Fusion of the Targets of AIS and Radar Based on a Stacked Auto-Encoder

Cao Xiufeng1* & Gao Shu12 & Jiang Zilong1 & Chen Liangchen13 & Wang Yan4

1School of Computer Science and Technology, Wuhan University of Technology, Wuhan/China
2Hubei Key Laboratory of Transportation Internet of Things, Wuhan University of Technology, Wuhan/China
3Department of Computer Application, China Institute of Industrial Relations, Beijing/China
4School of Information Engineering, Guizhou University of Engineering Science, Bijie/China

*E. Mail: cxf0417@hotmail.com

Received 11 January 2017; revised 26 April 2017

Automatic identification system(AIS) and radar have its own advantages and disadvantages and employ track association and fusion to complement each other. This study proposes a new model of target track association and fusion based on deep learning to overcome the drawbacks of AIS and radar target traces. First, the features of fusion data are selected; then, the standard deviation is employed to normalize the pre-selected data. Next, an algorithm and regression are added to the top layer of the stacked auto-encoder(SAE), and a back propagation (BP) algorithm is used to adjust the weights and thresholds of the cost function for track association. Finally, rule items are added to the softmax regression to conduct track fusion. Experimental results demonstrate that the proposed model improves the accuracy of AIS and radar tracking.

Keywords: track association; track Fusion; SAE; BP; Soft max

Introduction

Multi-sensor technology has become a topic of considerable in-depth research. Relying on a single source of information to understand an object has not met the needs of users. The fusion of multi-sensor information to understand an object provides multilevel all-around recognition and becomes important in the field of information fusion, particularly in marine and inland navigation.

Radar information acquisition is active; it contains various types of information, such as the distance, location, speed, course, closest point of approach (CPA) and time to CPA(TCPA). The traffic panorama of a water area, including moving, stationary and fixed targets, can be obtained. Echoes of a radar target reflect the size and shape of the target to a certain extent.

The detection of automatic identification system (AIS) targets is passive and can only be performed on ships equipped with AIS, which receives and broadcasts information. AIS information includes static information, such as a ship’s name, call sign, maritime mobile service identity(MMSI), length, width, international maritime organization(IMO) code, and cargo loading; dynamic information, such as GPS location, speed, Universal Coordinated Time(UCT), and course; and other information, such as the ship’s draft, type, destination and expected arrival time, as well as the port and weather.

Because radar is susceptible to weather and other environmental influences, and it is also vulnerable to radar distance limitations; as a result, blind spots exist in detection, making the measurement accuracy and reliability relatively low. The ship cannot receive and broadcast information if AIS equipment is not installed; it can only receive and broadcast the AIS information from the target of ships. AISs do not have a complete information validation mechanism. The acquired AIS information may have missing and erroneous data, which must be manually corrected. However, AISs are not susceptible to environmental influences, such as terrain, weather, radio and other factors. A comparison of the information provided by an AIS and radar is shown in Table 1. The information provided by an AIS and radar can be merged to fully utilize their advantages and complement their shortcomings to provide more accurate and reliable target tracking, particularly in inland river ports, where the density of ships is particularly high.

There are some achievements in the association and fusion of AIS and radar information; most of these achievements have used statistical methods, fuzzy mathematics and neural networks. However, deep
learning in the field of ships is still relatively lacking even though it has been widely used in image recognition, natural language processing, artificial intelligence and other aspects. Therefore, deep learning brings new prospects and direction for the fusion of AIS and radar information.

Heymann\(^4\) presented a particle filter algorithm to evaluate track performance and simulate synthetic radar images based on the use of radar information; however, their algorithm only estimated the target position and velocity. Abellard\(^5\) built an integrated technology system by fusing broadband radar, visible light and AIS information to detect and classify ships on the Hudson River; however, passive acoustic and wake analyses are not considered. Guerrero\(^6\) proposed a Bayesian model for multi-sensor fusion to quantify the values of radar and AIS information; however, they did not consider the ship’s dynamic information. Danu\(^7\) presented a method to fuse asynchronous track estimation from the Over-the-Horizon radar system and AIS measurements; however, AIS location information is not used. Abielmona\(^8\) utilized Level 0/1 and high-level information fusion techniques for synthetic aperture radar and AIS to remove unnecessary items from the operator’s workspace. Xiao\(^9\) used statistical methods in the Rotterdam and Su-Tong Bridge waterways to characterize the lateral position, speed, heading and interval times of ships; however, the data analysis did not consider indirect information. Tharmarasa\(^10\) proposed a joint probabilistic data association framework that used kinematic and non-kinematic Bayesian inference to handle AIS ID swaps between ships, but it mainly focused on the ship’s position and highly consistent AISs. Nguyen\(^11\) presented subjective judgment value and bumper safety criteria to predict a ship route and used a liberalization manoeuvring model to perform the tracking assessment. Xu\(^12\) presented a weighted fusion algorithm for radar and AIS information and analysed the position error of the fused trajectory; however, the weights used in the algorithm were fixed. Xiaoru\(^13\) introduced a BP neural network for multi-target association judgment and multi-factor fuzzy integration decision-making arithmetic to fuse target information.

