Journal of Marine Science and Technology

Volume 26 | Issue 1 Article 8

AUTOMATIC RECOGNITION OF MARINE TRAFFIC FLOW REGIONS BASED ON KERNEL DENSITY ESTIMATION

Wei-Feng Li
Navigation College, Dalian Maritime University, Dalian, P.R. China., dmuliweifeng@hotmail.com

Bin Mei
Navigation College, Dalian Maritime University, Dalian, P.R. China.

Guo-You Shi
Navigation College, Dalian Maritime University, Dalian, P.R. China.

AUTOMATIC RECOGNITION OF MARINE TRAFFIC FLOW REGIONS BASED ON KERNEL DENSITY ESTIMATION

Acknowledgements

This work was financially supported under Fundamental Research Funds for the Central Universities grant number 3132016002 and Natural Science Foundation of Liaoning Province of China grant number 20170540090.

Journal of Marine Science and Technology, Vol. 26, No. 1, pp. 84-91 (2018)
DOI: 10.6119/JMST.2018.02_(1).0014

AUTOMATIC RECOGNITION OF MARINE TRAFFIC FLOW REGIONS BASED ON KERNEL DENSITY ESTIMATION

Wei-Feng Li, Bin Mei, and Guo-You Shi

Key words: marine traffic flow, trajectory, kernel density estimation, clustering, automatic recognition.

ABSTRACT

The automatic recognition of traffic flow regions can provide decision support for ships' automatic route design and route planning. This study analyzes the characteristics of ships' trajectory structures and builds a course and distance model. Pearson correlation coefficients are used for measuring the similarities of the models and clustering trajectories, and kernel density estimation is used for estimating the probability density of clustered trajectories. An automatic recognition algorithm for traffic flow regions is proposed. This study examines ships' automatic identification system data in Laotieshan channel, China. The traffic separation scheme regions and traffic intersectional regions are recognized automatically, and the obtained results show good agreement with actual circumstances, thus verifying the applicability of the algorithm.

I. INTRODUCTION

Marine traffic refers to the combination of ships' motions and behaviors in designated water areas. An analysis of marine traffic characteristics provides decision-support information for ships' automatic routing design and route planning (Qi, 1991). In marine transportation, ships' traffic flow represents the summation of vehicles, such as ships, that are in continuous motion with certain fluid characteristics (Liu et al., 2014). Position, direction, width, density, and speed are the five basic characteristics of ships' traffic flow (Wu and Zhu, 2004).

According to the Safety of Life at Sea Convention, from 2002 onward, international voyage ships weighing over 300 gross tons and domestic cargo and passenger ships weighing over 500 tons have been required to install automatic identification system (AIS) equipment. AIS equipment is also being installed in an increasing number of fishing vessels. Ships' static, dynamic, and voyage information is transmitted automatically and continuously through such systems, which can also receive this information from surrounding ships and exchange information with shore-based AIS stations (IMO, 2005). Owing to the establishment of AIS base station networks in various countries and the emergence of satellite-borne AIS groups, these systems have become a nearly real-time source of marine traffic information (Pan et al., 2010). By using AIS data collected over various time periods such as a month or a year and performing various calculations, ships' trajectory distributions can be determined. From these results, traffic flow positions and distributions can be identified, marine traffic characteristics can be determined, and ships' behaviors can be analyzed to study the distribution of ships' speeds and courses (Tang et al., 2015). Considerable information about shipping traffic is hidden within AIS trajectory big data (Zhu et al., 2012). Through analysis of such data, hitherto unknown traffic information can potentially be extracted, and marine traffic flow regions can be recognized automatically, thus providing technical support and a decision-making basis for automatic routing design and route planning.

Clustering refers to the process of dividing sets of physical or abstract objects into various categories with similar objects. Cluster analysis separates groups of data objects into groups with higher and lower similarity (Chen et al., 2007). Trajectory clustering seeks out trajectories with the same motion mode and measures the similarity among them by analyzing their characteristics and the inner motion mode. Trajectories with high similarity can be placed in the same category. Common clustering algorithms such as K-MEANS, BIRCH, and DBSCAN mostly cluster sampling positions in a trajectory; however, characteristics and motion trends cannot be determined from this overall perspective (Lee et al., 2007).

