

# Study on Correlating Multiple Attribute Information for the Dynamic State of Inland River Ships

Yaotian Fan<sup>1,3,a</sup>, Xianzhang Xu<sup>2,b</sup>, Wang, Chi Wang<sup>1,c</sup>

<sup>1</sup>School of Navigation, Wuhan University of Technology, Wuhan, China, 430063

<sup>2</sup>Faculty of Engineering and the Environment, University of Southampton, Southampton, Hampshire, The United Kingdom, SO17 1BJ

<sup>3</sup>Hubei Key Laboratory of Inland Shipping Technology, Wuhan, China, 430063

<sup>a</sup>email: fan.yaotian@163.com, <sup>b</sup>email: xx1u14@soton.ac.uk, <sup>c</sup>email: 1058277631@qq.com

**Key words:** Information Correlation; AIS; Shipborne Radar; Inland River Ships.

**Abstract:** Based on the multi-sources, multi-dimension and isomerism of the ship-borne sensors, a model was constructed to correlate the ships' dynamic state information. It aims at enhancing the ability of inland river seaman on noticing and discriminating the target ship while the reliability of the model is improved at the same time. Since the vessels traffic in inland river is intensive and complex, it requires a more reliable information correlation based on the AIS and radar information fusion. This correlation model integrated the fuzzy Comprehensive method, BP neural network, grey theory and Euclidean method then used the redundancy technique to discriminate the correlation degrees and finally obtained the target radar information associated with the AIS data. An experiment was carried out to check the validity of the integrating model. It can not only increase the accuracy of information correlation but also provide the fundamental work on fusing AIS and radar data. In that case, this study is important as the theoretical references to guarantee the inland river navigation safety and enhance the traffic efficiency.

## Introduction

The information fusion research in ship-handing mainly focus on the data fusion for multiple ship-borne sensors currently. It can be applied into the Vessel Intelligent Collision Avoidance Decision Support System. Since Automatic Identification System (AIS) and ship-borne radar (radar) are two of the main shipborne navigation equipment, many studies address in correlating the track from AIS and radar then fusing the data. This paper presents the information correlation for AIS and radar data. Nowadays there are two primary algorithms of track correlation: Numeral Statistics and Fuzzy Mathematics.

Mutual K-Nearest Neighbour<sup>[1]</sup> (MK-NN), Central Clustering<sup>[2,3]</sup> and Sequential Correlation are the main algorithms applied in Numeral Statistics data correlation. Li (2007) presented the concept name Correlation Threshold<sup>[1]</sup>. When the correlation degree between AIS and radar track is beyond the threshold value, it judges that the track is successfully correlated. Luo (2003) proposed Central Clustering method to correlated the AIS and radar track of target vessel and provided the methodology which can achieve the data conversion, space-time unification and data correlation. Gan (2004) applied the Sequential Correlation algorithm to associate the real-time track with the historical track<sup>[4]</sup>. This method can reduce the errors when the target is in the intensive traffic environment or there are many track crossing.

Statistics Weighted approach, Back Propagation (BP) Neural Network, Grey Correlation, Kalman Filtering algorithms and Fuzzy Comprehensive method are the classical theories applied into Fuzzy Mathematics data correlation. Lin et al. (2006) correlated the AIS and radar data using fuzzy correlation algorithms<sup>[6]</sup>. Chang (2004) added Statistics Weighted approach into the track correlation according to the errors so that the accuracy could be improved. Hu and Lin (2011) correlated the AIS and radar data by applying BP neural network theory. They analyzed the structure of network and checked the validity of this method in simulation by MATLAB software<sup>[7]</sup>. Wang (2010) proposed to apply the Grey Correlation theory into the track correlation and it was proved by

writing the programme codes then forming the fusion images based on the AIS and radar data<sup>[8]</sup>. Sermi and Mugnai (2013) presented a set of AIS and radar data fusion theories according to the systematic concepts provided by International Maritime Organization (IMO). They checked the parameter analysis through two dispersion fusion method based on the research value and real data. Some suggestions on the covariance matrix calculating AIS and radar data were also presented.