The traditional technology and methods, such as statistical methods, fuzzy clustering methods, probabilistic methods, Bayesian methods, Kalman filters, and particle filters, have been used to associate or fuse information. Although some researchers have
adopted BP and artificial neural networks, their training methods largely focus on shallow training. Only part of the dynamic or static information provided by AISs and/or radar was considered.

The key problem in fusing information from radar and AISs is track association and track fusion. Track association determines whether two tracks from two sensors represent the same target, and track fusion can be used to obtain and evaluate the target state.

Because coordinate transformation and time calibration are a relatively mature area of research, Cartesian coordinates and information interpolation or extrapolation are used to perform coordinate transformation and time calibration, respectively; however, these tasks are not the focus of our research. The method of deep learning is employed to track association and fusion for AIS and radar targets. The fusion model is shown in Figure 1. The fusion of AIS and radar information provides accurate perceived services of ship traffic, such as navigation safety, anomaly detection, maritime surveillance, and so on, which provides new opportunities and developments.

Materials and Methods
Selecting the ship’s features
The acquisition of AIS and radar information is impacted by a variety of factors, including differences in information from different regions, geographical location and the crew’s technical level. Track association and fusion generally focus on a relatively small area of the ship to process and analyze. Therefore, three aspects are used to represent features in this paper. In part 1, radar provides the target’s distance, location, velocity, course, CPA, and TCPA. The dynamic information, such as location, destination, velocity, course, distance from the main fairway, and UCT, and the static information, such as the ship’s name, length and width, call sign, MMSI, IMO code, and GPS location, are provided by an AIS in part 2. In part 3, weather, water, obstacles and other auxiliary information are provided by an AIS or radar. The first two parts are the most intuitionist expression of ship information and state features. Part 3, as part of the environmental information, is auxiliary information, which must be quantified.

AIS and radar feature selection
The ship’s longitude, latitude, speed, course, length and width provided by AIS are selected, and the target ship’s location, speed, course, distance and CPA are supplied by radar, which is described as follows.

Assuming that there are \( T \) sampling moments that consist of \( m \) AIS targets and \( n \) radar targets, \( A_{\text{Cog}}^k, A_{\text{Sog}}^k, A_{\text{Lon}}^k, A_{\text{Lat}}^k, A_{\text{Len}}^k, A_{\text{Wid}}^k, \) and \( A_{\text{Typ}}^k \) \( (i \in m, k \in T) \) respectively represent the \( i \)-th AIS target, course over ground, speed over ground, longitude, latitude, length, width, and type, respectively, at time \( k \).

\( R_{\text{Cog}}^j, R_{\text{Sog}}^j, R_{\text{Cpa}}^j \) and \( R_{\text{Dis}}^j \), and \( R_{\text{Dea}}^j \) \( (j \in n) \) represent the \( j \)-th radar target, course over ground, speed over ground, CPA, and relation to the ship's distance and position, respectively, at time \( k \).

Quantification of the auxiliary factors
The principal auxiliary factors involved in the state of the ship are quantified. The Yangtze River is located between latitude 24° 30’ ~ 35° 45’, also has other characteristics, such as the summer rainfall is generally large, relatively fast water flow; summer and winter season two strong wind speed, while the spring and autumn relatively moderate; In addition, there are haze, rain and fog weather, especially in haze period, visibility is relatively low. Therefore, we manually quantify the visibility, wind, water flow and waterway density according to prior knowledge and expert advice.

The range of weather visibility according to the requirements of visibility of maritime traffic safety laws can be normalized. \( W_v \) is defined as

\[
W_v = \begin{cases} 
0.1 & Vb > 1500 \\
0.3 & 1000 < Vb \leq 1500 \\
0.6 & 500 < Vb \leq 1000 \\
0.99 & Vb \leq 500 
\end{cases}
\]  

where \( Vb \) denotes visibility.

