Research Article

Multilevel Association Rule Mining for Bridge Resource Management Based on Immune Genetic Algorithm

Yang Ou,[1] Zheng Jiang Liu,[1] Hamid Reza Karimi,[2] and Ying Tian[3]

[1] *College of Navigation, Dalian Maritime University, Dalian 116026, China*
[2] *Department of Engineering, Faculty of Engineering and Science, The University of Agder, 4898 Grimstad, Norway*
[3] *College of Engineering, Bohai University, Jinzhou 121000, China*

Correspondence should be addressed to Ying Tian; tianyingjzh@163.com

This paper is concerned with the problem of multilevel association rule mining for bridge resource management (BRM) which is announced by IMO in 2010. The goal of this paper is to mine the association rules among the items of BRM and the vessel accidents. However, due to the indirect data that can be collected, which seems useless for the analysis of the relationship between items of BIM and the accidents, the cross level association rules need to be studied, which builds the relation between the indirect data and items of BRM. In this paper, firstly, a cross level coding scheme for mining the multilevel association rules is proposed. Secondly, we execute the immune genetic algorithm with the coding scheme for analyzing BRM. Thirdly, based on the basic maritime investigation reports, some important association rules of the items of BRM are mined and studied. Finally, according to the results of the analysis, we provide the suggestions for the work of seafarer training, assessment, and management.

1. Introduction

With the development of shipping science, improvement of vessel technology, and the rising trend of multinational manning of ships, maritime safety and marine pollution prevention have been the hot spots in study of the maritime field by all the countries. Against this background, the focuses of maritime safety research are mainly on two aspects. One is the reasonable use of marine technology resources, which is termed as the seamanship. The other is the proper training of the communication and cooperation in the sailing team, which is named as the teamwork in the ship.

In fact, the good seamanship and the excellent teamwork in the deck crew management of the ship are both the most important factors to avoid the vessel accidents. The previous researches about the ship accidents show that more than 80% accidents are caused by human errors, which are inseparable with the level of seamanship and the teamwork. Therefore, in 2010, the IMO announced the sweeping amendments of the STCW convention, named STCW convention Manila

amendments. In the convention, marine resource management ability, which is the collectivity name of seamanship and the teamwork, becomes one of the mandatory requirements, and both bridge resource management (BRM) and engine resource management (ERM) are proposed formally. However, STCW Manila amendments only explain the marine resource management qualitatively. So far, a few literatures focus on the data analysis between the vessel accidents and items of BRM. Paper [1] promoted the rating model for seafarer based on BRM, but it gave the importance of items of BRM by the expert questionnaire without data analysis. Therefore, it is necessary to find some associations, which can reflect the deep inducements of accidents, from the collected maritime accident data to guide the seafarer training.

Generally speaking, a large amount of maritime accident data can be collected from the actual vessel accident; association rule mining, which is a kind of data mining, is a feasible method to analyze this data. During recent years many researches about the association rule mining are proposed [2–8]. Literature [2] is proposed to integrate