The current research on track information correlation are mainly applying single theory. That is because the traffic density on the open sea is not very big, one method is enough to the accuracy requirement. But the situations on inland river are quite different, the traffic flow is much more intensive and the state is very complex. So that requires a more reliable correlation method. This paper integrated Fuzzy Comprehensive method, BP Neural Network, Grey Theory and Euclidean Distance algorithms with the information correlation theory. An experiment was carried out to check the validity of the integrating model.

### Integrating Information Correlation Model

Fuzzy Comprehensive method can increase the fusion accuracy but it adds the complicity of the algorithm as well. Furthermore, the correlation algorithm depends on the data type. The parameters of BP Neural Network structure and the value of threshold may make the results quite different. The parameters of network is decided by the size of data and that requires to find the rule in research. The Gery Theory is still under study since it can only be applied into computational simulation. The Euclidean algorithm is calculating the Euclidean distance between AIS data and radar data then obtaining the most possible correlated radar data. However the accuracy of Euclidean method needs to be improved. This paper integrated those methods to avoid the unsatisfactory results from single algorithm. The model is shown in Figure 1.

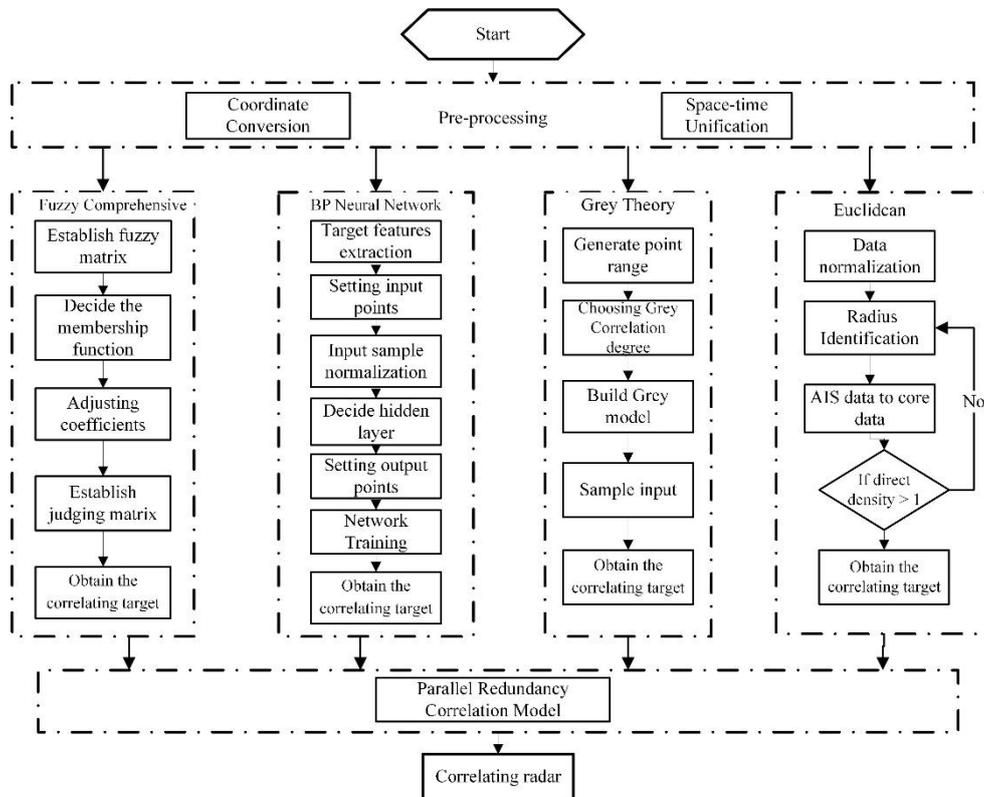


Figure 1. Methodology

The static and dynamic state data of the target vessels can be imported from the AIS on the own ship. However, when the radar tracking one target vessel it requires to eliminate the data from several other ships' radar until finding the target which is the same as what the AIS data refers to. Since the vessels on inland river are very intensive and busy, the updating data within one period is