The wind power level of 0-17 is a standard for quantitative processing, and \( W_w \) is defined as

\[
W_w = \frac{\text{wind}}{17} \quad \text{wind} \in (0,17)
\]
The ship is generally in a certain state of water flow, such as the state of upstream and downstream, and $W_f$ is defined as

$$W_f = \begin{cases} 0 & \text{upstream} \\ 1 & \text{downstream} \end{cases} \quad \ldots (3)$$

In a certain region, such as a port, the waterway density will affect the state of the target ship. The ship density is quantified by $W_r$, which is defined as

$$W_r = \frac{1000}{a \cdot L_0} \quad \ldots (4)$$

where $L_0$ denotes standard ship length and $a$ is typically taken to be 6.

Parameter integration

To summarize, $v$ denotes the feature vector of the input model, and vector $v$ is expressed as

$$v = \{A_{Cog}, A_{Sog}, A_{Lon}, A_{Lat}, A_{Wid}, A_{Typ}, R_{Cog}, R_{Sog}, R_{Cpa}, R_{Dis}, R_{Dea}, W_1, W_2, W_f, W_r\}$$

The vector dimension is $7M + 5N + 4$, illustrating that the parameters have high dimensionality. SAE is employed for unsupervised learning the encoding function of the high-dimensional vector to obtain the corresponding low-dimensional information representation, which can be directly used as the feature of the AIS and radar fusion model.

Design of the information fusion model based on deep learning

The fusion of radar and AIS information includes three processes. First, time and space are calibrated and unified. Second, the track association is determined. Finally, track fusion information is considered. Research on the unification of spatio-temporal calibration is mature. The main aspects of fusion radar and AIS information are track correlation and track fusion, which are executed in a preliminary study based on the theory of deep learning.

Stacked auto-encode

The auto-encoder (AE) $^5$ has three layers: the input layer, hidden layer and reconstruction layer. The output of the AE is another representation of the input, as shown in Figure 2(1); AE has been used in the greedy layer-wise unsupervised network$^{16}$. The SAE$^{17}$ is a neural network with multiple layers of AEs that has powerful expression ability, as shown in Figure 2(2); it is a special neural network that has been widely used for dimensionality reduction$^{18}$, feature learning and classification tasks$^{19,20,21,22}$.

The SAE contains two parts: the encoder and decoder$^{23,24}$. The encoder performs the mapping transformation from the input vector $x \in R^d$ to the hidden feature $y \in R^d$ by a deterministic function $f_e$ (see Equation (5)), also called encoding.

$$y = f_e(x) = s_e(w_1x + b_1) \quad \ldots (5)$$

where $\{w_1, b_1\}$ is the parameter sets with a weight matrix $w_1 \in R^{d \times d^*}$ and a bias vector $b_1 \in R^{d^*}$, and $s_e$ is the activation function, which is typically a non-linear function, such as a sigmoid function or hyperbolic tangent function.

The decoder performs the mapping transformation from the hidden feature $y \in R^d$ to a reconstruction $\hat{x} \in R^d$, by another deterministic function $f_d$ (see Equation (6)), also called decoding.

$$\hat{x} = f_d(y) = s_d(w_2y + b_2) \quad \ldots (6)$$

where $\{w_2, b_2\}$ is the parameter set with a weight matrix $w_2 \in R^{d^* \times d}$ and a bias vector $b_2 \in R^{d^*}$, the parameters $w_1, w_2$ are restricted to satisfy the relation $w_2 = w_1^T$, and $w_2 = w_1^T$ is the another activation function that is typically the same as $s_d$. 

Fig. 2 —AE and SAE Architecture
The AE trains the network parameters by making \( x \) and \( \hat{x} \) as close as possible, where \( x \) and \( \hat{x} \) are determined by a reconstruction error function \( f(x, \hat{x}) \), as shown in Equation (7).

\[
f(x, \hat{x}) = ||x - \hat{x}||^2 
\]  \hfill (7)

The SAE is trained in an input \( \rightarrow 1 - N \) and greedy layer-wise manner. The input vectors of the SAE are fed to the input of the AE (see Fig 2(a) white color,). After fulfilling the AE training, the output hidden features are propagated to a higher layer; the \( N \)-th hidden layer output of the AE is the output of the SAE, which can be further fed into other applications. Thus, the SAE can be constructed according to certain rules of the AE.

