Collision Hazard Identification of Unmanned Vessels in Inner River Based on Particle Swarm Parameter Optimization Support Vector Machine

Wenli Ma, Jie Yang*, Qingnian Zhang, Zhiqiang Guo and Rong Hu

Key Laboratory of Fiber Optic Sensing Technology and Information Processing, Ministry of Education, Wuhan University of Technology, Wuhan, China

* 794393254@qq.com

Abstract. In this thesis, Collision hazard identification of unmanned vessels in inner river is studied. The electronic chart, radar and AIS system of ship assembly can provide the data that affect the normal operation of unmanned ships. The fuzzy comprehensive evaluation method is used to evaluate the collision of unmanned ships Levels of danger. The particle swarm optimization algorithm is used to support Vector machines classify the ship's collision hazard classification. The data is based on data from 112 survey reports of collision between two ships issued by various provinces and local maritime authorities in China during 2004-2015 as a simulation sample. The collision risk of unmanned ships is divided into three levels, and Proposed a different level of unmanned ships coping methods. By comparing experimental results with ordinary support vector machine, particle swarm optimization, genetic algorithm and grid search parameters to optimize the performance of support vector machine, the support vector machine based on Particle Swarm Optimization can be used to predict the collision hazard identification of unmanned vessels, and it has a good performance.

1. Introduction
The recognition of the safety situation of unmanned vessels is an important guarantee for the navigation of ships in the river. The collision risk of ships is one of the important indicators of the safety situation of unmanned vessels. According to the statistics of some Chinese provinces and local maritime bureaus, a total of 165 water traffic accident data released on the official website, the collision accident accounted for as high as 68%, ranked first in water traffic accidents. Based on this data, it is particularly important to study the collision risk of unmanned ships.

Xue [1] adopted DCPA and TCPA weighted method to calculate the ship collision risk when multiple ships would encounter, and then determined the best path for collision avoidance decision. In 2000, Fan [2] put forward an evaluation method called Hazard Index Method, using the gas equation to get the collision accident indexes in all regions of the route, combined with the rest of the risk indicators affecting the normal operation of the ship and the ship flux weighting to get the ship Real-time risk index on a particular route. Yang [3] proposed a fuzzy Bayesian approach to the monitoring of shipping safety status. Taking the state as a node and the Bayesian network reasoning, the integrated safety probabilities and the risk factors of the ship under different environmental nodes were obtained Probability value. Ao [4] put forward the establishment of navigation safety assessment
index system, the use of fuzzy comprehensive evaluation method to achieve the definition of different levels of risk factors.

Based on the results of previous studies, this paper proposes to evaluate the collision risk of unmanned ships by using support vector machines optimized by particle swarm optimization. The basic concepts and calculation models of collision risk are studied. The main factors that affect the risk of collision are analyzed, and the risk of collision under different conditions of the ship is obtained by fuzzy comprehensive evaluation. Based on this, particle swarm optimization is used to model support vector machines Parameter optimization, and the data regression results and support vector machines, grid-based support vector machines and genetic algorithms based SVM classification regression results were compared.

2. Sample data extraction and risk assessment

2.1. Sample data extraction

The sample data used in this paper is from 112 Chinese-made survey reports of collision between two ships issued by various provinces and local maritime authorities in China during 2004-2015. The report details the size information, speed of navigation and navigation position of two ships colliding with each other. Weather and water conditions information. Static information such as the size of inland river unmanned ships and ship encounters can be obtained by AIS; the dynamic speed of ship and ship encountering ship and heading and other dynamic information are obtained through information fusion of AIS and radar; Real-time weather and water status information forms the environmental information on ship encounters through the Electronic Chart System (ECDIS). The Japanese scholar Iwasaki Kawai [5] proposed that the influence of collision probability of DCPA (shortest encounter distance) and collision risk of TCPA (shortest encounter time) with fuzzy reasoning method is proposed, which has been generally recognized by the sailing session.

2.2. Fuzzy comprehensive evaluation method

In this thesis, fuzzy comprehensive evaluation method is used to evaluate the collision risk level of unmanned ships. The original index data that affects the collision risk of ships mentioned above are used and the membership functions of different factors are given.

2.2.1. DCPA risk membership function. According to the definition of DCPA, the greater the value of the DCPA, the lower the probability of collision between the two ships, and the less dangerous it is. DCPA's risk membership function[6], such as equation (1):

\[
\tau_{DCPA} = \begin{cases} 
1 & \text{DCPA} \leq d_1 \\
\frac{1}{2} - \frac{1}{2} \frac{\pi}{d_2 - d_1} \left( \frac{d_2 + d_1}{2} - \text{DCPA} \right) & d_1 < \text{DCPA} < d_2 \\
0 & d_2 \leq \text{DCPA}
\end{cases}
\]  

(1)

In this case, d_1 is the safe meeting distance of the ship, d_2 is the absolute safe distance of the ship, and d_1=d_2 in general.