huge while the historical data is little. To find out the radar data which is associated with the own ship's AIS data, it is necessary to correlate the AIS data with radar data from several targets before data fusion. This paper applies four methods i.e. Fuzzy Comprehensive method, BP Neural Network, Grey Theory and Euclidean Distance algorithms in correlating data. It follows the Weighted Average approach to construct the combined data correlation model. The model is able to discriminate the radar signal related to AIS data. The parallel redundancy technique is applied to enhance the reliability of the model.

The redundancy technique can enhance the reliability of system output through a system parallel model. In Figure 2, the parallel model is formed with 4 parallel working units and outputs through a voter. This design is known as "working redundancy". The N modules redundancy system consists of N ( $N=2n+1$ ) sub-systems with the same properties and a voter. When most sub-systems have the same result, the voter would output it. If there are  $n+1$  or more sub-systems work well, the result is correct. That means only if half or more sub-systems work well, the system can complete the task. Under the assumption that the reliability value of each sub-systems is  $R_o$  and they are independent. The total reliability of the N modules redundancy system is:

$$p = 1 - \sum_{i=n+1}^m C_m^i R_o^i (1 - R_o)^{m-i} \gg R_o. \quad (1)$$

When two sub-systems are similar to the other two, it would compare all the four modules and rank them according to the similarity degree. The module with the lowest similarity would be excluded before inputting the new data into redundancy model and finally the output would be obtained. So the redundancy technique is able to greatly enhance the reliability of AIS and radar data correlation theoretically.

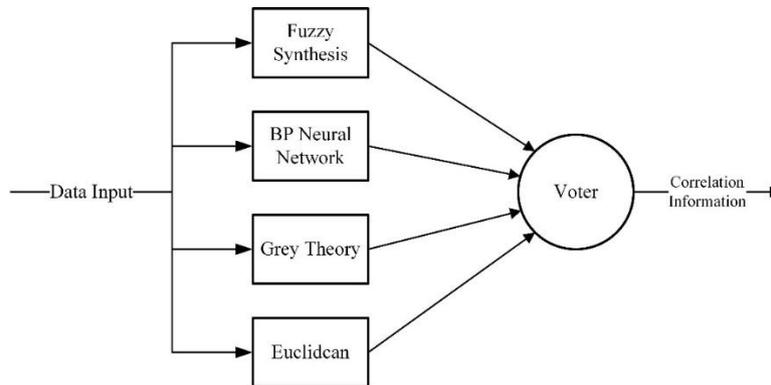


Figure 2. Multiple Modules Redundancy System

The four modules redundancy system in this paper consists of four sub-systems: Fuzzy Comprehensive method, BP Neural Network, Grey Theory and Euclidean Distance algorithms. In order to vote the decision, it ranks these four methods by weighted average approach according to their correlation degrees and then acquire the final correlation information. The details of this algorithm are:

Fuzzy Comprehensive ranking:  $C_{FC} = (c_{fc1}, c_{fc2}, \dots, c_{fc10})$ , where  $c_{fc1} > c_{fc2} > \dots > c_{fc10}$ ;

BP Neural Network ranking:  $C_{BP} = (c_{bp1}, c_{bp2}, \dots, c_{bp10})$ , where  $c_{bp1} > c_{bp2} > \dots > c_{bp10}$ ;

Grey Theory ranking:  $C_{GT} = (c_{gt1}, c_{gt2}, \dots, c_{gt10})$ , where  $c_{gt1} > c_{gt2} > \dots > c_{gt10}$ ;

Euclidean ranking:  $C_E = (c_{e1}, c_{e2}, \dots, c_{e10})$ , where  $c_{e1} > c_{e2} > \dots > c_{e10}$ .