Given a set \( \{ (x^1, \hat{x}^1), (x^2, \hat{x}^2), \ldots, (x^n, \hat{x}^n) \} \) of \( n \) training samples, for each of the input vectors of samples \( x^i \), the input layer of the SAE can be denoted as \( X^i = x^i \), and the encoding and decoding equation is defined as

\[
f_{l}(X^i) = s_i(w'X^i + b')
\]  \hfill (8)

Similarly, when \( \{w', b'\} \) is the parameter set, which is the same as \( \{w_i, b_i\} \) or \( \{w_2, b_2\} \), it can be obtained by greedy layer-wise training, where \( l \) denotes the number of layers and \( s_i \) is the activation function of encoding and decoding, which is the same as \( s_e \) or \( s_d \). This process establishes the SAE architecture.

**Track association method**

The process of radar and AIS data association determines whether the two tracks belong to the same target ship. The input form of the training network in the paper is based on AIS and radar sensor parameters to constitute the feature vector; the network learning function is to complete mining deep features of the vector, and the top of the SAE requires an additional regression layer for association.

After establishing the SAE, a training algorithm is required to obtain the SAE parameters. A backpropagation (BP) algorithm is a suitable choice because it is a common method used to train artificial neural networks in conjunction with optimization methods, such as batch gradient descent. BP calculates the gradient of the loss function considering all weights in the network. The AIS and radar track association model is shown in in Figure 3.

![Track association model](image)

Each node input in the model corresponds to an element of the feature vector \( v' \) and adds a constant term node +1 to the input node. According to different forecasting requirements, the number of target ships measured by the AIS and radar is \( M \) and \( N \), respectively, as the number of input nodes. The training set \( v = \{ v_1, v_2, \ldots, v_G \} \) is used for the learning network of the SAE, where each sample of \( v_x \) is a sampled sample and \( G \) is the total sample size. The model is composed of vector \( \{ v', H_1, H_2, \ldots, H_6, y \} \), where \( v' \) is a pre-input vector in the input of the model, \( H_1, H_2, \ldots, H_6 \) is the hidden layer of the SAE and \( y \) is the output layer of the BP algorithm. The purpose of this part is to determine whether a certain target on radar is associated with a certain AIS target and is a binary judgment at a certain time. Therefore, it can be set as a single node \( y \) in the output layer of the BP algorithm.

After the pre-training of SAE, the top layer carries a BP algorithm to perform the parameter adjustment for label data; this method is more effective than the traditional neural network using the BP algorithm in performing the gradient descent adjustment. The main reason for the superior performance of this method is that after the pre-training of the SAE is complete, its parameters are similar to the training values. Then, the BP algorithm is used to fine-tune the parameters, creating the input layer, hidden layer, and output layer of the SAE; a supervised learning algorithm is used to further adjust the pre-trained neural network. All of the weights and biases can be optimized after many iterations, and the training and convergence speed are relatively fast. The training steps for the model are as follows.

**Step 1:** Because there are differences among the dimensions and numeric range of features, some appropriate transformations of vector \( v \) should be carried out before the input of the deep learning...
network, which is called data standardization or normalization; the standard deviation method is employed for normalization.

Take the input vector \( R \_\text{Dis} = \{d_1, d_2, \ldots, d_n\} \) as an example. First, \( R \_\text{Dis} \) should be standardized by Equation (9).

\[
\hat{d}_i = \frac{d_i - \bar{d}}{s} \quad \ldots (9)
\]

where \( d_i \) is the \( j \)-th sample value, \( \bar{d} = \frac{1}{n} \sum_{i=1}^{n} d_i \) denotes the mean of elements in \( R \_\text{Dis} \), and 
\( s = \sqrt{\frac{1}{n-1} \sum_{i=1}^{n} (d_i - \bar{d})^2} \) is the standard deviation of elements. Next, \( R \_\text{Dis} \) should be handled by the method of maximum difference formalizations as 
\[
d'_i = \frac{\hat{d}_i - \min(d_1, d_2, \ldots, d_n)}{\max(d_1, d_2, \ldots, d_n) - \min(d_1, d_2, \ldots, d_n)} \quad \ldots (10)
\]

Finally, a sample of the input nodes of the neural network is \( R \_\text{Dis}' = \{d'_1, d'_2, \ldots, d'_n\} \); and other input vectors, such as \( A \_\text{Cog}_k \), \( A \_\text{Sog}_k \), \( A \_\text{Lon}_k \), \( A \_\text{Lat}_k \), \( R \_\text{Cog}_k \), \( R \_\text{Sog}_k \), \( R \_\text{Dis}_k \), and \( R \_\text{Dea}_k \), are transformed in a similar manner. The normalized input vector \( v' \) can be expressed as \( v' = [A \_\text{Cog}, A \_\text{Sog}, A \_\text{Lon}, A \_\text{Lat}, A \_\text{Lon}, A \_\text{Wid}, A \_\text{Typ}, R \_\text{Cog}, R \_\text{Sog}, R \_\text{Qua}, R \_\text{Dis}, R \_\text{Dea}, W_i, W_0, W_f, W_r] \).

