ISSN 2616-5783 Vol. 2, Issue 5: 16-30, DOI: 10.25236/AJHSS.020503

Port Congestion Evaluation Based on Gaussian Density Clustering——Taking Qingdao Port as an Example

Lingyi Li

College of Transport&Communications, Shanghai Maritime University, Shanghai 200000, China

ABSTRACT. *By extracting and processing the AIS data, the berthing and berthing records of a certain port are obtained for a period of time, and the time required for different ships to enter and leave the port under the congestion of such ports is calculated, and the grading is used as an evaluation index of the congestion degree of the port. The classification of the category boundary is completed by the clustering method, wherein the Gaussian fitting curve can determine the maximum and minimum values of the index, and finally determine the level limit by optimizing the maximum and minimum point points, and the ship data is obtained. Divided into different congestion states. It provides research ideas for the analysis of ship congestion, and also provides data support for mitigating port congestion.*

KEYWORDS: *Port congestion, Clustering, Gaussian density clustering*

1. Introduction

With the rapid development of China's economy, international trade is also developing continuously. As one of the country's important transportation facilities, the port plays a key role in the country's economic development, especially the development of international trade. However, due to the continuous increase in freight volume and various other reasons, the development of many ports has been restricted to a certain extent. According to statistics, most of the current ports are congested, and the production capacity of the port must be improved. Before this, it is necessary to make an objective quantitative evaluation of the congestion degree of the port, so as to formulate a targeted plan to alleviate port congestion.

The factors causing port congestion are complex, and there is uncertainty in time and space. The packaging, weight and category of goods have strong randomness.The arrival time of the vehicles arriving at the port , and different types of ships is relatively random. At the same time, meteorological factors will also have a great impact on the production of the port, which will affect the port handling

ISSN 2616-5783 Vol. 2, Issue 5: 16-30, DOI: 10.25236/AJHSS.020503

efficiency. The complexity of the operation of the internal agencies of the port will also affect the congestion of the port.

To solove the port congestion problems, many domestic and foreign scholars have conducted research. Mainly by strengthening the port production organization to improve the practice of port enterprises and administrative departments,or improve port management level and other methods to deal with port congestion. The universality of the method is poor, while the influencing factors of port congestion are complex, and the management of the port is only one of them. It is difficult to make an effective and targeted approach to a certain factor, and the situation of different ports is different. The method is hard to work. Therefore, it is necessary to analyze the basic situation of each port one by one in order to effectively alleviate port congestion. The situation of each port is different, so an objective evaluation method is needed to evaluate the congestion of each port. However, as time and external conditions are different, like the road traffic, the congestion of the port will also change. The evaluation method should be a real-time model.

GPCI (Global Port Congestion Index Report),published in the form of a weekly report, detailed and timely statistics reflecting the coal remaining in major ports around the world.The number of dry bulk carriers such as ore, analyzed due to the ship.The impact of port and port congestion on the supply and demand relationship in the dry bulk market and the rent of bulk carriers.The demurrage of dry bulk carriers such as coal and ore in major ports around the world covers the congestion of 80 coal and ore terminals in different countries in the world, including Australia, Brazil, China, India, Taiwan and South Africa. However, its timeliness is insufficient and it cannot reflect the congestion level of the port in time.

The Automatic Ship Identification System (AIS) is developed from the enemy-identification device of the ship's aircraft. It cooperates with the Global Positioning System (GPS) to combine the ship's actual position, ship speed, changing heading rate and heading with the ship's name, call sign and draught,ship scale. The information above can be broadcasted in real-time.Therefore, through the processing and analysis of AIS data, the movement of ships near the port can be obtained, and the congestion degree evaluation of the port is further calculated.