At present, various types of trajectory similarity measurement methods are available: methods based on the Euclidean distance can only be used for measuring the similarity of trajectories in ships of the same length, whereas methods based on longest common subsequence, dynamic time warping, Hausdorff distance, and one-way distance can be used for measuring the similarity of trajectories of ships with different lengths. Yuan et al. (2011) proposed a trajectory clustering algorithm based on structural

Paper submitted 05/03/17; accepted 11/09/17. Author for correspondence: Wei-Feng Li (e-mail address: dmuliweifeng@hotmail.com).
Navigation College, Dalian Maritime University, Dalian, P.R. China.

similarity. However, all trajectory measurement methods based on similarities have disadvantages in that they overemphasize the algorithm efficiency; that is, they neglect the characteristics of the trajectory and cannot effectively manage abnormal situations that arise because of data acquisition.

Pearson correlation coefficients (PCCs) can be used to quantitatively measure the similarity of two variables X and Y. They have been used by Zhang (2015) to quantitatively analyze relationships among various living species, Jiao (2009) to analyze data on land use in 40 counties in Chongqing, China, and Yang (2015) to analyze pollution indexes in Beijing. Wang (2012) introduced PCCs in multicomponent seismic exploration and used them to detect the rotation angle and then separate fast and slow waves. Compared with the traditional cross-correlation method, the PCC method has higher accuracy, noise immunity, and computational efficiency.

For extracting trajectory characteristics, this study developed a course and distance model, used PCCs for measuring the trajectories' similarity, clustered ships' trajectories, applied kernel density estimation (KDE) to predict trajectory clusters' boundaries, and finally proposed a systematic approach for the automatic recognition of traffic separation schemes and traffic flow intersectional regions.

The proposed model and algorithm afford numerous advantages, such as a track model characterized by fewer parameters, a consistent structure, and ease and speed of manipulation. The PCC-based clustering algorithm is simple and adaptable, and the use of KDE helps determine the boundaries of every discrete point set. This study examined traffic flow in the Laotieshan channel, China. Two weeks of sample data were used to build the track model, and a clustering algorithm was used to automatically identify the main track clusters in these waters. A comparison between the obtained results and the currently used traffic lane boundaries in the channel showed good agreement, indicating that the boundary estimation algorithm that tracks clusters can accurately reflect the geographical position of boundaries; this method is therefore highly applicable.

II. SHIP TRAJECTORY MODEL

Traffic speed and traffic density are two key parameters for measuring traffic flow. Each traffic flow trajectory can be described through three parameters: course, distance, and position. This study determined the relationship between course and distance to build a relevant model.

1. Data Preprocessing

In this study, AIS data for the Bohai Straits was collected and recorded between September 14 and 27, 2015, from a single AIS ship station, and was then decoded using selected data in the longitude range of 120°52′.00 E to 122°30′.00 E and latitude range of 37°30′.00 N to 39°00′.00 N. Some AIS information sent by ships beyond the AIS receiving range of coast stations or due to AIS equipment failure resulted in mistakes in the trajectory and potentially the omission of some waypoints (Graveson, 2004). These factors affected the classification and recognition of a single ship's manipulation behavior. Thus, AIS data cleaning, interpolation, and transfer were necessary to obtain data sets that satisfied the requirements. Fig. 1 shows the preprocessing steps for AIS data.

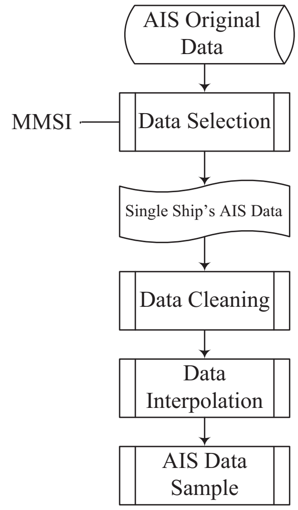


Fig. 1. AIS data preprocessing.

As Maritime Mobile Service Identity (MMSI) has been implemented worldwide, the trajectory data from the same MMSI must be implemented in the following situations:

(1) As the maximum update interval of a ship's AIS dynamic information is 3 min, when the interval between these ships' position information exceeds 3 min, a segmenting process is needed.
(2) As the length of Laotieshan's traffic separation scheme is 9 nm, these should be segmented if the distances between ships' positions exceed 9 nm.

2. Calculation of Trajectory

When a ship proceeds along a constant course, its trajectory is a rhumb line. It crosses all meridians of longitudes at the same angle, that is, the path has constant bearing as measured relative to the true or magnetic north. A ship will typically take a course represented by a rhumb line between two waypoints. Therefore, connecting all position points of the ships' AIS data forms rhumb lines, as shown in Fig. 2. The trajectory segments O_1O_2, O_2O_3, and O_3O_4 are rhumb lines. The course and distance of the rhumb line between two waypoints can be calculated by Eqs. (1) and (2), respectively.