2.2.2. TCPA risk membership function. TCPA can be expressed as the urgency of unmanned ships meeting the danger of collision with ships. In the event of a minimum encounter distance, TCPA indicates the time required for the collision between the two if the unmanned vessel took no collision avoidance measures. It can be seen that the smaller TCPA, the greater the risk of collision. TCPA membership function of the degree of risk [7] as formula (2):

\[
\tau_{TCPA} = \begin{cases} 
1 & \text{TCPA} \leq d_1 \\
\frac{1}{2} - \frac{1}{2} \frac{\pi}{d_2 - d_1} \left( \frac{d_2 + d_1}{2} - \text{TCPA} \right) & d_1 < \text{TCPA} < d_2 \\
0 & d_2 \leq \text{TCPA}
\end{cases}
\]  

(2)
2.2.3. Collision risk assessment. Based on the fuzzy comprehensive evaluation method, the weighted expression of five factors that defines the collision risk of unmanned ships[8]:

\[
\text{CRI} = w_{\text{DCPA}} r_{\text{DCPA}} + w_{\text{TCPA}} r_{\text{TCPA}} + w_{D_f} r_{D_f} + w_{Q_f} r_{Q_f} + w_k r_k
\]  

(4)

Among them, according to different factors, the degree of impact on the ship collision risk is different, based on the three methods of evaluation index weight: the Dofei method, expert survey method and AHP. This article defines the different influencing factors as:

\[
W = \{ w_{\text{DCPA}} = 0.36, w_{\text{TCPA}} = 0.32, w_{D_f} = 0.14, w_{Q_f} = 0.1, w_k = 0.08 \}
\]  

(5)

In this thesis, the collision risk is divided into three different stages, and the unmanned ships in different stages should be adjusted to varying degrees.

(1) Collision hazard level three
CRI is in the range of 0 to 0.4. The automatic control center of the unmanned ship is provided with ship data meeting with the ship and real-time tracking of the target ship, and real-time data are returned to re-evaluate the collision risk.

(2) Collision hazard level two
CRI is in the range of 0.4 to 0.7. Said to unmanned automatic control center for risk warning, apply for unmanned craft autopilot heading fine-tuning, return data for real-time risk assessment.

(3) Collision hazard level one
CRI is in the range of 0.7 to 1. Said the automatic control center for unmanned vessels to apply for emergency braking, this route is defined as not feasible, to re-routing to avoid collision.

3. Support vector machine
SVM (Support Vector Machine) model has great advantages in pattern recognition such as small sample, high dimension data and non-linearity. It is a modeling model based on VC dimension theory and minimization of structural risk, and can be solved under certain circumstances. Dimension disaster, over learning, owe learning and other machine learning some shortcomings. In practice, the choice of SVM parameters directly determines the performance of the whole model, and the choice of parameters is often based on actual experience, that is, through a large number of experiments choose the optimal parameters.

Grid search parameter optimization is known as a primitive programming approach to solve constrained nonlinear extremum problems [9], similar to the "exhaustive method." The optimization method is to divide the sample into grids within a certain range and then traverse all grid points to find the optimal parameters. However, due to the low classification accuracy of each parameter group in the grid, it is a waste of time to traverse the entire region when the accuracy of classification in a particular parameter group is high.

GA (Genetic Algorithm) has strong robustness and global optimization search ability [10]. It is a randomized search algorithm based on Darwin's theory of evolution and Mendel's genetic variation
theory. Including coding, population initialization, selection, cross, mutation and other steps. The algorithm needs to acquire parameters such as the crossover rate and mutation rate, and the choice of these parameters seriously affects the quality of the solution. At present, the choice of these parameters is mostly based on experience. There is no feedback information that can make timely use of the network.

Particle Swarm Optimization (PSO) is a kind of optimization algorithm for simulating swarm intelligence. It uses some particles without mass and volume as individuals to initialize the particles randomly by optimizing the search to form a population, so that the particles The group presents complex features, changing the position of each particle, the entire algorithm iteratively. Particle swarm optimization, as an efficient optimization tool, not only has excellent global search ability and fast convergence speed, but also has a good effect on identifying collision risk of inland river unmanned ships. The parameters can get accurate danger through training Degree level.

We define the input sample dataset as T, and the steps of hazard algorithm based on PSO-SVM are as follows:

1. Divide sample data set T into training data set T1 and test data set T2, and separate the first rank of danger levels of data sets T1 and T2 as the identification sample set, leaving the remaining 8 data sample elements, Data normalization;
2. A set of SVM classifier parameters can be obtained by using each support vector in the normalized training dataset T1 (including the representative samples and the element samples) to form a particle and form a particle population X;
3. The population of particles set initialization parameters C1, C2, the initial velocity matrix, the optimal individual particle initial individual position and the global optimum position;
4. (6) is the fitness function of particle swarm, calculate the fitness function value of each particle in particle swarm;
\[
F(x) = \frac{1}{m} \sum_{i=1}^{m} (f_i - y_i)^2
\]  \(6\)
5. According to the value of the fitness function, the particle's best position and the global optimum position are adjusted, and the state of the example is updated to obtain a new set of SVM classifier parameters;
6. When the fitness function value meets the requirement or reaches the maximum number of iterations, proceed to step ⑦, otherwise, return to step ④;

Using the best parameters for SVM network training, network prediction; Table 1 is the simulation of the input data sample profile.