2. Research methods

2.1 Research at home and abroad

The research on port congestion is still in a blank stage, and there are few studies and evaluations for causing port congestion. Cai Lijuan, Luo Bencheng, etc,through the comprehensive analysis of the internal and external conditions of the inland waterway, pointed out that the key reason for the congestion is the interaction of the ship, environment, management, crew and other factors, and accordingly make regularity on the channel congestion problem. Zhou Junhua and Chen Xianqiao, on the basis of in-depth study of Bode's modern accident causal chain theory, analyzed in detail the causes and causes of inland waterway congestion. Bayesian network models have unique advantages in the field of congestion warning in inland waterways. They also model Bayesian network model based on the inland river channel congestion warning system. Then relying on GIS technology, from the three aspects of the overall architecture design, technology development and workflow of the system, a design of the inland river channel congestion prediction and warning system under the Bayesian network model is proposed. Wang Shujun mainly analyzes the navigational norms of various types of vessels in the shrimp gate channel, and proposes five improvements to implement safety navigation regulations, increase cruising efforts, standardize navigation, further develop VTS functions, and strengthen publicity and education and joint law enforcement. It is recommended to reduce the traffic pressure of the navigation channel and improve the safety level of navigation. Liu Yuxiao analyzed the ship positioning data, established the channel congestion grading index model, and proposed the channel congestion identification method based on fuzzy comprehensive evaluation. Liu Sailong et al proposed the channel service level index and grade, combined it with the ship traffic flow theory to establish the inland waterway passage. Capacity calculation model; Liu Mingjun et al. analyze the factors affecting the passage capacity of the channel, select the correction coefficient, and establish a calculation model for the passage capacity of the ship based on the ship flow; Zhu Fuling draws on the Highway Capacity Manual (HCM) to establish urban road traffic. Congestion evaluation index system; Based on three-phase traffic theory, KNORR proposes communication detection and control congestion technology to make road traffic congestion obsolete and controllable; LAKAS and other based on Geocast protocol to manufacture vehicle-to-vehicle communication system, which can effectively determine road traffic conditions. The method for judging highway traffic congestion is mature and has a scientific judgment index system. Most of the channel traffic congestion judgments are based on the channel passing ability to determine the congestion grading principle. It is a manual judgment and an after-the-fact judgment. It lacks real-time performance, and it is difficult for the maritime department to process the channel traffic in time. Congestion problem. Jiang Mingyu and Zhao Renyu made research on judging whether the channel is congested. They proposed using cluster analysis to find a more congested channel. The method forms a ship cluster by clustering AIS ship data, compares the calculated value of the water density of the segment waters in the ship cluster with the actual optimal density, and judges accordingly.

However, the causes of port congestion in actual production are various, and it is difficult to quantify by using specific indicators. Therefore, it is necessary to conduct statistics and analysis on the inbound and outbound data of ships, and then obtain a relatively objective evaluation criteria for describing the results of congestion. There is also no uniform quantitative standard for the evaluation of road traffic conditions. The service level is usually used to describe the congestion status of the road. The definition of the service level generally includes the driving speed of the road user, the driving time, and the perception of the comfort level of the driving situation. Defined. Therefore, based on the evaluation of road congestion status, it is proposed to study the method of clustering ship data according to its congestion degree.

ISSN 2616-5783 Vol. 2, Issue 5: 16-30, DOI: 10.25236/AJHSS.020503

2.2 Cluster analysis

Cluster analysis, referred to as clustering, is an unsupervised learning in which it is not necessary to know the distribution of data sets in advance. Clustering is the process of dividing a set of data objects into subsets. Each subset is a cluster, and the data objects in the cluster are similar to each other, but not similar to the data objects in other clusters. For the two concepts of “similarity” and “dissimilarity”, the specific meanings must be clarified before similar dissimilar measures can be made. From the purpose of cluster analysis and the whole process, the cluster analysis mainly includes three steps. Before clustering, it is necessary to preprocess the collected data, including extracting the characteristics of the data or calculating the distance and similarity between the data objects. Depends on the specific issue. The clustering algorithm is executed, and the clustering result is explained and the clustering algorithm is adjusted by the clustering result feedback. The whole process is shown in the figure.