The course A between every pair of adjacent waypoints in the trajectory can be calculated as follows:

$$\tan A_1 = \frac{\sin(\lambda_{o2} - \lambda_{o1})}{\cos\varphi_{o1}\tan\varphi_{o2} - \sin\varphi_{o1}\cos(\lambda_{o2} - \lambda_{o1})} \tag{1}$$

Here, λ and φ are the longitude and latitude of the waypoint, respectively.

Table 1. Detailed steps in model for course and distance.

Algorithm: Model of Course and Distance	
Input Parameter: Points $P_1, P_2, \ldots, P_n$	% One Point (φ, λ);
Output Parameter: Array[360]	% Size 360 One Unit [A, S];
Array[·] = [0, 0];	% initial Array state: zero
for i = 1:1:n-1	
[A_i, S_i] = GetCourseAndDistance (P_i, P_{i+1});	% formula (1) & fomula (2);
A_i = int(A_i)	% Get the integer number of A_i;
Array[A_i] = [A_i, S_i];	
end	
return Array;	

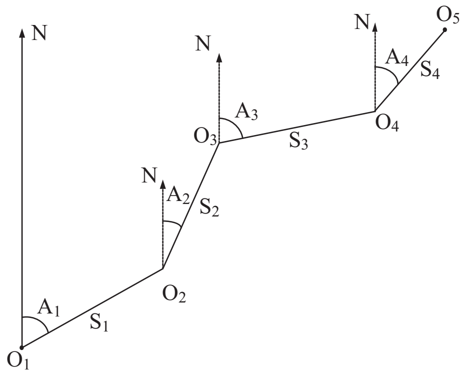


Fig. 2. Example of trajectory.

The distance S between two adjacent waypoints in the trajectory can be calculated as follows:

$$\cos S = \sin\varphi_{o1}\sin\varphi_{o2} + \cos\varphi_{o1}\cos\varphi_{o1}\cos(\lambda_{o2} - \lambda_{o1}) \quad (2)$$

3. Construction of Course and Distance Model

The course and distance model shows the statistical relationship between the course and distance of a single trajectory. This model consists of 360 data points, the abscissa of which is 0, 1, 2, …, 358, 359; each data point describes the distance component of the trajectory in the course.

For example, if a trajectory contains n points, Eqs. (1) and (2) can be used to calculate each segment's course and distance (A_i, S_i), $i \in [1, 2 \ldots, n]$. The course and distance model consists of 360 points. Assume that point P is the ith point and that its coordinate is (A_i, S_i), $i \in [1, 2 \ldots, n]$; other points' coordinates can be calculated similarly using Eq. (3). Table 1 shows the algorithm.

$$P_{\lfloor A_i \rfloor} = P_{\lfloor A_i \rfloor} + \left(A_i, S_i\right) \quad (3)$$

Here, $[A]$ is the round down for course A. $[A_i] \in [0, 1, 2 \ldots, 359]$, $i \in [1, 2 \ldots, n]$.

The model of course and distance has the following advantages:

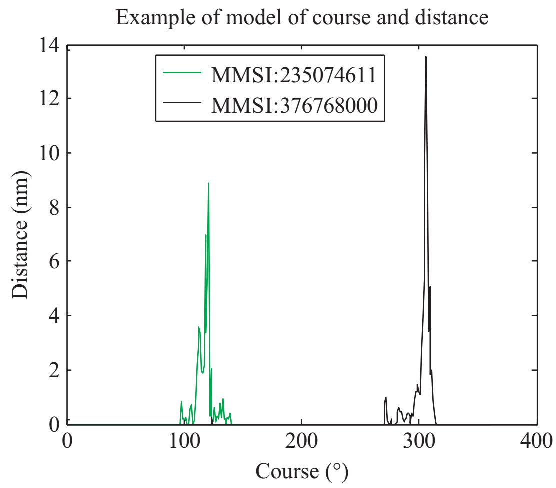


Fig. 3. Example of the course and distance model.

(1) It does not compress the AIS data, thus maintaining the integrity of the data and the topological property of the trajectory points.
(2) The model has a simple structure, and just 360 data points are needed to describe any trajectory.
(3) The data points of every two trajectories are equal, being 360, and thus, there is no need to perform interpolation or normalization processing for trajectory big data; instead, similarity measurements can be performed directly.

Fig. 3 shows the model of course and distance for two trajectories in the Bohai Strait, China.

III. SHIP TRAJECTORY CLUSTERING

After course and distance models for every trajectory are built, the similarity between models of different trajectories and clusters should be measured. This study used PCCs to measure the similarities in the models in terms of the course and distance of every trajectory and as a parameter for clustering.