Step 2: Taking vector \( v' \) as the input parameter, the weights are random in the initial network, the weight from \( v' \) to the first \( h_1 \) hidden layer is \( w_1 \), the threshold is \( b_1 \), and \( v' \) is brought into the network as
\[
z_i = w_i \cdot v' + b_i \quad \ldots (11)
\]

The value obtained is brought into the activation function Sigmoid\(^{35}\), as shown in equation (12). The function will be a transform kernel function of hidden layers.

\[
f (z_i) = \frac{1}{1 + \exp(-z_i)} \quad \ldots (12)
\]

where \( f (z_i) \) is the output matrix value of the hidden layer, the weights from the first \( h_1 \) hidden layer to the second \( h_2 \) hidden layer are \( w_2 \), the threshold is \( b_2 \), the output value matrix \( z_2 \), and the other hidden layers can be handled in a similar manner.

Step 3: The output weight of the last \( h_n \) hidden layer is \( w_n \), the threshold is \( b_n \), and the output value is \( z_n \). \( \{h_1, h_2, \ldots, h_n\} \) is repeated until a given number of layers is completed and as an input of the top-level regression layer of the BP algorithm. \( y_j^k = z_n + 1 \) is an output of the BP algorithm that represents the track association judgment of the \( i \)-th radar target and \( j \)-th AIS target at time \( k \).

Step 4: Using the supervised BP algorithm, the weights \( w_1, w_2, \ldots, w_n \) and threshold \( b_1, b_2, \ldots, b_n \) are fine-tuned by the constructed cost function \( f_{\min}(w, b) \), as shown in Equation (13), until the changes are sufficiently small to be neglected.

\[
f_{\min}(w, b) = f(w, b) + \beta \sum_{j=1}^{s} \left( \rho \cdot \hat{p}_j \right) \quad \ldots (13)
\]

\[
f(w, b) = \frac{1}{G} \sum_{i=1}^{G} \left( \frac{1}{2} \sum_{h_k} \left( x_{h_k} - x_{h_k}' \right)^2 \right) + \frac{\lambda}{2} \sum_{i=1}^{G} \sum_{j=1}^{s} \left( w_i^j \right)^2
\]

where \( \beta \) is the control weight, \( \sum_{j=1}^{s} \left( \rho \cdot \hat{p}_j \right) \) is the constraint term, \( s \) is the number of nodes in the \( h_i \)-hidden layer, \( \rho \) is a parameter that is typically taken as a value close to 0, such as \( \rho = 0.05 \), \( \hat{p}_j \) denotes the average activity degree of the \( h_i \)-hidden layer node, \( \frac{1}{G} \sum_{i=1}^{G} \left( \frac{1}{2} \sum_{h_k} \left( x_{h_k} - x_{h_k}' \right)^2 \right) \) is a reconstruction error term, \( \frac{\lambda}{2} \sum_{i=1}^{G} \sum_{j=1}^{s} \left( w_i^j \right)^2 \) is a regular expression used to prevent over-fitting, and \( x_{i,j} \) denotes the \( i \)-th node parameter of the \( h_i \)-hidden layer.

Because the number of layers and nodes in the hidden layer are uncertain, the number of nodes used in the hidden layer is smaller than that in the input layer, and the nodes in each layer are reduced gradually. The training model can be used as a model of track association. When a set of input vectors are given, the associated results of a corresponding ship object are obtained.

Track fusion method

In the problem of track fusion, a label learning sample set must be provided as the reference of the learning machine. The classifier can classify the new input feature vector sample after learning to obtain the ability. This approach is superior to logistic regression because logistic regression is suitable for two
classification problems. The learning feature of the SAE network can be used to replace the original vector input with the classifier for pattern classification, which can greatly increase the classification accuracy.

To make the SAE have the function of classifying and recognizing, the classifier must be appended after completing the output layer of the neural network. The softmax regression will be employed at the top layer of SAE in the paper; it is a classification method that generalizes logistic regression to multiclass problems and provides more choices for class labels. Neural network training can achieve the multi-perception of hierarchical feature extraction and data classification to establish the track classification and fusion model. The basic steps can be summarized as follows.