classification and association mining techniques according to the target of discovery are not predetermined in association rule mining. The integration is done by focusing on mining a special subset of association rules, called class association rules (CARs). An efficient algorithm is also given for building a classifier based on the set of discovered CARs. In [3], an efficient representation scheme, called general rules and exceptions, is proposed. This paper focused on using the representation to express the knowledge embedded in a decision tree. In this representation, a unit of knowledge consists of a single general rule and a set of exceptions. This scheme reduces the complexity of the discovered knowledge substantially. It is also intuitive and easy to understand. Authors in literatures [4, 5] deal with the algorithmic aspects of association rule mining, and they explain the fundamentals of association rule mining and moreover derive approaches. Paper [6] presents effective techniques for pruning the search space when computing optimized association rules for both categorical and numeric attributes. However, to mine the association rules effectively, it needs to tackle the problems on the very large search space of candidate rules during the rule discovery process. Some intelligent searching methods are used in the rule discovery process, such as literatures [7, 8]. The paper [7] seeks to generate large item sets in a dynamic transaction database using immune clone algorithm, where intratransactions, intertransactions, and distributed transactions are considered for mining association rules. The paper [8] proposes an approach, which is based on artificial immune systems, for mining association rules effectively for classification. Instead of massively searching for all possible association rules, this approach only finds a subset of association rules that are suitable for system effective in an evolutionary manner. However, producing too many rules have been regarded as a major problem with many data mining techniques. The large number of rules makes it very hard for manual inspection of the rules to gain insight of the domain. Many data mining methods mentioned above typically generate a list of rules with no further organization. It is like writing a book or a paper by randomly listing out all the facts and formulas. The resulting book will be difficult for anyone to read because one cannot see any relationships among the facts and formulas. In papers [9, 10], they proposed a technique to intuitively organize and summarize the discovered rules. With this organization, the discovered rules can be presented to the user in the way as the thinking and talking about knowledge in the daily lives. A common strategy for organizing the large number of rules is by hierarchy. In such an organization, the rules are summarized into topics and presented at different levels of details. The main advantage of the hierarchical organization is that it allows the user to view the problem from general to specific and to focus his/her attention on those interesting aspects. Papers [9, 10] manage the complexity of the discovered rules in a hierarchical fashion, but these methods are too complex to deal with the maritime accident data. In our paper, we use the hierarchical thinking, organizing lots of items into multiple levels by the cross level coding. Additionally, the immune genetic algorithm is used in the rule mining process. By this way, the major association rules, which reflect deep inducements of accidents, will be mined effectively in an evolutionary manner.

This paper mainly researches BRM. Through the analysis of the association rules between items of BRM and vessel accidents, the importance of BRM are evaluated quantitatively and the relationship of its items are mined. The contributions of this paper are (1) proposing a cross level coding scheme for mining the multilevel association rules; (2) presenting the immune genetic algorithm using our cross level coding scheme for the research of BRM; (3) finding the important association rules of the items of BRM based on the basic maritime investigation reports; (4) providing suggestions to alleviate the incentive of shipwreck in aspect of seafarer training, assessment, and management.

The rest of this paper is organized into the following sections. In Section 2, our multilevel association rule mining based on immune genetic algorithm is proposed. In Section 3, the proposed approach is simulated and analyzed; then the feasibility of the approach is illustrated by comparing with other methods; some suggestions are developed toward the mining results at the last part of this section. In Section 4, we conclude our work.

2. Multilevel Association Rule Mining Based on Immune Genetic Algorithm

In this paper, we analyze items of BRM based on immune genetic algorithm. The goal is to mine the association rules among the items of BRM and the vessel accidents. However, due to the indirect data that can be collected, the cross level coding scheme needs to be proposed. In this section, a multilevel association rule mining will be introduced, where the basic data is considered as the bottom of the data mining, and the summarized data, that is, items of BRM in this paper, is the middle level, which has certain corresponding relationship with the basic data.

2.1. Multilevel Association Rule Mining. Association rule is the hiding relationship among the data items, which is the relevance of different items appearance in the same event [1]. Let $I = \{x_1, x_2, \ldots, x_m\}$ be a set of distinct literals, called items. A set $X \subseteq I$ with $k = |X|$ is called a k-item set or simply an item set. Let a database D be a multiset of subsets of I. Each $T \in D$ is called a transaction. We say that a transaction $T \in D$ supports an item set $X \subseteq I$ if $X \subseteq T$ holds. An association rule is an expression $X \Rightarrow Y$, where X and Y are item sets and $X \cap Y = \emptyset$ holds. The fraction of transactions T supporting an item set X with respect to database D is called the support of X, $\mathrm{supp}(X) = |\{T \in D \mid X \subseteq T\}|/|D|$. The support of a rule $X \Rightarrow Y$ is defined as

$$\mathrm{supp}\,(X \Longrightarrow Y) = \mathrm{supp}\,(X \cup Y). \tag{1}$$

The confidence of this rule is defined as

$$\mathrm{conf}\,(X \Longrightarrow Y) = \frac{\mathrm{supp}\,(X \cup Y)}{\mathrm{supp}\,(X)}. \tag{2}$$
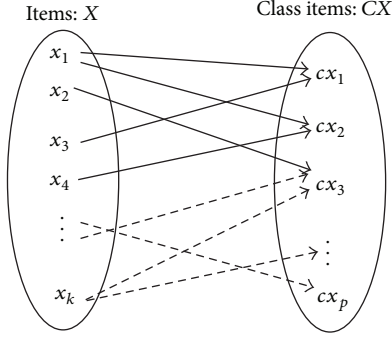
FIGURE 1: Mapping relationship of items and class items.

