any noticeable differences. Thus, we can conclude that the intentional attacks cause greater harm to the container ship transportation network than the oil tanker transportation network and the bulk carrier network.

6 Discussion

According to the definition of network fragmentation (S), when the fitting curve for the proportion of removal ports (f) approaches the network's fragmentation (S), with $S = 1 - f$, then the removal of ports does less harm to the network's integrity. On the other hand, the closer the fitting curve approaches $f = 0$, the greater harm the removal of ports may do to the network's integrity. Next, we analyze the results shown in section 5.

From the variation of S in Figure 7, we can calculate that for random attack, the fitting curve is almost $S = 1 - f$ when the removal ports are less than 10%, which means that a random attack will not cause great damage to the marine transportation network. From the variation of S in Figures 8, 9 and 11, we can calculate that for the three intentional attacks, the fitting curve of the bulk carrier transportation network and the oil tanker transportation network is $S = 1 - 1.15f$, whereas the container ship transportation network is $S = 1 - 1.6f$ when 10% of the ports are removed from each network. Tables 3–5 show that when 10% of the ports are removed, approximately 90 ports (6%) disconnect from the largest remaining component of the container ship transportation network. Meanwhile, only approximately 30 ports (1.5%) disconnect from the largest remaining component of the bulk carrier transportation network and the oil tanker transportation network. Therefore, the removal of a small proportion of ports has a greater impact on the integrity of the remaining ports of the container ship transportation network but has a relatively smaller impact on the bulk carrier transportation network and the oil tanker transportation network.

In addition, compared with the other transportation networks, the removal of a smaller proportion of ports does relatively less damage to the integrity of the maritime transportation network. For instance, when 10% of the ports are removed based on degree-based attack for the worldwide air transportation network, the fitting curve is $S = 1 - 4f$, which means 30% (approximately 1200 airports) of the airports disconnect from the largest remaining component of the network (Woolleymeza *et al.*, 2011).

In summary, the possible reasons why the three selected networks seem to hardly split under all four of the attack strategies include the following:

1) The connection between each port in the cargo ship transportation network is very dense, as depicted in section 3. Both the network connectivity and high mean degree value confirm this result, so a small failure will not affect the network's integrity.

2) An important port (high degree value or high betweenness value) has one or more neighboring ports with the same function. Thus, when an important port is malfunctioning, its neighboring port can be used for transit, such as the Shanghai port and Zhoushan port, Singapore port and the Port Kelang, Hong Kong port and Shenzhen port, and Suez and Port Said.

Ranking ports based on flux is another strategy to evaluate their importance, similar to the ranking strategy based on topological structure. From the experimental results in section 4.3, we can conclude that the flux-based attack is more destructive than random attack but less harmful than the intentional attack, according to the topological structure. In addition, we

conclude that high betweenness centrality value ports are extremely important for maintaining the stability of the cargo ship transportation network. Therefore, we rank the port importance of the three selected cargo ship transportation networks based on their betweenness centrality value and flux value, and then pull out the top ten ports. We then compare this with the top 10 ports of the three selected cargo ship transportation networks, which are ranked based on their flux value.

In Figure 13, we compare the three selected networks according to the topological structure and the port flux approaches. In the three columns of the port flux, we find that only Khawr Fakkan, Zhoushan, Shanghai, Pusan and Europa Point are also in the topological structure columns, which means that although some ports have a huge amount of flux, namely cargo handling capacity, they do not play a dominant role in the global cargo ship transportation network. In addition, by comparing the ranking of ports according to the two ranking strategies, we find that most of the ports with a topological structure also have a huge amount of flux. Therefore, according to the topological structure, the ranking of ports may result in the flux-based attack appearing more destructive than random attack but less harmful than the intentional attack.