The weights of the correlation degrees of each sub-systems are:

$$W_{c_{fc1}} = W_{c_{bp1}} = W_{c_{gt1}} = W_{c_{e1}} = 0.19 - 0.02(i - 1). \quad (2)$$

When each correlation ranking of the ship 'i' ( $1 \leq i \leq 10$ ) corresponding to  $C_{FC} - j$ ;  $C_{BP} - u$ ;  $C_{GT} - k$ ;  $C_E - n$  ( $1 \leq j, u, k, n \leq 10$ ), the weighted average of ship 'i' is:

$$Z_i = W_{C_{fcj}} + W_{C_{bpu}} + W_{C_{gtk}} + W_{C_{en}} \quad (3)$$

If  $Z_j = \max\{Z_1, Z_2, Z_3, \dots, Z_{10}\}$ , (4)

the ship 'j' can be regarded as the most correlated target ship.

When there are more than one  $Z_j$ , it would compare the ranking of the similarity of those four methods then exclude one method in the lowest similarity. Following that, the system would start a three modules redundancy ranking.

The similarity between two methods is calculated by the Euclidean Distance. A larger Euclidean distance indicates a lower similarity. The equation of Euclidean Distance is:

$$distance(d_i, d_j) = \sqrt{\sum_{k=1}^n (W_k(d_i) - W_k(d_j))^2} \quad (5)$$

Based on the Euclidean Distance among  $C_{FC}$ ,  $C_{BP}$ ,  $C_{GT}$  and  $C_E$ , a matrix can be obtained:

$$\begin{bmatrix} 0 & d_{FCBP} & d_{FCGT} & d_{FCE} \\ d_{FCBP} & 0 & d_{BPGT} & d_{BPE} \\ d_{FCGT} & d_{BPGT} & 0 & d_{GTE} \\ d_{FCE} & d_{BPE} & d_{GTE} & 0 \end{bmatrix}$$

According to the Euclidean Distance matrix, the sum of the Euclidean distance between one method and the other three can be derived:

$$D_{FC} = d_{FCBP} + d_{FCGT} + d_{FCE} \quad (6)$$

$$D_{BP} = d_{FCBP} + d_{BPGT} + d_{BPE} \quad (7)$$

$$D_{GT} = d_{FCGT} + d_{BPGT} + d_{GTE} \quad (8)$$

$$D_E = d_{FCE} + d_{BPE} + d_{GTE} \quad (9)$$

Finally, the  $\max\{D_{FC}, D_{BP}, D_{GT}, D_E\}$  would be excluded, the remaining three methods would be ranked using three modules redundancy model.

## Experiments Setting

In order to fully indicate the features of the inland river vessel traffic, not only the AIS and data of real ship in straight line should be collected but also the data of ships' substantial motivation. However, considering the operation space limit on inland river and the complex environment and traffic situation, the experiment taken on inland river may easily make accidents. So in November 2015 some experiments were carried out on the seaboard open water in daytime with good visibility. These experiments aimed to collect AIS and radar data for simulation and verification. Since big ship is slow in course turning and speed changing, in consideration of the maritime safety the experimental ship is a pilot vessel with good manoeuvrability. The experiment route starts from No. 6 wharf in Qingdao, China to Tuan Island, Jiaozhou Gulf, China then returns.

The target vessel is "Pilot No.2" pilot vessel in Qingdao Pilot Station. On the one hand, in Qingdao Vessel Traffic Service (VTS) Centre the own ship radar data was simulated by VTS radar and it was used to follow the track of "Pilot No.2" pilot vessel in real time. Meanwhile, the longitude,

latitude, navigational speed and course data of the vessel from AIS and radar were collected. On the other hand, a Differential Global Positioning System (DGPS) was installed on the target vessel, the antenna was placed near the vessel's own GPS antenna to collect the data of longitude, latitude, navigational speed and course. Since the accuracy of DGPS is quite high, the data collected can be used to compare the correlation model with target vessel in the subsequent verification procedure. The photo of experimental vessel and its track plotting in VTS display is shown in Figure 3.