Step 1: Construct the training sets. In the softmax regression model, the target variable is divided into \( k \) classes, labels can be the different values \( y \in \{1,2,\ldots,k\} \), and a multinomial \( k \)-type distribution is used. According to prior knowledge, class labels \( y^{(i)} \in \{1,2,3,4\} \) are set for each vector, which represents the four output states of distance, relative speed and course, and position. Thus, the label learning vector is obtained as sets \( L = \{ (x^{(1)}, y^{(1)}), \ldots , (x^{(m)}, y^{(m)}) \} \).

Step 2: Solving classifier. For a given set of training samples \( L \), a hypothesis function can be adopted to estimate the probability value \( p(y = i | x) \) for each class \( i \). The function outputs a vector of \( k \) dimensions, as follows

\[
h_{\theta}(x) = \begin{bmatrix} p(y^{(i)} = 1 | x^{(i)}; \theta) \\ p(y^{(i)} = 2 | x^{(i)}; \theta) \\ \vdots \\ p(y^{(i)} = k | x^{(i)}; \theta) \end{bmatrix} = \frac{1}{\sum_{j=1}^{k} e^{	heta_j x^{(i)}}} \begin{bmatrix} e^{	heta_1 x^{(i)}} \\ e^{	heta_2 x^{(i)}} \\ \vdots \\ e^{	heta_k x^{(i)}} \end{bmatrix} \quad \cdots (14)
\]

where \( \frac{1}{\sum_{j=1}^{k} e^{	heta_j x^{(i)}}} \) normalizes the probability distribution such that the sum of all probabilities is 1. \( \theta \) is a model of a required matrix, and each row of the matrix can be seen as a category corresponding to the classifier parameters, where there are a total of \( k \) lines. Therefore, \( \theta \) can be written as follows

\[
\theta = \begin{bmatrix} \theta_1^T \\ \theta_2^T \\ \vdots \\ \theta_k^T \end{bmatrix} \quad \cdots (15)
\]

Its neural network’s loss function is as follows

\[
J(\theta) = -\frac{1}{m} \sum_{i=1}^{m} \sum_{j=1}^{k} \{ y^{(i)} = j \} \log \frac{e^{	heta_j x^{(i)}}}{\sum_{j=1}^{k} e^{	heta_j x^{(i)}}} \quad \cdots (16)
\]

where \( 1 \{ \cdot \} \) is an indicator function; that is, the result of the function is 1 when the value in the brace is true and 0 otherwise.

The Newton method is used to solve the loss function. The partial derivative function of the loss function is obtained as follows

\[
\nabla_{\theta_j} J(\theta) = -\frac{1}{m} \sum_{i=1}^{m} \sum_{j=1}^{k} \{ y^{(i)} = j \} \left[ p(y^{(i)} = j | x^{(i)}; \theta) \right] \quad \cdots (17)
\]

where \( \nabla_{\theta_j} J(\theta) \) denotes the partial derivative of the loss function for the \( i \)-th parameter of the \( j \)-th class.

Because there is more than one optimization for the parameter in softmax, a rule item will be added as follows

\[
J(\theta) = -\frac{1}{m} \sum_{i=1}^{m} \sum_{j=1}^{k} \{ y^{(i)} = j \} \log \frac{e^{	heta_j x^{(i)}}}{\sum_{j=1}^{k} e^{	heta_j x^{(i)}}} + \frac{\lambda}{2} \sum_{j=1}^{k} \theta_j^2 \quad \cdots (18)
\]

Its partial derivative function is expressed as follows

\[
\nabla_{\theta_j} J(\theta) = -\frac{1}{m} \sum_{i=1}^{m} \sum_{j=1}^{k} \{ y^{(i)} = j \} \left[ p(y^{(i)} = j | x^{(i)}; \theta) \right] + \lambda \theta_j \quad \cdots (19)
\]

where \( \lambda \) will be obtained by the maximum likelihood method and the parameter \( \theta \) is determined to yield the final track classification fusion model.

**Results and Discussion**

In this section, the above fusion methods based on deep learning will be verified by simulation experiments with data from the Yangtze Maritime Bureau. The AIS and radar information within a certain period of time is extracted. First, the data are pre-processed and standardized for the ship’s longitude, latitude, speed, course, length, width, type,
relative location, and distance and other track information in the Chongqing City waterway of the Yangtze River. The given ship (MMSI:713825613) (Figure 4) is selected as a reference point, and other ships’ information from AISs and/or radar are acquired for training and verification.

Data samples and association
A portion of the target ships’ data has been collected from AIS and radar equipment, as shown in Tables 2 to 5. The location of the own ship to be the coordinate origin is N 29.153990 E 106.182116, the speed is 0, the course is 67°, and its type is unknown. A portion of the data for target ship 2 collected by radar or AIS is shown in Tables 2 and 3, respectively, and similar information has been acquired from otherships.