1. Pearson Correlation Coefficient

The PCC quantitatively measures the similarity of variables X and Y in the range [−1, 1]. X has an entirely positive linear

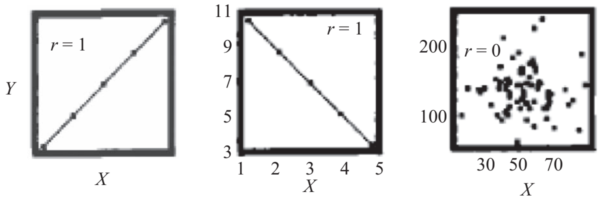


Fig. 4. PCC values of +1, –1, and 0.

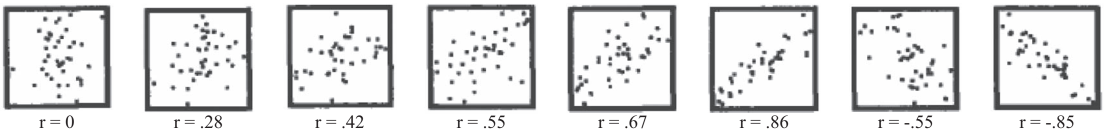


Fig. 5. Other PCC values.

correlation with Y if the correlation coefficient is +1, as shown in Fig. 4(a). X has an entirely negative linear correlation with Y if the correlation coefficient is −1, as shown in Fig.e 4(b). X and Y have no linear dependency if the correlation coefficient is 0, as shown in Fig. 4(c). However, in most cases, the correlation coefficient obtained using Eq. (4) is neither +1 nor −1; Fig. 5 shows some examples of the linear dependency of X and Y for other correlation coefficient values.

$$
\begin{aligned}
W_{12} &= \frac{\mathrm{COV}(X_1, X_2)}{\sigma_{X_1}, \sigma_{X_2}} \\
&= \frac{E(X_1, X_2) - E(X_1)E(X_2)}{\sqrt{E(X_1^{\,2}) - E^2(X_1)}\sqrt{E(X_2^{\,2}) - E^2(X_2)}} \qquad (4) \\
&\approx \frac{\sum X_1 X_2 - \dfrac{\sum X_1 \sum X_2}{N}}{\sqrt{(\sum X_1^{\,2} - \dfrac{E(X_1)^2}{\mathrm{N}})(\sum X_2^{\,2} - \dfrac{E(X_2)^2}{\mathrm{N}})}}
\end{aligned}
$$

In Eq. (4), X and Y are the set of x and y coordinates of all points, respectively; σ is the variance, COV is the covariance, E(·) is the standard deviation, and N is the number of samples.

The necessary and sufficient condition for the existence of PCCs is that the standard deviation of the variable is not zero. Considering that the standard deviation of the course and distance model can be zero only when a ship tracks a circle or when it is still, and because of the existence of inaccuracy in the locator measurement and in the numerical precision of computers, the probability of the correlation coefficient not existing is extremely low. When the geographic position or course are not considered, similarities in trajectory are mainly determined by measurement of structural similarity (Yuan et al., 2011).

2. Similarity Measurement

Three groups of trajectories can be compared. If the structures of trajectories are similar and the courses are close, the correlation coefficient of the course and distance model is greater than 0. If the structures of trajectories are similar and the courses differ from each other by 180°, the correlation coefficient is less than 0. If the structures of trajectories and the courses are both different, the correlation coefficient is again less than 0. This is because the necessary and sufficient condition for the existence of PCCs is that the standard deviation of the variable is not zero.

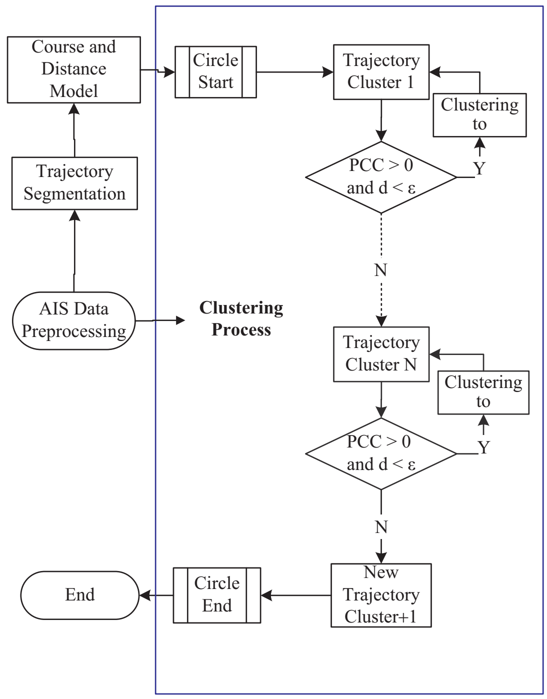
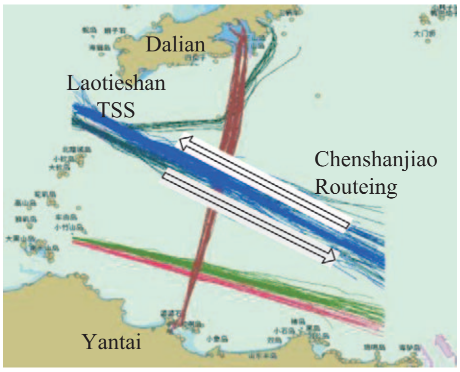