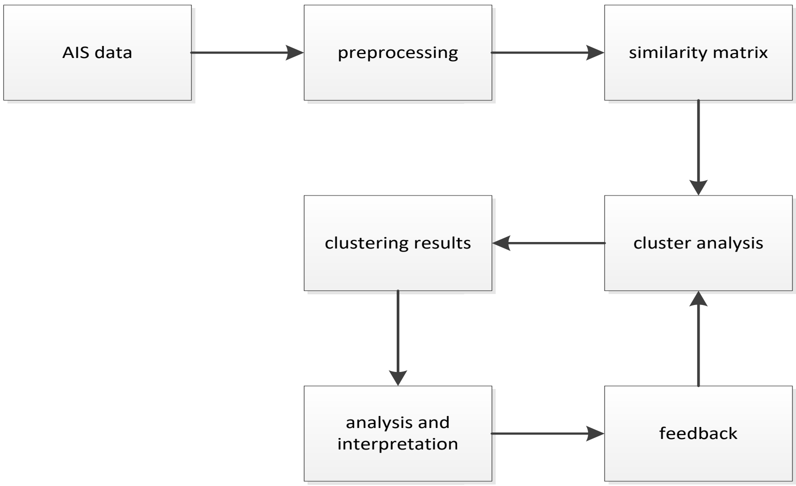

Figure. 1 Clustering algorithm flow chart

3. The port congestion evaluation method

3.1 AIS data processing

3.1.1 Port area division

Qingdao Port consists of four major port areas, namely Qingdao port, Huangdao Oil Port, Qianwan New Port Area and Dongjiakou Port Area. It has 15 docks and 72 berths. According to the 2016 GPCI (Global Port Congestion Index Report) report, the port congestion days of Qingdao Port are also much higher than those of other ports in China. As an important import and export port in China, the port congestion

ISSN 2616-5783 Vol. 2, Issue 5: 16-30, DOI: 10.25236/AJHSS.020503

degree is representative. Therefore, Qingdao Port is used as a research object to evaluate the degree of congestion in the port. First, the port area is delineated based on the AIS data near the port.

The original AIS data format is as follows:

Table 1 AIS data

mmsi	utc	status	lon	lat
212276000	1504307416	1	120911736	35797157
212276000	1504313357	1	120908615	35796768
212276000	1504341438	1	120911446	35798424
212276000	1504341617	1	120911575	35798531
212276000	1504358920	5	120210098	36011982
212276000	1504359100	5	120210098	36011856
212276000	1504359110	5	120210098	36011982
212276000	1504367920	5	120210098	36011616

The file is processed by python's pandas library, and the ship near Qingdao Port is found from the latitude and longitude data, and the polygon area of Qingdao Port is obtained. The ship information in the intercepted area is the research sample.

The AIS format obtained after screening is as follows:

Table 2 AIS data after processing

mmsi	utc	status
412379520	1504224938	0
412349146	1504262215	0
412349146	1504260984	0
412349146	1504261607	0
100900066	1504264809	0

The state of the ship is expressed as "status" in the data, and "status" is 0, which is 1, and 5 indicates that the ship is in travel, berth, mooring state, and "utc" is the time recorded for this ship.

3.1.2 AIS data screening

Due to the subjectivity of the port area and the lack of some AIS data, it is necessary to screen the AIS data to obtain AIS data that can be used for calculation, and reduce the error. The screening rules are:

(1) Delete data with unreasonable sailing speed;

(2) Screening out the ship that has berthing and anchoring behavior in Qingdao Port during the observation period;

ISSN 2616-5783 Vol. 2, Issue 5: 16-30, DOI: 10.25236/AJHSS.020503

(3) Sorting the ship mmsi in ascending order for the calculation of the ship's dwell time in the next step;

(4) Match the ship information in the ship file through mmsi, add teu (standard box), ship arrival time (converted to Beijing time), ship type, dwt (load capacity) and other information;

(5) According to the matched ship data, the ship data that is insensitive to the congestion degree of the port, such as the towed ship and the official ship, are excluded.

By further processing the data, AIS data in the following format is obtained:

Table 3 AIS data after screening and processing

shipnumber	mmsi	time
190	412469640	2.160556
189	255805615	1.098889
188	356413000	3.531667
187	538002864	0.964444
186	413354480	0.059722
184	416456000	4.704167
182	373061000	5.537222
181	413695770	5.017222
179	220164000	6.173333
178	412330570	0.031111

3.2 K-means clustering

The road traffic congestion index, also known as the traffic operation index, has a value range of 0 to 10, and each level is one level, corresponding to “unblocked”, “basic smooth”, “lightly congested”, “moderately congested”, “ Five levels of severe congestion, the higher the value, the more serious the traffic congestion. According to the road traffic situation, some cities have comprehensively reflected the conceptual index value of the smooth or congested road network, which is equivalent to digitizing the congestion situation. In the specific evaluation grading, the ship data of different congestion states have no definite grading limit. Therefore, this paper uses the clustering method to classify ship data with similar congestion levels into different categories and complete the classification of congestion degree.