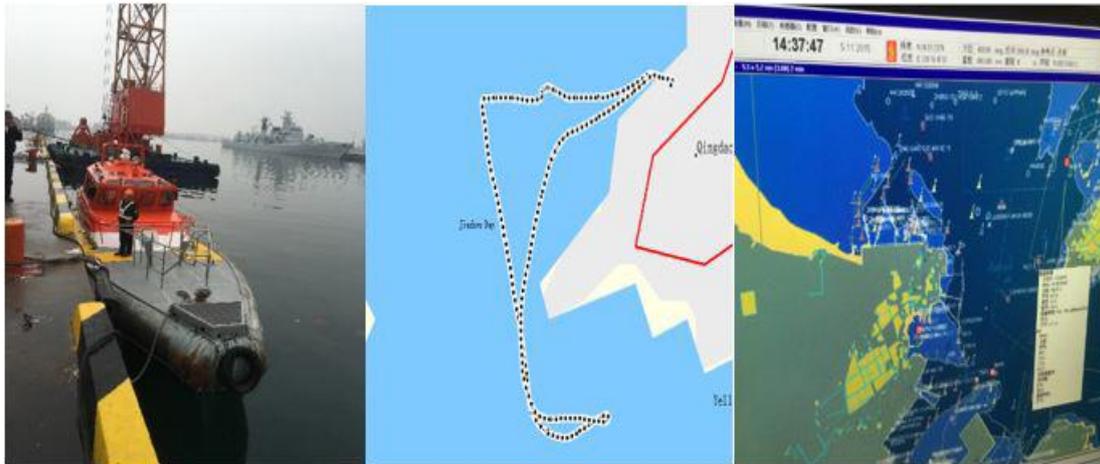


Figure 3. Experiment and Data Collection

### Simulation and Verification

The process of simulation and verification can be listed:

1. Collect data of the target vessel's from AIS, radar and DGPS;
2. Generate several virtual AIS and radar data simulating other vessels, curved channel and shoal;
3. Load the collected data and construct the complex and intensive traffic environment on inland river (see Figure 4);
4. Start the data correlation model calculation by parallel redundancy technique based on the simulated vessel data and real data collected so that the reliability can be verified.