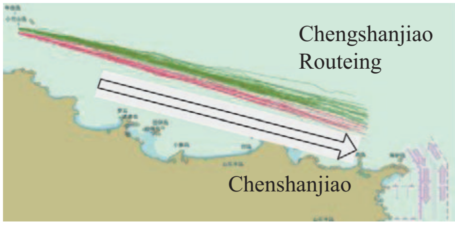
Fig. 6. Clustering process.

3. Examples of Clustering

With regard to the trajectory, not only its structure and course but also its geographic position should be considered. If the structural similarity between two trajectories is high but they are far away from each other, it would be meaningless to study them. The course and distance model only considers the trajectory's structure and course without considering the impact



(a)



(b)

Fig. 7. Clustering result.

of geographical position while clustering, because it measures the distances between the geometric centers of trajectories with the same structures and courses in different geographic areas. These distances between geometric centers are related to the traffic flow width. In this paper, the threshold for the distances of geometric centers is 20 nautical miles.

After a course and distance models are built and the similarities are measured, ships' trajectories should be clustered. Fig. 6 shows a flowchart of the process used for clustering ship trajectories in the Bohai Strait, and Fig. 7 shows the results.

IV. ESTIMATION OF TRAFFIC FLOW REGION BOUNDARIES

Assorted trajectory clusters obtained through clustering represent different traffic flows and different ships' behavior. Ships' trajectory points are a discrete point set distributed in Mercator coordinates signifying two-dimensional space. The point set distributes in a close and orderly manner. In probability statistics, it is a basic challenge for a given point set to solve the distribution probability density function. Picard (2000) proposed a density estimation approach by summing up the loose boundaries of a point set. In this study, KDE was used to estimate the distribution of a trajectory cluster in two-dimensional space, thereby obtaining the boundaries of trajectory clusters and recognizing intersectional regions of traffic flow and traffic lanes.

1. Kernel Density Estimation

In statistics, KDE is a nonparametric method used for estimating the probability density function of a random variable (Li et al., 2004). KDE is a fundamental data smoothing problem in which inferences about a population are made based on a finite data sample (Parzen, 1962). In some fields, such as signal processing and econometrics, it is also called the Parzen-Rosenblatt window method. The KDE formula of the density function of a point set $z_j (j \in [1, 2 \ldots, N])$ is given as follows:

$$\hat{f}(x) = \frac{1}{N} k\left(\frac{x - x_j}{h}\right) \frac{1}{h} \tag{5}$$

In Eq. (5), $k(\cdot)$ is the kernel function, and the most useful kernel function is the Gauss kernel function $k(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2}$, and h is the window width or smoothing function. Kroese et al. (2014) presented a linear diffusion kernel density estimator that solves problems such as the deviation of boundary estimation in the Gaussian kernel function and excessive smoothness of density between point sets.

When Eq. (5) was set to be Gaussian KDE, Eq. (6) was obtained:

$$\hat{f}(x; t) = \frac{1}{N} \sum \frac{1}{\sqrt{2\pi t}} e^{-\frac{1}{2}\frac{(x - x_i)^2}{t}} \overset{\text{def}}{=} \frac{1}{N} \sum_{i=1}^{N} \varphi(x; x_i; t) \tag{6}$$

When $\lim_{x \to \infty} \hat{f}(x; t) = 0$, $\hat{f}(x; t) = \Delta(x)$, and Eq. (7) is obtained:

$$\frac{\partial}{\partial x} \hat{f}(x; t) = \frac{1}{2} \frac{\partial^2}{\partial x^2} \hat{f}(x; t), t > 0, x \in R \tag{7}$$

Eq. (7) satisfies the heat equation, and $\Delta(x)$ is the empirical density formula.

From Eq. (7), $\hat{f}(x; t) = \Delta(x)$ is the heat equation based on the initial Neumann boundary condition $\frac{\partial}{\partial x} \hat{f}(x; t)\Big|_{x=1} = \frac{\partial}{\partial x} \hat{f}(x; t)\Big|_{x=1} = 0$ with the defining field of $[0, 1]$.