Clustering is the process of dividing a collection of physical or abstract objects into multiple classes of similar objects. According to different calculation methods, clustering can be divided into partitioning methods (representative algorithms including K-MEANS algorithm, K-MEDOIDS algorithm, CLARANS algorithm), hierarchical methods (representing algorithms including BIRCH algorithm, CURE algorithm, CHAMELEON algorithm), density-based Methods (representative algorithms include STING algorithm, CLIQUE algorithm, WAVE-CLUSTER

ISSN 2616-5783 Vol. 2, Issue 5: 16-30, DOI: 10.25236/AJHSS.020503

algorithm), model-based methods (representing algorithms including statistical schemes and neural networks).

The core idea of K-means is to set the K initial cluster centers in the space, and calculate the similarity of the data objects separately. The data sentences with high similarity are new, and the values of the cluster centers are updated in turn. In the iterative process, when the displacement of the cluster center point is less than the set threshold, the algorithm ends. The function for calculating the similarity is the Euclidean distance formula, as follows:

$$J_n = \sum_{k=1}^{k} \|x_n - c_k\|^2 . \quad (1\text{-}1)$$

c_k: Cluster center point vector

x_n: Data object

In the K-means algorithm, the first to determine the number of clusters K, the second is the initial cluster center. According to the characteristics of port production, the cluster center is initially positioned 4, which reflects the degree of congestion and reduces the time complexity. The K-means clustering map made with Qingdao Port data is as follows:

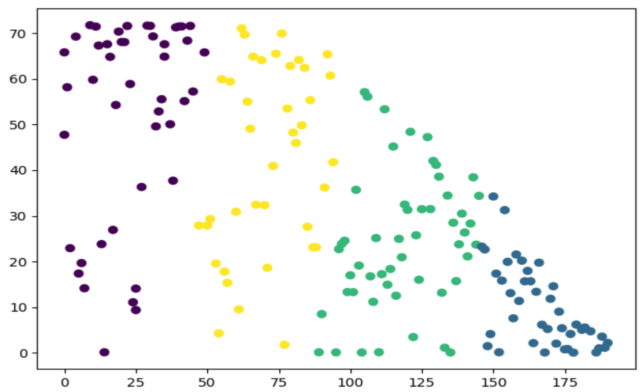

Figure. 2 K-means clustering diagram

The cluster center points are distributed as follows:

Table 4 Cluster Center Distribution Points

Number of ships to be operated in the port area	Ship waiting time
118.71153846	24.24432692
23.17777778	52.31682716
166.95238095	9.93531746
71.33333333	42.49928775

ISSN 2616-5783 Vol. 2, Issue 5: 16-30, DOI: 10.25236/AJHSS.020503

At this time, K-means classified the ship data into four categories, but it can be seen that the number of ships and waiting time are relatively scattered, but in actual port production, industry workers are often very sensitive to waiting time, and the results of K-means are relatively rough. The congestion status cannot be clearly defined. Therefore, the Gaussian fitting algorithm based on density clustering is used to classify the congestion level.

3.3 Density Clustering-Gaussian Fitting Algorithm

3.3.1 Calculating the optimal data width

The main purpose of the first step is to present the traffic elements in the form of histograms, using the Silverman Rule of Thumb algorithm, an elementary method based on statistical experience, using the Epanechnikov kernel function.

Taking the ship waiting time as an example,$X_1 \cdots X_n$ is taken as input, N is the total data amount, M is the mean value, S_d is the standard deviation:

$$s_{d=}\sqrt{\frac{1}{N\times[(X_1^2+\cdots+X_N^2)-2M(X_1+\cdots+X_N)+N\times M^2]}} \quad (1\text{-}2)$$

The optimal data width:

$$Bw = k \times S_d \quad (1\text{-}3)$$

$$k = -4.57 \times 10^{-6}N + 0.1867 \quad (1\text{-}4)$$

Through calculation, the optimal width of the ship waiting time is obtained by S_d, k , and the optimal width of the waiting time is calculated in the same way.