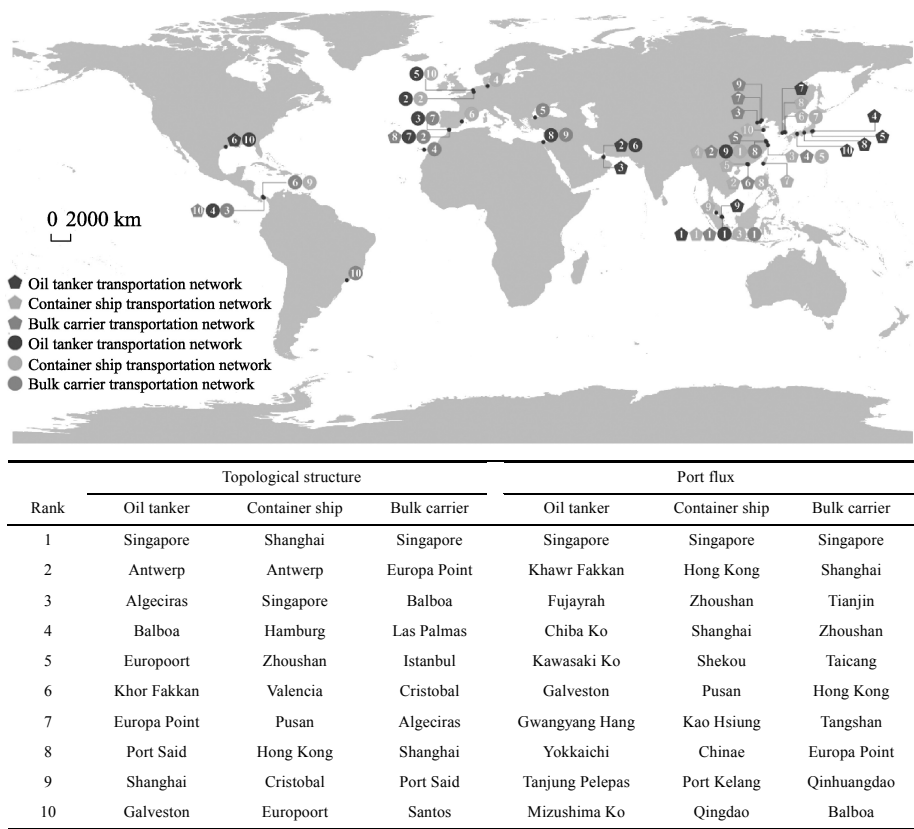


Figure 13 The top 10 ports of different networks in the world

7 Conclusions

In this paper, we build different types of cargo ship transportation networks by using AIS data

and then quantitatively study these networks' robustness. The conclusions are as follows:

(1) In contrast to the network constructed with the container liner OD information, the global cargo ship transportation network constructed with ships' AIS trajectories can fully reflect the pattern and variation process of the global maritime transportation network.

(2) Quantitative robustness evaluations are applied to these types of cargo ship transportation networks. By analyzing the networks' robustness under four attack strategies, their ranking was as follows, from weakest to strongest: container ship network, oil tanker network, and bulk carrier network.

(3) Small-scale port failures have a major impact on the connectivity of the container ship transportation network but have almost no impact on the connectivity of the bulk carrier transportation network and the oil tanker transportation network. Therefore, an opportunity remains to optimize the route to improve the robustness of the container ship transportation network.

We study the robustness of the cargo ship transportation network by subdividing the cargo ship types to characterize the fine-grained structure of the marine transportation network. Our findings can be considered a first step in cargo ship transportation network robustness optimization.

References

- Albert R, Barabási A, 2002. Statistical mechanics of complex networks. *Review of Modern Physics*, 74(1): xii.
- Barabási, A, 2009. Scale-free networks: A decade and beyond. *Science*, 325(5939): 412–413.
- Barabási A, Albert R, 1999. Emergence of scaling in random networks. *Science*, 286(5439): 509–512.
- Barthélemy M, 2004. Betweenness centrality in large complex networks. *The European Physical Journal B*, 38(2): 163–168.
- Berche B, Ferber C V, Holovatch T *et al.*, 2009. Resilience of public transportation networks against attacks. *The European Physical Journal B*, 71(1): 125–137.
- Callaway D S, Newman M E J, Strogatz S H *et al.*, 2000. Network robustness and fragility: Percolation on random graphs. *Physical Review Letters*, 85(25): 5468–5471.
- Colizza V, Barrat A, Barthélemy M *et al.*, 2006. The modeling of global epidemics: Stochastic dynamics and predictability. *Bulletin of Mathematical Biology*, 68(8): 1893–1921.
- Deng G, Wu P, Tian W, 2008. Research on robustness and vulnerability of global shipping network. *Journal of Dalian University of Technology*, 48(5): 765–768. (in Chinese)
- Duan Y, Lu F, 2013. Structural robustness of city road networks based on community. *Computers Environment & Urban Systems*, 41(9): 75–87.
- Duan Y, Lu F, 2014. Robustness of city road networks at different granularities. *Physica A Statistical Mechanics & Its Applications*, 411(411): 21–34.
- Ducruet C, 2013. Network diversity and maritime flows. *Journal of Transport Geography*, 30(2): 77–88.
- Ducruet C, 2017. Multilayer dynamics of complex spatial networks: The case of global maritime flows (1977–2008). *Journal of Transport Geography*, 47–58.
- Ducruet C, Lee S W, Song J M, 2011. Network position and throughput performance of seaports. *Current Issues in Shipping Ports & Logistics*, 189–201.
- Ducruet C, Rozenblat C, Zaidi F, 2010. Ports in multi-level maritime networks: Evidence from the Atlantic (1996–2006). *Journal of Transport Geography*, 18(4): 508–518.
- Ferber C V, Holovatch T, Holovatch Y, 2009. Attack Vulnerability of Public Transportation Networks. Berlin Heidelberg: Springer.
- Goodhart C A E, 2008. The background to the 2007 financial crisis. *International Economics and Economic Pol-*