Set

$$\theta(x; x_i; t) = \sum_{k=-\infty}^{\infty} \varphi(x; 2k + x_j; t) + \varphi(x; 2k - x_j; t) \tag{8}$$

Then,

$$\hat{f}(x; t) = \frac{1}{N} \sum_{k=-\infty}^{\infty} \theta(x; x_j; t), x \in [0, 1] \tag{9}$$

If h (window width) in Eq. (5) is sufficiently large,

$$\theta(x; x_i; t) = 1 + \sum_{k=1}^{\infty} e^{-\frac{1}{2}k^2\pi^2 t/2} \cos(k\pi x) \cos(k\pi x_j) \tag{10}$$

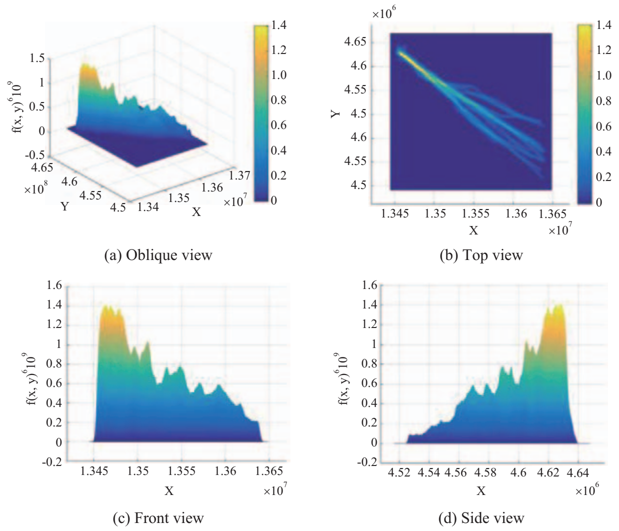


Fig. 8. KDE results of trajectory cluster.

When t $\to \infty$ and n is sufficiently large,

$$\hat{f}(x;t) \approx 2\sum_{k=1}^{n-1}\left(\sum_{i=0}^{n-1}\cos\left(k\pi\frac{i}{n}\right)\hat{f}_i\frac{1}{n}\right)e^{-k^2\pi^2 t/2}\cos(k\pi x) \quad (11)$$

where $\hat{f}_i$ is the number in the domain $\left(\frac{i}{n},\frac{i}{n}\right)$, $i = 0, \ldots, n-1$.

The heat equation of the kernel density function in two-dimensional space is

$$\frac{\partial}{\partial x}\hat{f}(Z;t) = \frac{1}{2}\left(\frac{\partial^2}{\partial {z_1}^2}\hat{f}(Z;t) + \frac{\partial^2}{\partial {z_2}^2}\hat{f}(Z;t)\right) \quad (12)$$
$$\forall t > 0,\ Z \in R,\ \hat{f}(x;t) = \Delta(x),\ Z = (z_1, z_2)$$

2. Example of Boundary Estimation

KDE can be applied to determine the closeness of trajectory points and reflect the traffic flow density through the numerical probability density (Rosenblatt, 1956). As shown in Fig. 8, in the Mercator coordinate system, the X–Y axis indicates the positions of track points and the Z axis indicates the probability density function $f(x, y)$. The value of $f(x, y)$ represents the density of traffic flow. As the update interval of position information in AIS data is changeable, it is necessary to interpolate AIS data equidistantly before estimating the kernel density.

To verify the KDE algorithm, this study uses the traffic flow density data from Fig. 9 to estimate the boundary of trajectory clusters. It is easy to obtain the probability density boundary by setting the probability density $\alpha = 0.05$, as shown in Fig. 9. The boundaries contain the complete trajectory, and they accurately estimate the traffic flow region trajectory points described by the trajectory cluster.

Table 2. Analysis of probability density contours.

	Match Probability Density of the Channel Length	Max Density	Percen-tage
North	1.59*10e-9	1.83*10e-9	86.9
South	1.22*10e-9	1.41*10e-9	86.5

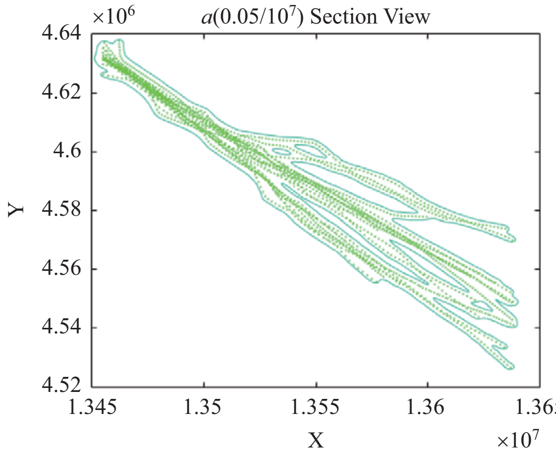


Fig. 9. Boundary of trajectory cluster.