In our work, for a hierarchical organization to work effectively and feasibility, numbers of items in database are summarized into class items as the higher level of the information, which are interested by the user. Then the items $I = \{x_1, x_2, \ldots, x_m\}$ are mapped into the class items $CI = \{cx_1, cx_2, \ldots, cx_q\}$, where q is the number of class items with $q < m$. A set $CX \subseteq CI$ with $p = |CX|$ is called a p-class item set or simply a class item set. Figure 1 shows us the mapping relationship of items and class items. Generally, there is more than one item corresponding to one class item. In our paper, binary coding scheme is used. Each item will be 0 or 1 in our scheme. The link of items with different values forms a basic antibody. The antibody is formed by the link of class items with different values. The value of class item is decided by the product of the corresponding cross level weights and the items which compose the class item. Set X is a k-item set, where each item is represented by x_i, $x_i = \{0, 1\}$. CX is a p-class item set containing the class item cx_j. The contribution of x_i to cx_j is α_{ij}, called cross level weight, where $\sum_{j=1}^{n} \alpha_{ij} = 1$, n is the total number of class items affected by the item x_i. As shown in Figure 1, items x_1 and x_3 compose the class item cx_1; the corresponding cross level weights are α_{11} and α_{31}, while items x_1 and x_4 compose the class item cx_2; the corresponding cross level weights are α_{12} and α_{41}, where $\alpha_{11} + \alpha_{12} = 1$ and $\alpha_{31} = \alpha_{41} = 1$. The value of cx_j is calculated using the following formula:

$$cx_j = \left[\frac{\sum_{i=1}^{M} \alpha_{ij} x_i}{\left(\sum_{p=1}^{K} \alpha_{pq} x_p\right)_{\text{mas}}} \times L \right], \quad (3)$$

where L is the bit size of cx_j, M is the number of items who form the class item, $\left(\sum_{p=1}^{K} \alpha_{pq} x_p\right)_{\text{max}}$ is the maximum sum value of the class item, where K is the number of items contained by the class item. In formula (3), the symbol $[\cdot]$ represents the round calculation. The code of cx_j can be obtained by converting decimal of cx_j into binary number.

Definition 1 (information entropy). Because the immune system in the process of evolution is the uncertain one which is composed by antibodies, the irregular degree of the immune system can be presented by average information entropy of Shannons.

TABLE 1: An example of basic antibody.

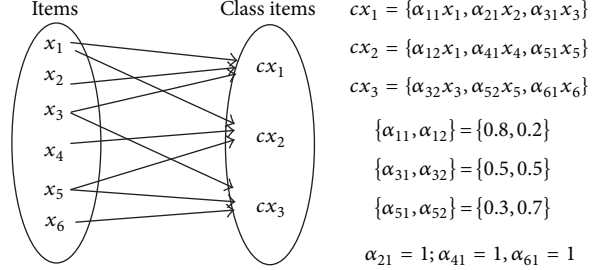| Basic antibody	x_1	x_2	x_3	x_4	x_5	x_6
$X(1)$	1	1	0	0	0	1
$X(2)$	0	1	1	0	1	1

FIGURE 2: An example of mapping relationship with 6 items and 3 class items.

Assuming an immune system is consisted with N antibodies, the number of gene for each antibody is M. The S numerals is used on encoding each gene, where the binary is used in this paper; that is, $S = 2$, and the ith code $S_i \in \{0, 1\}$, where $i = 1, 2$; then the information entropy of these N antibodies is as follows:

$$H(N) = \frac{1}{M} \sum_{j=1}^{M} H_j(N), \quad (4)$$

where $H_j(N) = -\sum_{i=1}^{S} p_{ij} \log_2 p_{ij}$, defined as the information entropy of the jth gene; p_{ij} is the probably of the ith code appeared in the jth gene.