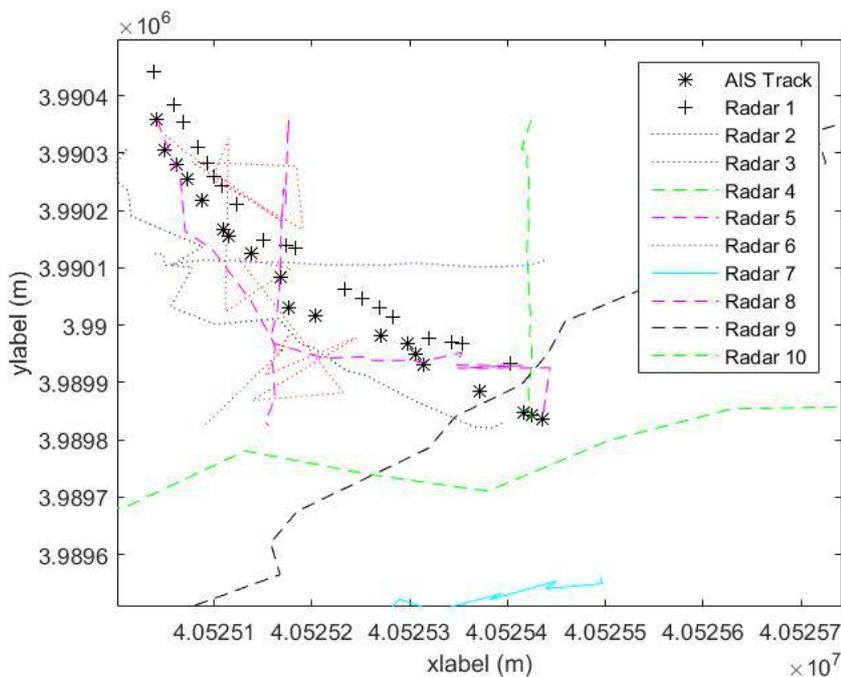


Figure 4. Track of the Real Vessel Data and Virtual Data

As shown in Figure4, the \* symbol indicates the AIS data collected while the + shows the radar data from the VTS, it was named as Radar 1. Radar 2 to Radar 10 is the virtual data applied to simulate

other vessels motivation and track so that there is enough distraction like the real situation on inland river.

### Fuzzy Comprehensive Correlation

In this paper, the adjustment coefficients of Cauchy distribution membership function were set as  $\alpha_1 = 7487$ ,  $\alpha_2 = 1111$ ,  $\alpha_3 = 172$ ,  $\alpha_4 = 133$ ,  $\beta_q = 2$  respectively; the single-factor fuzzy weight set can be expressed as  $B = [b_1 \ b_2 \ b_3 \ b_4] = [2 \ 2 \ 1 \ 1]$ . The square of Euclidean distance between the AIS data and radar data of the target vessel's speed, course, position etc. was calculated to evaluate the single-factor fuzzy set. The final correlating rank result calculated by Fuzzy Comprehensive method is: Target vessel 1, Target vessel 2, Target vessel 4, Target vessel 5, Target vessel 7, Target vessel 3, Target vessel 8, Target vessel 6, Target vessel 9, Target vessel 10.

### BP Neural Network Correlation

The characteristic factors of vessel were chosen as the longitude, latitude, speed and course from both radar data and AIS data. So there are 8 input points totally. Since the data format from these two sensors are different, the input sample was normalized before the network training. For the network training, the time of training, learning rate, time-limited iteration process and training accuracy need to be defined based on the network requirement. In the simulation, the training times were set as 6000, learning rate was 0.4, and training accuracy was  $1e-9$ . However, after 4778 times training the network reached the correlation requirement. The correlation results between 10 radar data with the collected AIS data are presented in the Table 1.

Table 1. BP Neural Network Correlation Result

Vessel	CD*									
1	0.8192	1.0446	1.0114	1.0236	0.2531	-0.0015	0.565	0.9398	0.9398	0.1652
2	0.6971	1.0247	0.8019	0.8014	0.8597	0.8793	0.7878	0.7341	0.8505	0.6568
3	0.4608	1.0766	-0.1301	0.2508	0.6509	0.0671	0.765	0.0342	-0.0819	-0.0983
4	-0.4446	-0.4409	-0.233	-0.1821	-0.1737	0.0033	0.1282	0.1923	0.2164	-0.0408
5	0.1348	0.1774	0.2362	0.1417	-0.0639	-0.3722	-0.4564	-0.4437	-0.4346	-0.0877
6	-0.2084	0.0355	-0.0008	-0.3542	0.064	-0.1454	-0.1896	0.4198	0.486	0.103
7	0.2152	0.0495	0.0076	0.1788	0.1737	0.0335	0.0869	-0.4784	0.0704	0.2222
8	-1.44	-0.9953	-1.4695	-0.8972	-0.9564	-1.2867	-0.1025	-0.7283	-0.3834	-0.2045
9	-0.3543	-0.0576	-0.4758	-0.0142	-0.4779	-0.0421	-0.0966	-0.1335	0.2721	0.4547
10	-0.0464	-0.0206	0.1802	0.3578	0.4175	0.6039	0.7349	0.7772	0.4645	0.3764
Vessel	CD*									
1	0.0925	0.0474	0.9589	0.1619	0.0465	-0.0496	0.0078	-0.0831	0.9241	1.0006
2	0.2951	0.0186	-0.1221	-0.033	0.0117	0.0863	0.0811	0.2792	0.0058	0.0016
3	-0.1193	-0.1443	-0.0315	0.0092	-0.1984	0.0338	0.4805	0.2755	-0.4237	0.2217
4	-0.3571	-0.6536	-0.799	-0.8015	-0.7345	-0.5984	-0.1729	0.1252	0.394	0.3504
5	0.3322	0.2658	0.394	0.4099	0.606	0.9986	0.9103	1.1136	1.4432	1.2066
6	0.3867	0.2904	0.729	0.4665	0.4127	0.6012	0.5937	0.7935	0.6649	0.3881
7	0.1218	-0.3576	0.1344	0.0014	0.2758	0.0313	-0.2266	0.129	0.5342	0.5425
8	-0.199	-0.1663	-0.1496	-0.2159	-0.2109	-0.2184	-0.1833	-0.167	-0.1655	-0.1559
9	0.1616	0.1731	0.2539	0.2413	0.4196	0.5834	1.0021	1.0306	0.6504	-0.5766
10	-0.0279	-0.0703	-0.2659	-0.2078	-0.2433	-0.2392	-0.3972	-0.8476	-1.0383	-1.1068