A histogram is obtained from the Gaussian fitting curve. The y-axis of the histogram is the number of times the data appears within the width, and the x-axis is the latency data divided by the optimal width. The Gaussian fitting rule is as follows, and the ship number-frequency histogram is made according to the Gaussian function. The ordinate of the t-th histogram is set to n(t), the total data amount is N, and σ is the standard deviation. Then, the t-th histogram Gaussian distribution function is as follows:

$$f(t|x) = \frac{n(t)}{N} \times \frac{1}{\sqrt{2\pi}\sigma} exp\left(-\frac{\left(x-\frac{2t-1}{2}\times Bw\right)^2}{2\sigma^2}\right) \quad (1\text{-}5)$$

$$F(x) = \sum_{t=1}^{k} f(t|x) \quad (1\text{-}6)$$

To get a smooth curve, let $\sigma = Bw;\ K = \lceil max(x)\ /\ Bw \rceil$. Through the calculation of the Gaussian fitting function, according to the optimal width obtained above, the ship number-frequency histogram can be obtained, and the ship waiting time-frequency histogram can be obtained in the same way.

3.3.2 Gaussian fitting to calculate the density curve

Based on the characteristics of shipping traffic, the maximum waiting time of the ship and the number of anchoring ships are obtained in the data processing stage, and the optimal width obtained in the previous step is divided.

ISSN 2616-5783 Vol. 2, Issue 5: 16-30, DOI: 10.25236/AJHSS.020503

$$F(\chi_i) = \sum_{t=1}^{k} f_t(\chi_i) \tag{1-7}$$

$$K = \lceil max(X) \ / \ Bw \rceil \tag{1-8}$$

$$i = 1,2,3 \ldots$$

After the discretization of the data, the $max(X) \ / \ Bw$ can represent a complex Gaussian fitting curve, which can reduce the computational complexity. The frequency table obtained by the above formula.

According to the formula (1-7), the Gaussian fitting function is discretized, and only the result of t being 1, 2, 3...n is calculated, and the frequency map of the number of ships after the dispersion is obtained. At this time, the Y-axis coordinate of the frequency map represents the percentage of the number of occurrences of a certain number of ships, and the ship's waiting time frequency histogram is obtained.

3.3.3 Extreme point optimization

The gradient between two points can be expressed as $Di = x_{i+1} - x_i$, when Di changes from positive to negative, its corresponding x_i is the maximum value of the function, and vice versa; The two endpoints are judged separately. If Di is negative, then x_i is the maximum value; otherwise, it is the minimum value.$Di - 2$ is negative, it is a minimum value; otherwise, it is a maximum value, and the classification position is set at a minimum value. Considering the actual situation, the minimum value at the end of the curve should be deleted. In summary, the minimum value of the maximum value can be obtained from the formula (1-7).

$$\mathrm{T(j)} = n_1\sqrt{(Y[t_b(j)] - Y[t_s(j)])^2 + [t_b(j) - t_s(j)]^2} + n_2\sqrt{(Y[t_b(j+1)] - Y[t_s(j)])^2 + [t_b(j+1) - t_s(j)]^2} \tag{1-9}$$

According to the number of classifications n, the weighted values are arranged in descending order, and the minimum value x coordinate corresponding to the first n weights is the classification boundary value. The calculated weighting value corresponding to the minimum number of ships.

4.The example verification

4.1 Qingdao Port AIS Data Preprocessing

The characteristic attributes of the data are obtained through data processing, including the waiting time of the Qingdao Port ship from July to September 2016 and the number of ships to be operated in the port area.