In order to explain the cross level coding scheme, an example of cross level coding from the basic antibody to antibody is shown as follows. The two complete basic antibodies are given in Table 1. As shown in Figure 2, we summarized the 6 items into 3 class items, the cross level weights are also given. When $cx_1 = \{\alpha_{11}x_1, \alpha_{21}x_2, \alpha_{31}x_3\}$ and $\{\alpha_{11}, \alpha_{21}, \alpha_{31}\} = \{0.8, 1, 0.5\}$, the corresponding fragments of the basic antibodies $X(1)$ and $X(2)$ are $\{0.8, 1, 0\}$ and $\{0, 1, 0.5\}$, respectively. Similarly, the results can be derived as $\{cx_1, cx_2, cx_3\} = \{1.8, 0.2, 1\}$ for the antibody $CX(1)$ and $\{cx_1, cx_2, cx_3\} = \{1.5, 0.3, 2.2\}$ for the antibody $CX(2)$. Then using the formula (3), where $L = 2$, $M = 3$, and cx_1 has the maximum sum value among all of the class items in $CX(1)$, then we have the equation:

$$\left(\sum_{p=1}^{K} \alpha_{pq} x_p\right)_{\text{max}} = \alpha_{11}x_1 + \alpha_{21}x_2 + \alpha_{31}x_3 = 1.8, \quad (5)$$

where $K = 3$. While cx_3 has the maximum value of the same one in $CX(2)$, then we have the equation:

$$\left(\sum_{p=1}^{K} \alpha_{pq} x_p\right)_{\text{max}} = \alpha_{32}x_3 + \alpha_{52}x_5 + \alpha_{61}x_6 = 2.2. \quad (6)$$

Finally, taking both values of (5) and (6) into the formula (3), the fragments of the antibodies $CX(1)$ and $CX(2)$ are coded

TABLE 2: The corresponding antibody of Table 1.

Antibody	cx_1	cx_2	cx_3
$CX(1)$	11	00	10
$CX(2)$	11	00	11

as {1, 1} and {1, 1}, which are shown in the second column of Table 2. The corresponding cross level codes of antibodies are shown in Table 2.

After the cross level coding, the mining of association rules can be executed on the encoded population based on the basic data. However, frequent patterns meeting the support requirements need to be found for association rule mining. When there are multiply levels and dimensions, lots of thresholds should be defined. Meanwhile, the search space, storage space, and running time will increase with the increasing of the complexity of algorithm [11, 12]. Some intelligent algorithms such as genetic algorithm and artificial immune algorithm can make the rule mining process more efficient.

2.2. Immune Genetic Algorithm for Association Rule Mining. Genetic algorithm (GA) is a search heuristic that mimics the process of natural selection, which generates solutions to optimization problems using techniques inspired by natural evolution, such as inheritance, mutation, selection, and crossover. In GA, a population of candidate solutions to an optimization problem is evolved towards better solutions. Each candidate solution has a set of properties which can be mutated and altered. The evolution usually starts from a population of randomly generated individuals and is an iterative process, with the population in each iteration called a generation. In each generation, the fitness of every individual in the population is evaluated, and the fitness is usually the value of the objective function in the optimization problem being solved. However, literature [13] notes the drawbacks of GA such as the convergence direction is unable to control and the object function does not correspond to constraints and does not have memory function.

Artificial immune systems (AIS) are concerned with abstracting the structure and function of the immune system to computational systems and investigating the application of these systems towards solving computational problems from mathematics, engineering, and information technology [14–17]. Considering the characteristics of both GA and AIS, many researchers have proposed the immune genetic algorithm (IGA) with immune function [7, 18]. The aim of leading immune concepts and methods into GA is theoretically to utilize the locally characteristic information for seeking the ways and means of finding the optimal solution when dealing with difficult problems. To be exact, it utilizes the local information to intervene in the globally parallel process and restrain or avoid repetitive and useless work during the course, so as to overcome the blindness in action of the crossover and mutation.

Based on IGA, the association rule mining algorithm for BRM is researched in this paper. We design a multilevel association rule mining algorithm especially for BRM using IGA, which is called IGA-BRM for simplicity. There are some definitions listed below about IGA-BRM algorithm.

Definition 2 (antigen). Antigen is the molecule, which can be identified by any antibody or T cell receptors, corresponding to the problem. Usually, different problems have different forms of data encoding. In this paper, it will be the attribute values which we are interest in.