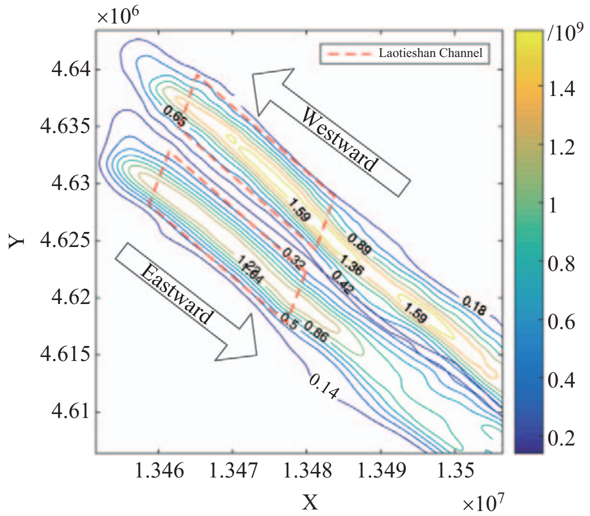


Fig. 10. Contours of probability density.

3. Recognition of Traffic Lane Region

The automatic recognition of traffic separation scheme regions can provide technical support for auto-routing designs and route planning. This study examines ships' traffic flow data in the Laotieshan Traffic Separation Scheme (TSS) and determines the probability density contours. Fig. 10 shows the estimated width and boundary of the Laotieshan TSS; it shows good agreement with the actual boundary and has high precision. As shown in Table 2, the width of the traffic lane is the estimated boundary corresponding to the Min density, and the length is the estimated boundary corresponding to 86% of the Max density.

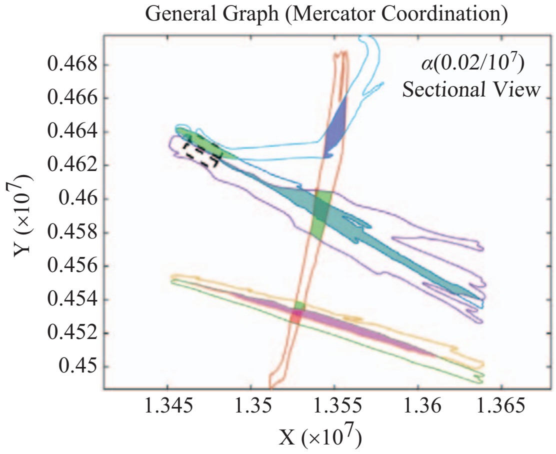


Fig. 11. Intersection of traffic flow (a).

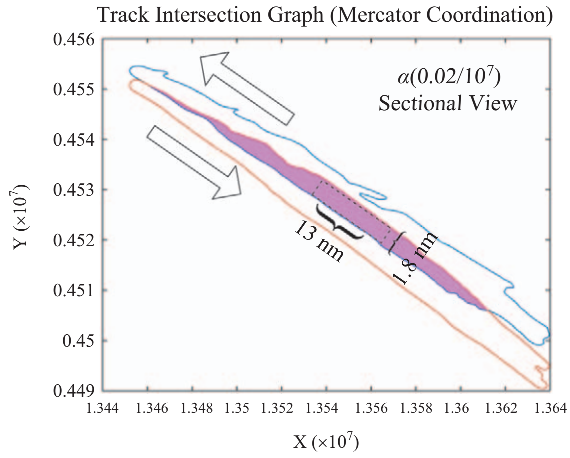


Fig. 12. Intersection of traffic flow (b).

4. Recognition of Traffic Flow Intersection Regions

Different trajectory clusters reflect different ships' behaviors, and different ships' behaviors produce traffic flow with different characteristics. Thus, the automatic recognition of traffic flow intersection regions is helpful for marine traffic management. Here, we choose the sectional view of 0.02×10^{-9} density (= 2% of max($f(x, y)$)), as shown in Fig. 11.

The recognition of traffic flow intersection regions can provide technical support and suggestions for route planning. A separation zone can be set up in intersection regions with opposite traffic flows, and when the impact of the crossing traffic flow shown in Fig. 12 is considered, the length and width of the separation zone can be set as 13 nm and 1.8 nm, respectively.