-CD\* means Correlation Degree

The first threshold value is 0.9 and the second threshold value is 12. The final correlating rank of the 10 radar is Target vessel 1, Target vessel 5, Target vessel 9, Target vessel 2, Target vessel 6, Target vessel 10, Target vessel 7, Target vessel 3, Target vessel 4, Target vessel 8.

## Grey Theory Correlation

The Grey Theory was calculated through the data of the position, speed and course of the target vessels. According to the data collected, the correlation results are given in Table 2.

Table 2. Grey Theory Correlation Results

Target Vessel	1	2	3	4	5	6	7	8	9	10
Speed Correlation	14.4	14.6	7.34	10.3	11.8	9.97	10.9	13.0	11.1	7.84
Course Correlation	14.7	8.27	4.86	7.34	7.57	5.57	3.57	5.85	4.87	3.13
Y label Correlation	17.1	19.0	15.6	20	15.8	17.7	9.37	19.3	16.1	12.0
X label Correlation	18.0	17.8	19.8	14.4	18.0	17.0	17.8	16.2	11.8	13.3
Total Correlation	64.4	59.8	47.6	52.1	53.3	50.3	41.6	54.5	44.0	36.3

The total correlation degree rank is Target vessel 1, Target vessel 2, Target vessel 8, Target vessel 5, Target vessel 4, Target vessel 6, Target vessel 3, Target vessel 9, Target vessel 7, Target vessel 10.

## Euclidean Correlation

The theory of Euclidean correlation is actually calculating the Euclidean distance between two points in the n-dimensional space. In this paper, the Euclidean distance between 10 radar data and the collected AIS data. The correlating rank is Target vessel 2, Target vessel 1, Target vessel 3, Target vessel 4, Target vessel 10, Target vessel 5, Target vessel 6, Target vessel 8, Target vessel 9, Target vessel 7.

## Results Discussion

Table 3 indicates the correlation ranking by the four methods.

Table 3. Correlation Results of Four Methods

Fuzzy Comprehensive	1	2	4	5	7	3	8	6	9	10
BP Neural Network	1	5	9	2	6	10	7	3	4	8
Grey Theory	1	2	8	5	4	6	3	9	7	10
Euclidean	2	1	3	4	10	5	6	8	9	7

According to the Parallel Redundancy Correlation model and equation 2 to 9 presented in the previous section, the final vessel rank is listed as follow: Target vessel 1, Target vessel 2, Target vessel 5, Target vessel 4, Target vessel 3, Target vessel 7, Target vessel 6, Target vessel 8, Target vessel 9, Target vessel 10. That indicates that the radar data of Target vessel 1 is the same as the collected AIS data, which is correct.