Table 1. Sample input data

| Risk rating | CRI   | DCPA | TCPA | Relative distance | Target ship Angle | Vt/Vo | Vr  | Cr  |
|------------|-------|------|------|-------------------|-------------------|-------|-----|-----|
| 1          | 0.93516 | 1.8  | 3.24 | 2.1               | 60                | 0.98019 | 20  | 0.99 |
| 2          | 0.43792 | 0.84 | 12.0 | 1.77              | 35                | 0.69565 | 7.73 | 6.56 |
| 3          | 0.30525 | 1.01 | 11.4 | 1.69              | 40                | 0.47916 | 7.08 | 3.24 |

4. Comparison and analysis of Simulation results
Taking the continuous CRI value of Table 1 as output data, 400 data samples of 500 sample data are taken as training samples and 100 data yang samples as test sample. Using the discrete Risk rating in the sample data as output data, the input data is the remaining attribute data values in the sample data. Taking 250 data samples from 500 sample data as training samples and 250 data yang samples as test samples. Comparisons of classification regression results of four different SVMs.
Table 2. Regression data analysis

| Method          | Best c | Best g | MSE  | Mean squared error | Regression |
|-----------------|--------|--------|------|--------------------|------------|
| SVM             | 2      | 1      | 0.1  | 0.0419285          | 0.850686   |
| GridSearch-SVM  | 8      | 1.442  | 0.04 | 0.0281851          | 0.895708   |
| GA-SVM          | 9.4316 | 1.6682 | 0.03 | 0.0255490          | 0.906331   |
| PSO-SVM         | 12.4562 | 1.7090 | 0.04 | 0.0237895          | 0.912771   |

Table 3. Multiple hazard level forecast results

| Method           | SVM  | GridSearch-SVM | GA-SVM | PSO-SVM |
|------------------|------|----------------|--------|---------|
| The first time   | %93.6| %96.4          | %96   | %97.6   |
| The second time  | %93.6| %96.4          | %97.2 | %98     |
| The third time   | %93.6| %96.4          | %96.8 | %98.4   |
| Average value    | %93.6| %96.4          | %96.7 | %98     |
5. Conclusion
This article explores the classification of unmanned vessels in river collision hazard classification and forecasting problems. The system of unmanned ships is used to acquire the relevant data that affects the safe driving of the ship. The classification of ship collision hazard classification is carried out by using SVM based on Particle Swarm Optimization. After a large number of simulation experiments and comparison of simulation results, we can clearly see the advantages of PSO in global search ability, and can build SVMs with better classification and recognition effects, and improve the classification accuracy of support vector machines.
Acknowledgments
This work was supported by National Science Foundation of China (NSFC) under Grand 51479159 and the Fundamental Research Funds for the Central Universities under 185209005.

References
[1] Cai.Y.Z, Wei.Y, Sun.M. “Multi-ship Collision Avoidance Simulation Based on Collision Avoidance Ship Key Algorithm”. Journal of Dalian Maritime University:2014, 40(1):13-16.
[2] Fan. Y.T. “Traffic Safety Evaluation Method for Inland Waterways - Dangerous Index Method”. Wuhan Jiaotong University:2000.3.
[3] Yang.J, Zhang.Q.N. “Research on Shipping Safety Status Monitoring Based on Fuzzy Bayesian”. Wuhan University of Technology:.2016.04.018.
[4] Ao.J, Hu.Q.Y, Zhao.R.Y. “Analysis of real time safety index in navigable waters”. Shanghai maritime university. Ship sea engineering:2008.8.
[5] A.N.Cockeroft. “The circumstance of sea collision”. The journal of Navigation: 1989:212-221.
[6] Li B, Pang F W. “An approach of vessel collision risk assessment based on the D–S evidence theory”. Ocean Engineering: 2013, 74(7):16-21.
[7] Ren Y L, Mou J M, Yan Q X, et al. Study on Assessing Dynamic Risk of Ship Collision[C]// International Conference on Transportation Information and Safety. 2011:2751-2757.
[8] Xin.H, Zhou.Z.S, lv.A.Q,. “Shipwreck Risk Early Warning Based on Fuzzy Comprehensive Evaluation Method ”. Chinese Journal of Safety Science: 2011, 21(4):115-120.
[9] Li.Q.Y ,Zhou.H, Lin.A.P. “Gray Melting Point Prediction Based on Grid Search and Support Vector Machine”. Journal of Zhejiang University:2011:45-12.
[10] Chen.W.G, Ten.L, Liu.J, Peng.S.Y, Sun.C.X. “Prediction Model of Transformer Winding Hot Spot Temperature Based on Genetic Optimization Support Vector Machines”. Transactions of China Electrotechnical Society:2014:29-1.