Geographic position of water in separation zone:

North boundary:

37°50.80 N, 121°38.00 E
37°47.00 N, 121°54.00 E

South boundary:

37°49.50 N, 121°35.08 E
37°45.00 N, 121°53.00 E

V. CONCLUSION

By using AIS data and data mining technology, this study proposes a new approach for the automatic recognition of marine traffic flow regions. In comparison with the traditional data mining and clustering method, the proposed traffic model affords advantages such as few parameters, minimal data, and convenient calculations. The automatic recognition of traffic flow is highly applicable and can provide technical support for automatic route design and route planning. Additionally, it can help support supervisory and management departments with decision making. Upcoming research will be conducted on the key technologies for traffic efficiency of KDE in traffic separation schemes and the estimation of the coverage and communication efficiency of AIS ship stations.

ACKNOWLEDGEMENTS

This work was financially supported under Fundamental Research Funds for the Central Universities grant number 3132016002 and Natural Science Foundation of Liaoning Province of China grant number 20170540090.

REFERENCES

Chen, J. D., X. F. Meng and C. F. Lai (2007). Clustering objects in a road network. Journal of Software 18(2), 332-344.

Graveson, A. (2004), AIS-An Inexact Science, Journal of Navigation, 57, 339-343.

Jiao, Q. D., Q. Y. Yang and Y. B. Feng (2009). Research on land-use division of Chongqing based on pearson hierarchical clustering. Journal of Southwest University 31(6), 173-178.

Kroese, D. P., C. Joshua and C. Chan (2014). Statistical Modeling and Computation: Joint Distributions, Springer, New York.

Kroese, D. P., T. Taimre and Z. I. Botev (2011). Handbook of Monte-Carlo methods. John Wiley & Sons, New York. ISBN: 978-0-470-17793-8.

Lee, J. G., J. W. Han and K. Y. Whang (2007). Trajectory Clustering: A Partitionand-group Framework. Proceeding of the 2007 ACM SIGMOD International Conference on Management of Data. Beijing, China, 593-604.

Li, C. H., Z. H. Sun, G. Chen and Y. Hu (2004). Kernel Density Estimation and Its Application to Clustering Algorithm Construction, Journal of Computer Research & Development 41(10), 1712-1719.

Liu, Z., J. X. Liu, F. Zhou and Z. W. Guo (2014). Behavior Characteristics of Vessel Traffic Flow and its Realization in Marine Traffic Organization, Journal of Dalian Maritime University 40(2), 22-26.

Pan, J. C., Z. P. Shao and Q. S. Jiang (2010). Application of Data Mining Technology in Analysis of Marine Traffic Characteristics, Navigation of China, 33(2), 60-62.

Parzen, E. (1962). On Estimation of a Probability Density Function and Mode, The Annals of Mathematical Statistics, 33(3), 1065.

Picard, N. and A. Bar-Hen (2000). Estimation of the envelope of a point set with loose boundaries, Applied Mathematics Letters 13(07), 13-18.

Qi, C. X. (1991). Ship Transportation Safety at inland waters, Dalian: Dalian Maritime University, 48-50.

Rosenblatt, M. (1956). Remarks on Some Nonparametric Estimates of a Density Function, The Annals of Mathematical Statistics 27 (3), 832.

Tang, C. B., B. L. Zhong, Q. R. Tang (2015). Fulfillment on Diagrams of Position Distribution for Ship Speed Based on AIS, Journal of Guangzhou Maritime Institute 23(1), 5-8.

Wang, K., X. Feng and C. Liu (2012). Wave Filed Separation of Fast-slow Shear Waves by Pearson Correlation Coefficient Method, Global Geology

31(2), 371-376.
Wu, Z. L. and J. Zhu (2004). Marine Traffic Engineering, Dalian: Dalian Maritime University, 57-58.
Yang, N., D. G. Ji and S. J. Li (2015). The Application of Pearson Correlational Analysis Method in Air Quality Analysis of Being- Tianjin-Hebei Region. Agricultural Science & Technology 16(3), 590-592.
Yuan G., S. X. Xia, L. Zhang, Y. Zhou (2011). Trajectory Clustering Algorithm Based on Structural Similarity, Journal of Communications 32(9), 103-109.
Zhang, Y. L., Y. Dang, P. G. He (2015). Quantitative Analysis of the Relationship of Biology Species Using Pearson Correlation Coefficient, Computer Engineering and Its Application 33, 79-83.
Zhu, F. X., Y. J. Zhang, Z. J. Gao (2012). Research on ship behaviors based on data mining, Journal of navigation of China 35(2), 50-54.