Definition 3 (antibody). Antibody is the receptor which is produced by B lymphocytes that react with antibodies. Usually, the antibody corresponds to the optimal solution to the problem. In this paper, it refers to association rules that will be mined.

Definition 4 (affinity). The strength of the bind between antibodies is represented by affinity. In AIS, the more similar antibodies will have the bigger values of the affinity. In order to control the density of antibodies, the new antibody is created only if the affinity between the two exceeds the affinity threshold.

The affinity between both antibodies can be calculated using (4), where $N = 2$. Both antibodies are similar when the affinity between both antibodies is beyond the affinity threshold λ, which is called the similarity constant, $0.9 \leq \lambda \leq 1$. The population similarity of antibodies also can be achieved using the same equation, where N is the number of antibodies in the population.

Definition 5 (fitness). The strength of the bind between the antigen and the antibody, which depends on how closely the two match, is called fitness. The more matching of the antibody and the antigen is, the stronger the molecular binding will be, and the better the antibody can be recognized. In IGA, the fitness indicates the adaptability of the candidate solutions of objective function to the problem.

Actually, different fitness functions are chosen according to the different problem. Therefore, fitness function is the key of convergence rate of the algorithm. Assuming the support of the antibody is Supp(A) and the confidence is Conf(A), the fitness of antibody A is as follows:

$$\text{fitness}(A) = w_s \text{Supp}(A) + w_c \text{Conf}(A). \qquad (7)$$

Definition 6 (antibody concentration). It is the proportion of the similar antibodies and the total number of antibodies in one generation. We use the symbol C_i to express the antibody concentration of the ith antibody.

With the development of the IGA, the antibodies in the populations are more and more similar, and the diversity of the antibodies is no longer maintaining the original level. In order to maintain the diversity of the antibodies, improve the global search ability, and avoid the immature convergence of the algorithm, the fitness and concentration of the generation should be controlled. The parameter called polymerization fitness is introduced to control the selected probability of antibody in the genetic process.

Definition 7 (polymerization fitness). It is the balance of antibody concentration and the fitness of antibody, which is expressed by the following equation:

$$\text{fitness}_p\left(A_i\right) = \text{fitness}^{\mu C_i}\left(A_i\right), \tag{8}$$

where μ is the correction factor of polymerization fitness and $\mu \leq 0$.

Based on the above definitions and cross level encoding scheme, steps of the IGB-BRM algorithm are described as follows.

Step 1. Initialize all the parameters such as the size of population N, number of generation T, cross level weight α_{ij}, and affinity threshold λ.

Step 2. Process the data in database including encoding the basic antibody and then converting to the antibody code using our cross level coding scheme.

Step 3. Build the initial antibody population with half randomly generated and the other half taken from the coded data in Step 2.

Step 4. Calculate the values of support, confidence, and fitness of each antibody using the coded data in Step 2.

Step 5. Produce the new antibodies; execute the GA algorithm which includes selection, crossover, and mutation.

Step 6. Calculate the affinity of the new antibodies, and then compare them with the corresponding thresholds λ.

Step 7. If the comparing result of Step 6 is smaller than the thresholds, it will execute Step 4, which means the antibodies satisfying the requirement of diversity of population; else it will go to the next step.

Step 8. Produce the other new antibodies which make the information entropy and the fitness satisfy the required thresholds.

Step 9. Calculate the concentration and polymerization fitness of the antibody population, and then select the N antibodies with the highest polymerization fitness as the new generation.

Step 10. If the count number of generation $t_{\text{count}} \geq T$, the optimal solution of antibodies will be achieved, and then the algorithm is over; else t_{count} adds 1 and goes to the Step 4 until it is over.

In the algorithm, the antibodies with the highest polymerization fitness will be selected as the new increasing antibodies when the antibody population is refreshed. It is the concentration based population refresh in IGA-BRM. From formula (8), when the concentration is maintained, the bigger the fitness is, the more the polymerization fitness will be, which will increase the selected probability of the antibody. While the fitness is unchanged, the polymerization fitness will decrease with the increasing of the antibody concentration, which will lead the selected probability to be smaller. This scheme will not only retain the antibodies which have excellent fitness, but also inhibit the ones whose concentration is too high.