Target ship 5 will be simulated and discussed, that is, whether ships 1 to 4 are associated with ship 5; if one of ships 1 to 4 is associated with ship 5, the tracks will be fused.

Figure 5 illustrates that the association factor between ship 5 and the other ships can be obtained. The association factor of AIS target ship 5 and radar target ship 3 is the largest. Although the positions of ship 5 and ship 1 are close to that of ship 3, the association factors are extremely low, with an average value of approximately 0.12, because the ship type in the input vector had been added, where ship 5 is a cargo ship and ship 1 is a passenger ship. If the factor value is greater than 0.93, two ships can be judged as associated; because the mean value of the non-association factor is less than 0.5, thus the AIS information of ship 5 and radar information of ship 3 should be fused.

In Figure 6 (1), (2), (3), and (4), the AIS and radar distance, speed, course, and position have been fused. The simulation illustrates that the fusion target’s information is smoother than the target ships’ AIS or radar information. In Figs. 6 (1), (2), and (3), the horizontal axis represents time and the vertical axis represents the association relationship after fusion. The fused Cartesian coordinates are represented by X-Y, and the other axis is still time in Fig. 6 (4). Furthermore, the fusion information

Table 2 — A portion of the sample data on ship 2 from radar

<table>
<thead>
<tr>
<th>NO</th>
<th>Sampling Time</th>
<th>Distance(km)</th>
<th>Location( )</th>
<th>Speed(kts)</th>
<th>Course( )</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2015/1/11 15:22</td>
<td>0.983</td>
<td>36.445</td>
<td>23.56</td>
<td>11.21</td>
</tr>
<tr>
<td>2</td>
<td>2015/1/11 15:25</td>
<td>0.967</td>
<td>36.728</td>
<td>24.85</td>
<td>12.23</td>
</tr>
<tr>
<td>3</td>
<td>2015/1/11 15:28</td>
<td>0.950</td>
<td>36.185</td>
<td>24.10</td>
<td>10.89</td>
</tr>
<tr>
<td>4</td>
<td>2015/1/11 15:31</td>
<td>0.932</td>
<td>36.804</td>
<td>22.55</td>
<td>12.24</td>
</tr>
</tbody>
</table>

Table 3 — A portion of the sample data on ship 2 (MMSI: 413774233) from AISs

<table>
<thead>
<tr>
<th>NO</th>
<th>Sampling Time</th>
<th>Longitude(N)</th>
<th>Latitude(E)</th>
<th>Speed(kts)</th>
<th>Course( )</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2015/1/11 15:22</td>
<td>106.592858</td>
<td>29.616435</td>
<td>21.67</td>
<td>62.2</td>
</tr>
<tr>
<td>2</td>
<td>2015/1/11 15:24</td>
<td>106.594165</td>
<td>29.617087</td>
<td>22.04</td>
<td>61.4</td>
</tr>
<tr>
<td>3</td>
<td>2015/1/11 15:26</td>
<td>106.595147</td>
<td>29.617585</td>
<td>21.85</td>
<td>60.7</td>
</tr>
<tr>
<td>4</td>
<td>2015/1/11 15:28</td>
<td>106.596065</td>
<td>29.618013</td>
<td>20.37</td>
<td>63.4</td>
</tr>
</tbody>
</table>

Table 4 — A portion of the sample data on ship 4 from radar

<table>
<thead>
<tr>
<th>NO</th>
<th>Sampling Time</th>
<th>Distance(km)</th>
<th>Location( )</th>
<th>Speed(kts)</th>
<th>Course( )</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2015/1/11 15:22</td>
<td>0.158</td>
<td>36.103</td>
<td>23.56</td>
<td>150.23</td>
</tr>
<tr>
<td>2</td>
<td>2015/1/11 15:25</td>
<td>0.175</td>
<td>36.185</td>
<td>23.61</td>
<td>151.43</td>
</tr>
<tr>
<td>3</td>
<td>2015/1/11 15:28</td>
<td>0.201</td>
<td>36.245</td>
<td>23.56</td>
<td>150.85</td>
</tr>
<tr>
<td>4</td>
<td>2015/1/11 15:31</td>
<td>0.220</td>
<td>36.250</td>
<td>24.01</td>
<td>150.98</td>
</tr>
</tbody>
</table>