- icy*, 4(4): 331–346.
- Holme P, Kim B J, Yoon C N *et al.*, 2002. Attack vulnerability of complex networks. *Physical Review E Statistical Nonlinear & Soft Matter Physics*, 65(2): 634–634.
- Hu Y, Zhu D, 2009. Empirical analysis of the worldwide maritime transportation network. *Physica A Statistical Mechanics & Its Applications*, 388(10): 2061–2071.
- Kay E, 1977. Graph theory with applications. *Journal of the Operational Research Society*, 28(1): 237–238.
- Li Z, Xu M, Shi Y, 2015. Centrality in global shipping network basing on worldwide shipping areas. *GeoJournal*, 80(1): 47–60.
- Meng Q, Wang S, 2011. Liner shipping service network design with empty container repositioning. *Transportation Research Part E Logistics & Transportation Review*, 47(5): 695–708.
- Notteboom T E, 2004. Container shipping and ports: An overview. *Review of Network Economics*, 3(2): 86–106.
- Notteboom T E, 2007. Container shipping and ports: An overview. *Review of Network Economics*, 3(2): 1–21.
- Pallotta G, Vespe M, Bryan K, 2013. Vessel pattern knowledge discovery from AIS data: A framework for anomaly detection and route prediction. *Entropy*, 15(6): 2218–2245.
- Ristic B, La Scala B, Morelande M *et al.*, 2008. Statistical analysis of motion patterns in AIS data: Anomaly detection and motion prediction. *International Conference on Information Fusion*, 29: 109–122.
- Tsiotas D, Polyzos S, 2015. Analyzing the maritime transportation system in Greece: A complex network approach. *Networks and Spatial Economics*, 15(4): 1–30.
- Wang J, Mo H, Wang F *et al.*, 2011. Exploring the network structure and nodal centrality of China's air transportation network: A complex network approach. *Journal of Transport Geography*, 19(4): 712–721.
- Wang J, Rong L, Zhang L *et al.*, 2008. Attack vulnerability of scale-free networks due to cascading failures. *Physica A Statistical Mechanics & Its Applications*, 387(26): 6671–6678.
- Wang N, Dong L, Wu N *et al.*, 2016. The change of global container shipping network vulnerability under intentional attack. *Acta Geographica Sinica*, 71(2): 293–303. (in Chinese)
- Woolleymeza O, Thiemann C, Grady, D *et al.*, 2011. Complexity in human transportation networks: A comparative analysis of worldwide air transportation and global cargo-ship movements. *The European Physical Journal B*, 84(4): 589–600.
- Wu P, 2008. Research on topology character of container shipping network. *Journal of Wuhan University of Technology*, (4): 665–668. (in Chinese)
- Yin Y, Madanat S M, Lu X, 2009. Robust improvement schemes for road networks under demand uncertainty. *European Journal of Operational Research*, 198(2): 470–479.
- Zong K, Hu Z, 2016. Maritime association of countries along One Belt and One Road based on the perspective of social network analysis. *Journal of Dalian University of Technology*, 42(4): 84–90. (in Chinese)