ISSN 2616-5783 Vol. 2, Issue 5: 16-30, DOI: 10.25236/AJHSS.020503

Table 5 July-September 2016 Qingdao Port Ship Data

number	mmsi	time
0	477127100	47.76
0	413303860	65.82778
1	413322480	58.18167
2	412427620	22.90806
4	412330130	69.27333
5	232006421	17.35139
6	413526540	19.64889
7	413668000	14.13028
9	413320420	71.77944
10	413444680	59.81111

The optimal data width is calculated by using the Silverman Rule of Thumb algorithm, and the optimal data width of the ship number and waiting time is obtained respectively, and the Epanechnikov kernel function is used. Calculate the optimal width of the data according to equations (1-2), (1-3), (1-4):

The histogram is made according to the optimal width, and the X-axis of the data is divided into several parts with the optimal width BW by the density function obtained by Gaussian fitting, and the Y-axis is the number of times the data appears in the interval, and the ship can be obtained. Quantity, wait time histogram. According to equations (1-5), (1-6), the histogram Gaussian distribution function is obtained, and because the characteristics of the traffic data are discretely distributed with time, the function is adjusted by the equation (1-7) in order to smooth the fitting curve.

First calculate the optimal width of the number of ships and waiting time.

Calculate the average number of ships available according to formulas (1-2), (1-3), (1-4), M=95.27624, k=0.185872, S_d=55.52287, Bw_{nu}=10.32, the same is true, the optimal width of waiting time Bw_{ti}=4.37.

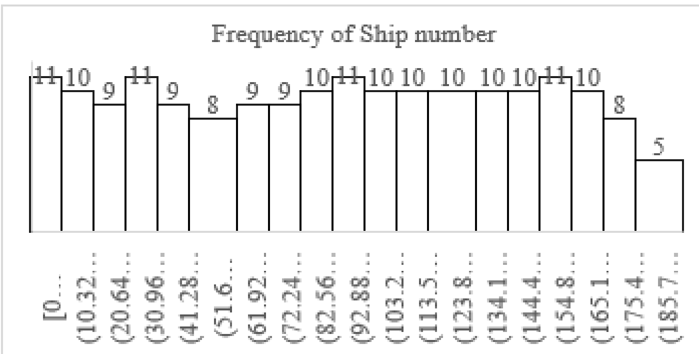

Figure. 3 Ship Frequency Table

ISSN 2616-5783 Vol. 2, Issue 5: 16-30, DOI: 10.25236/AJHSS.020503

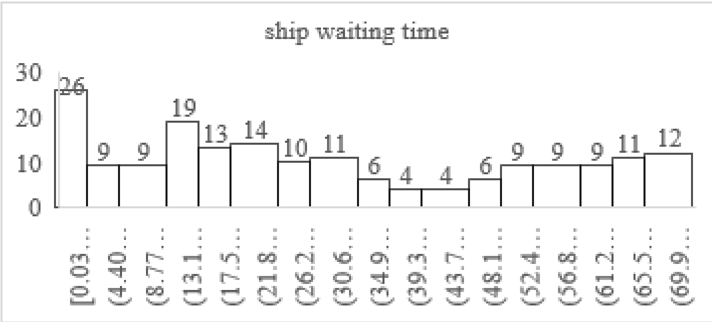

Figure. 4 Waiting timetable

4.2 Gaussian fitting

Gaussian fitting of the ship number histogram is performed by equations (1-5) and (1-6). $\sigma = Bw_{nu} = 10.32$,$k = \left\lceil \frac{\max(X)}{Bw_{nu}} \right\rceil = 19, N = 181$

The frequency table is as shown:

Table 6 Frequency Table

number	1	2	…	19
Interval	0-10.32	10.32-20.64	…	185.76-196.08
Frequency	11	10	…	5

After Gaussian fitting, to simplify the calculation, the calculation process is discretized and only calculated $x = 1, 2 \ldots 181$, F(x), After obtaining the result, the frequency chart of the ship number is drawn. Similarly, the waiting time frequency map can be obtained:

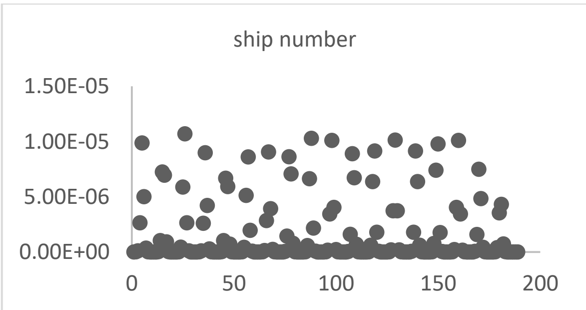

Figure. 5 Number of ships

ISSN 2616-5783 Vol. 2, Issue 5: 16-30, DOI: 10.25236/AJHSS.020503

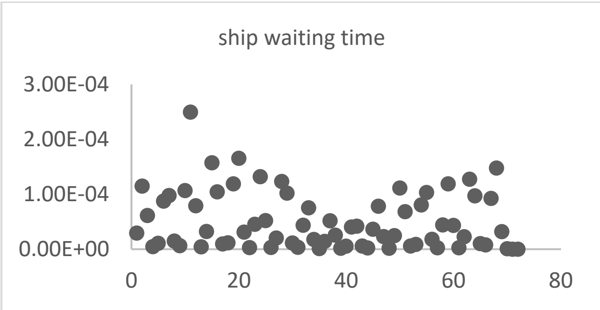

Figure. 6 Waiting time chart

4.3 Congestion degree division

According to the discrete derivative $D_i = x_{i+1} - x_i$ the minimum value of the number of ships 48, 112, 179 is obtained, and the minimum value weight is calculated according to the formula (1−9), $\mathrm{T}(48) = 4.36$, $\mathrm{T}(112) = 3.15, \mathrm{T}(179) = 1.32$, the weight is selected to be larger, that is, 48 and 112 are the demarcation points of the number of ships; similarly, the minimum value of waiting time is 2, 41, $\mathrm{T}(2) = 1.12, \mathrm{T}(41) = 2.08$, the weight is selected to be larger, so the demarcation point of the ship waiting time is 41 hours.

Figure. 7 Number of ships

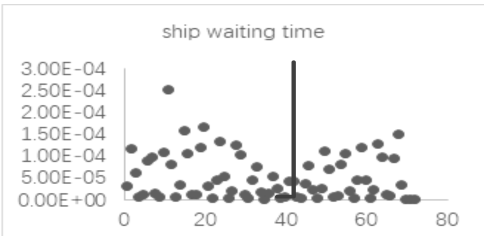

Figure. 8 Waiting time chart

ISSN 2616-5783 Vol. 2, Issue 5: 16-30, DOI: 10.25236/AJHSS.020503

From the above calculation results, the port congestion can be divided into the following: the number of ships in the port is bounded by 48, 112. When the number of ships is below 48, the port is relatively smooth. When the number reaches 48-112, the port becomes busy, and the number becomes When it becomes 112 or more, the port enters a relatively congested state; the ship's port time is bounded by 41 hours, the ship's port transit time is more than 41 hours, and the port's operating conditions become congested.

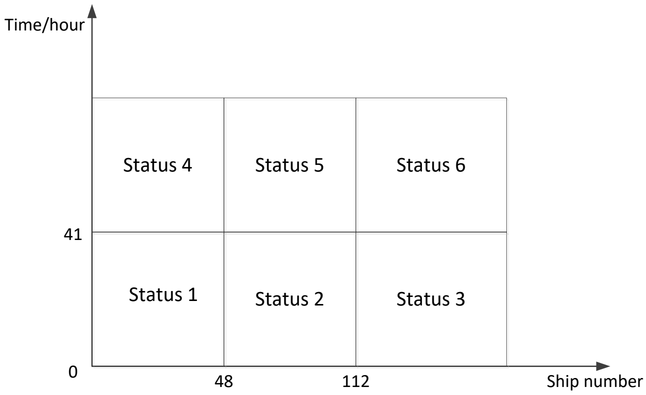

Figure. 9 Congestion status division diagram

5. Conclusion

The port production caused by port congestion is blocked, and the problems of overdue goods have become a major problem in various regions of various countries. At present, the evaluation criteria for port congestion are not uniform. Different ports have different geographical locations and operating conditions. It is not suitable for mechanical port congestion. Objective port congestion evaluation standards can help improve the accuracy of port congestion prediction. The distribution of port machinery has important guiding significance. In order to further improve the effect of port congestion assessment, this paper based on ship AIS data, GPCI data, etc., comprehensively use these data to carry out innovative research on multi-source data fusion and port congestion evaluation methods, and through the 2016 Qingdao Port ship access data The design method was verified.