3. Bridge Resource Management Analysis Based on IGA-BRM

Resent years, researches have shown that the composition of marine traffic is seaman, ship, and environment, while the human error is the main factor which causes the accident among them. In this paper, based on vessel collision accidents on the sea, we research the deep reasons induced by human errors. The goal of our work is to reduce the maritime collision accidents caused by human error in aspect of seafarer training, assessment, and management.

3.1. Model of IGA-BRM. Traditionally, the collision accidents of vessels are analyzed by the data from the maritime investigation reports. But the data does not reflect deep inducements which we want to guide seafarer training. For example, a collision happened which was caused by the improper lookout, when a duty officer is on duty while sailing. Maritime survey results showed that the accident is mainly due to the improper lookout which led to the vessel missing the prime time for the prevention of collision; that is, collision accident cannot be avoided despite using the avoidance measures and good seamanship when the vessel sailed into the distance to closest approach (DCPA). Usually, as the maritime competent authority, we want to make certain policy to reduce or even avoid the vessel collision accidents through analyzing and summarizing these reasons of accidents. However, in the surface of the maritime accident investigation report the only reason of the accident and the malpractice is classified as a human error accident, which lacks the analysis of the human error itself. In fact, throughout the numerous collision accidents caused by the "improper lookout," there are many causes hiding the "improper lookout." The author, who works in seafarers' competency examination management and maritime investigation for several years, found that the main factors of "improper lookout" may be due to the fatigue, poor awareness of risk situation caused by overconfidence and the bad work attitude, and so on. Based on the above reasons, IMO in STCW Manila amendments proposed concept of BRM, which summarized the causes of the accident into six aspects, called items of BRM in this paper, listed in Table 3. These items of BRM are only proposed and researched qualitatively in STCW Manila amendments. In order to make the relationship between these items and the causes collected in the daily maritime investigation clear, we sorted data of the causes into ten surface reasons of collision accident, which are listed in Table 4, formed data mining chain in this paper, and then analyzed the data using the proposed IGA-BRM algorithm quantificationally.

3.2. Simulation Results and Analysis. The data analyzed is chosen from the maritime investigation files in resent ten years in north China. There are about 1032 items. Among them, we randomly select 100 of ship collision accident

Table 3: Corresponding relationship between items in BRM and attribute.

Attribute	Items in BRM
cx_1	Team work abilities and working attitude
cx_2	Abilities of communication
cx_3	Decision-making capacity
cx_4	Pressure and fatigue
cx_5	Situational awareness and ability of risk assessment
cx_6	Educational background
cx_7	Severity of the accident

Table 4: Corresponding relationship between human errors and basic attributes.

Basic attribute	Human error
x_1	Improper lookout
x_2	Nautical instruments use error
x_3	Errors in judgment
x_4	Inopportune navigating
x_5	Improper communication
x_6	Inappropriate maneuvering
x_7	Not at safe speed
x_8	Disobey navigation rules
x_9	Unable to use good seamanship
x_{10}	Error in positioning and route plan

Table 5: Values of α_{ij} between human errors and items in BRM.

	x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8	x_9	x_{10}
cx_1	0.2	0.2	0	0.1	0	0.2	0.7	0	0.6	0.4
cx_2	0.2	0	0.1	0.1	0.4	0	0	0.7	0	0
cx_3	0.1	0.2	0.2	0.2	0	0.2	0	0	0	0
cx_4	0.4	0.5	0.3	0.4	0.5	0.3	0.1	0.1	0.2	0.5
cx_5	0.1	0	0.4	0.2	0	0.3	0.2	0.2	0.1	0
cx_6	0	0.1	0	0	0.1	0	0	0	0.1	0.1

investigation report and build a statistical table about the value of cross level weight, that is, α_{ij}, as shown in Table 5. The ratio of cross level weight is counted by asking the maritime investigation officer who takes part in the investigation of the accident directly or calling to the crew in the accident report.