Table 5 — A portion of the sample data on ship 5 (MMSI: 413829077) from AISs

<table>
<thead>
<tr>
<th>NO</th>
<th>Sampling Time</th>
<th>Longitude(N)</th>
<th>Latitude(E)</th>
<th>Speed(kts)</th>
<th>Course( )</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2015/1/11 15:22</td>
<td>106.60547</td>
<td>29.621018</td>
<td>15.93</td>
<td>85.1</td>
</tr>
<tr>
<td>2</td>
<td>2015/1/11 15:24</td>
<td>106.608203</td>
<td>29.621135</td>
<td>15.93</td>
<td>87.4</td>
</tr>
<tr>
<td>3</td>
<td>2015/1/11 15:26</td>
<td>106.609558</td>
<td>29.621125</td>
<td>15.74</td>
<td>92.2</td>
</tr>
<tr>
<td>4</td>
<td>2015/1/11 15:28</td>
<td>106.614998</td>
<td>29.620478</td>
<td>15.93</td>
<td>107.6</td>
</tr>
</tbody>
</table>
Fig. 5 — Association factor between ship 5 and the other ships

Fig. 6 — Fusion of the distance, speed, course, and position from AIS and radar information

The association factor between ship 5 and the other ships is shown in Figure 5. The association factor is calculated based on the distance, speed, course, and position of the ships. The figure shows that the association factor is higher for ships that are closer in distance, speed, course, and position.

In Figure 6, the fusion of the distance, speed, course, and position from AIS and radar information is depicted. The fusion process is used to improve the accuracy of the track information. The tracks from AIS and radar are combined to create a more accurate and reliable track.

Track fusion and timeliness analysis

In Figure 7 (1), (2), and (3), green denotes the AIS track, purple denotes the radar track, and red denotes the track after fusing. The arrows symbol “” in the red line indicates the course of the ship. The red fused track is smoother and more in line with the state of ship navigation. Figure 7 (1) presents the track fusion of AIS target ship 5 and radar target ship 3, called the tracking of ship A_R35. In Figure 7(2), the red track of ship 2 is smoother than the red track of ship A_R35, and the lines of 3 colors are correspondingly more compact relative to ship A_R35. Along the long lines, ship 2 is moving in the Yangtze River and stays in the downstream state, as in Fig. 7(3). Through the fusion of more than 400,000 AIS and radar ship data points on January 11, 2015, ships in the downstream course are chosen to sample, and these tracks on the AIS and/or radar are fused by using a red line. A large number of ship traces are fused to derive a ship’s track; an entire track is drawn by using a black line, which corresponds to the actual navigation in the waterways of the Yangtze River, as shown in Figure 7(4). The fusion accuracy is 35% higher than that of radar and 8% higher than that of the AIS. The effectiveness and correctness of our model are verified by extensive simulations, which can improve the safety of ship navigation.

Over 200 ships were sampled from more than 2 GB of data per day and 250,000 pieces of AIS and radar
The training time on half a day of data is listed in Table 6, where the training times of deep learning, the BP algorithm and softmax regression are given under the parameter (\( \rho = 0.05 \)).

The neural network established in this paper uses the deep learning algorithm, which considerably shortens the learning time compared with the BP algorithm and softmax regression.
Table 6 — Comparison on Learning Way and Time

<table>
<thead>
<tr>
<th>Learning way</th>
<th>Deep learning SAE</th>
<th>BP</th>
<th>Deep learning SAE</th>
<th>Softmax</th>
<th>Softmax</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of iterations</td>
<td>25</td>
<td>10</td>
<td>105</td>
<td>18</td>
<td>8</td>
</tr>
<tr>
<td>Learning time(s)</td>
<td>52.3</td>
<td>104.2</td>
<td>879</td>
<td>37.8</td>
<td>80.6</td>
</tr>
</tbody>
</table>

Conclusions

According to the features of AISs and radar, a new model based on deep learning is proposed for target ship association and fusion. Selecting the characteristic vector of ship fusion, the data are preprocessed by selection and the standard deviation method. The SAE is combined with the BP algorithm and softmax, the association factors between ships are realized by fine-tuning parameters, and the traces are fused by adding rule items to softmax. The experimental results show that the model is effective for track fusion and exhibits improved accuracy and a reduced learning time compared with the BP algorithm and softmax. The fusion of AIS and radar information can provide more accurate perceived services of ship traffic. In the future, association and fusion will be processed for ship navigation in real time.

Acknowledgement

The work was partially supported by National Natural Science Foundation of China(Grant No.51479155), Natural Science Foundation of Hubei Province of China(Grant No.2014CFB190), Science and Technology Department of Bijie City of China(Grant No.2014CFB190), and Fundamental Research Foundation of the Central Universities(Grant No. 16ZY006).

References