In the research process of this paper, there are still some problems that need to be further studied and analyzed due to various reasons, including:

(1) The inbound and outbound port of the ship has certain randomness, and the distribution is difficult to describe with a unified model. At the same time, the weather, policy and other factors also have a great impact on the congestion of the port. The ship waiting time and the ship to be operated the quantity can only be objectively described as the current degree of congestion in the port, and does not

ISSN 2616-5783 Vol. 2, Issue 5: 16-30, DOI: 10.25236/AJHSS.020503

have the ability to predict and multi-factor analysis. Therefore, this aspect can be used as a new research area.

(2) In this paper, the classification of congestion is only based on the 2016 Qingdao Port data, and the results are lower. At the same time, the factors affecting congestion are considered less. The selection of relatively representative data from different ports and different time periods for analysis may result in more accurate, objective and comprehensive results.

(3) There is a phenomenon of “predictive or intervention” in the pre-processing of big data, that is, the phenomenon that the forecast will affect the future after the release, and the situation that the original congestion will become very congested. How to command the ship to choose the port through the calculation results without "side effects" is a problem to be considered.

References

[1] Reasons for the congestion of container terminal yard and countermeasures [J]. Wang Weibin. China Water Transport. 2019 (04)

[2] Xu Jianhua.Exploration of the Causes of Congestion in Shanghai Port Container Terminal and Suggestions [J]. Chinese ship inspection, 2017 (07): 16-18+108-109.

[3] Wang Furong, Ding Yi, Lin Guolong. Optimization of Coastal Shipping Network Considering Port Congestion Factors [J]. Computer Engineering and Applications, 2017, 53 (10): 218-224.

[4] Lin Guolong, Ma Libo, Zhang Chenyan, He Hongdi. Multifractal Study on Congestion Index of Ports in China and Brazil [J]. Journal of Chongqing Jiaotong University (Natural Science Edition), 2016, 35 (06): 148-152.

[5] Zhijia Tan, Wan Li, Xiaoning Zhang, Hai Yang. Service charge and capacity selection of an inland river port with location-dependent shipping cost and service congestion [J]. Transportation Research Part E, 2015, 76.

[6] Matteo Balliauw, Peter M. Kort, Hilde Meersman, Eddy Van de Voorde, Thierry Vanelslander. The case of public and private ports with two actors: capacity investment decisions under congestion and uncertainty [J]. Case Studies on Transport Policy, 2019.

[7] Wan, Zhang, Li. Port competition with accessibility and congestion: a theoretical framework and literature review on empirical studies [J]. Maritime Policy & Management, 2018, 45 (2).

[8] Lin Tianyi, Lu Chunxia.A Model of Hub-Ship Maritime Network Hub Port Selection Based on Congestion Control [J]. Journal of Shanghai Maritime University, 2013, 34 (04): 59-66.

[9] Gao Weibo, Du Taili, Zhang Yong, Huang Lianzhong.Application of Clustering Algorithm in Ship Energy Efficiency Data Mining [J]. Journal of Wuhan University of Technology (Transportation Science and Engineering Edition), 2019, 43 (02): 286-290.

ISSN 2616-5783 Vol. 2, Issue 5: 16-30, DOI: 10.25236/AJHSS.020503

[10] Zhang Yonglai, Zhou Yaojian. A Survey of Clustering Algorithms [J/OL]. Computer application: 1-14 [2019-05-22]. http://kns.cnki.net/kcms/detail/51.1307.TP.20190415.1412.004.html.

[11] Li Zhifeng, Zhang Wei. Analysis and Evaluation of Clustering Analysis Algorithm [J]. Electronic Technology and Software Engineering, 2019 (07): 157.

[12] Huang Bingsen, Chen Yuzhong, Guo Kun. A Community Discovery Algorithm Based on Improved Density Peak Clustering [J]. Microcomputer system, 2019, 40 (04): 782-786.

[13] Wang Feiyu, Hu Zhixiang, Huang Wei. Modal Parameter Identification Based on Density Peak Clustering Algorithm [J]. Vibration and shock, 2019, 38 (02): 172-178.

[14] Sun Zhijun, Xue Lei, Xu Yangming, Wang Zheng. A Review of Deep Learning Research [J]. Application Research of Computers, 2012, 29 (08): 2806-2810.

[15] Chen Shuwen. Port congestion problems endanger the global shipping system [J]. China Ocean Shipping, 2007 (03): 24-25.