In this section, an inland river traffic environment of which the vessel density is large and the vessel heading changing frequently was simulated. After applying the four correlation method and the parallel redundancy correlation model, it successfully excluded the fault correlation and found the correct radar data which is associated with the collected AIS data under the distraction of other 9 virtual vessel radar data. The validity of the parallel redundancy correlation model is checked.

## Conclusions

This paper constructed a data correlation model based on the characteristics of shipborne sensors' data. The data was pre-processed by normalizing format and time before correlated by four methods: Fuzzy Comprehensive, BP Neural Network, Grey Theory and Euclidean Distance. A parallel redundancy technique was added to improve the reliability of the correlation result. After the experiment and simulation, it found that single correlation method may easily make mistake on inland river vessel but the parallel redundancy model can enhance the accuracy of correlation ranking and reduce the possibility of making mistakes. This model is more stable and accurate than traditional single method correlation especially when the questions relating to inland river where the

traffic is very intensive, the distance between vessels are small and the vessel state changes frequently. In that case, this model can support to enhance the maritime safety on inland river.

## References List

- [1] Lee, Weiyun. Discussing Information Fusion about Radar and AIS in VTS. (Dalian Maritime University, 2007.) In Chinese.
- [2] Suo, Jidong. Study on AIS and Radar Information Fusion in VTS. (Dalian Maritime University, 1999). In Chinese.
- [3] Luo, Suyun. Research on Fusion Method of AIS and Radar Objects Location Data. (Wuhan University of Technology, 2003). In Chinese.
- [4] Gan, Xing. Study on AIS and Radar Information Fusion in VTS. (Dalian Maritime University, 2003). In Chinese.
- [5] Lin, Changchuan; Sun, Tengda; et, al. Algorithm and Simulation of Fuzzy Correlation of Tracks from Radar and AIS. (Journal of System Simulation, Vol. 18, 2006). In Chinese.
- [6] Chang S. Development and analysis of applications as an efficient tool for vessel traffic service. (MTS/IEEE OCEANS 04, U.S.:2004:2249-2253).
- [7] Hu Xiaorui, Lin Changchuan 2013A Preliminary Study on Targets Association Algorithm of Radar and AIS Using BP Neural Network. (International Conference on Advanced in Control Engineering and Information Science, China,2011: 18-19).
- [8] Wang, Binglin. The Researches of AIS and Radar Information Fusion Based on the Grey Theory. (Dalian Maritime University, 2010). In Chinese.
- [9] Sermi Francesco, Mugnai Clio, Cuccoli Fabrizio, Facheris Luca. Analysis of the radar coverage provided by a maritime Radar Network of Co-operative Vessels based on real AIS data. (European Radar Conference 10th, Nuremberg, 2013:9-11).
- [10] Zou, Qihua. Study on AIS and Radar Information Fusion Based on Fuzzy Comprehensive Function. (Dalian Maritime University, 2013). In Chinese.
- [11] Hu, Xiaorui. Research on Targets Information Fusion of Radar and AIS Based on BP Neural Network. (Jimei University, 2012). In Chinese.
- [12] Satish Kumar. Neural Networks. (Tsinghua University Press, Beijing, 2006).
- [13] Adam P. Piotrowski. Differential Evolution algorithms applied to Neural Network training suffer from stagnation. (Applied Soft Computing, Vol.21. 2014).
- [14] Manoel V. M. França, Gerson Zaverucha. Fast relational learning using bottom clause propositionalization with artificial neural networks. (Machine Learning, Vol.94 (1), pp.81- 104 2004)
- [15] Lu Xia, Hou Muzhou, A new constructive neural network method for noise processing and its application on stock market prediction. (Applied Soft Computing, Vol.15. 2014).
- [16] Li Cunbin, Wang Kecheng. A new grey forecasting model based on BP neural network and Markov chain. (Journal of Central South University of Technology, 2007).