In the simulation, the items of BRM are considered as attributes of the algorithm, and the ten surface reasons coming from the database are as the basic attributes which have certain contributions to the attributes. Additionally, to find the relationship between items of BRM and the vessel collision accidents, the severity of the accident is added to the attributes, which can be achieved from the maritime investigation reports, denoted as cx_7 in Table 3. It is expressed by the 2 bit binary code, where $cx_7 = 11$ represents the worst accident. We set the following: size of population $N = 50$, correction factor of polymerization fitness $\mu = 0.8$, generation number $T = 100$, affinity threshold $\lambda = 0.1$, support threshold is 0.2, and confidence threshold is 0.8. The IGA-BRM algorithm is executed ten times in Matlab 6.5. Through the simulations, about 27 association rules are mined, some interested rules with high support and confidence are listed in Table 6.

From the results in Table 6, the item of pressure and fatigue in BRM is the most important factor which leads to the worst or most serious accident, and team work abilities and working attitude are the second important factor that causes the serious accident, while the bad situational awareness and ability of risk assessment are the third one cause of the general accident.

In order to illustrate the reasonability of our method, we compare our results with the literature [1], which obtained the importance of BRM items by the expert questionnaire. In [1], the item of pressure and fatigue in BRM is the most important factor in the management capacity evaluation, situational awareness and ability of risk assessment are the second one, and team work abilities and working attitude are the third one. From both results, whether in the vessel accident analysis or management capacity evaluation, the importance of BRM items is basically consistent. It means that our data mining method for BRM analysis is feasible and reasonable. We should pay more attention to the items of pressure and fatigue, situational awareness, and team work abilities of BRM in the maritime management.

3.3. Suggestions. Comparing with the actual situation, the simulation results in Table 6 are reasonable. In terms of the item of pressure and fatigue in BRM, there are many factors that make the crew always in the state of pressure and fatigue, such as the long-term work and boring life in the navigation, the worst natural conditions when sailing in the sea, and the contract period signed with the shipping company of the crew being too long. It is difficult to release the mental stress that causes every crew to have certain psychological subhealth more or less. Facing these problems, many crews usually adopt the way of self-injury, such as avoidance, depression, and denial, which make the crew always in the worst mental state. In order to alleviate or solve these problems from the aspect of management, we provide five suggestions as below.

(1) Based on the Manila amendments, the item about rest time of crew needs to be built; the competent authority should control the rest time longer than the minimum requirement of STCW strictly. The Port State Control (PSC) should make the reasonable and feasible ways to protect the rest and relax right of the seaman.

(2) In terms of the vessel operators and owners, considering both the culture and character of the company, the control of the crew's rest time and education of mental health should be incorporated in the safety management system, in order to make the crew rest enough and be happy.

(3) According to the "pressure and fatigue," there are two ways to solve the problem. One is that IMO and the competent authority add the courses of mental health to the standard of competency ability examination

TABLE 6: Association rules and the meaning of them based on BRM.

Association rule with support and confidence	Meaning based on BRM
$cx_1 = 11 \Rightarrow cx_7 = 10$ [supp = 65.2%, conf = 88.1%]	The worst team work abilities and working attitude lead to the serious accident
$cx_4 = 11 \Rightarrow cx_7 = 11$ [supp = 69.3%, conf = 89.0%]	The worst pressure and fatigue lead to the worst accident
$cx_5 = 10 \Rightarrow cx_7 = 01$ [supp = 55.8%, conf = 86.3%]	The bad situational awareness and ability of risk assessment lead to the general accident
$cx_1 = 11 \cap cx_5 = 10 \Rightarrow cx_7 = 11$ [supp = 32.6%, conf = 82.1%]	The worst team work abilities and attitude added to bad situational awareness and ability of risk assessment lead to the worst accident
$cx_3 = 10 \cap cx_4 = 10 \Rightarrow cx_7 = 10$ [supp = 21.8%, conf = 85.4%]	The bad pressure and fatigue added to bad decision-making capacity lead to the serious accident
$cx_2 = 10 \cap cx_6 = 01 \Rightarrow cx_7 = 01$ [supp = 25.2%, conf = 81.9%]	The bad abilities of communication and poor educational background lead to the general accident
$cx_4 = 11 \cap cx_5 = 10 \Rightarrow cx_7 = 10$ [supp = 33.6%, conf = 81.9%]	The worst pressure and fatigue added to situational awareness and ability of risk assessment lead to the serious accident

for the crew, and the other is to modify the item of the minimum safety requirements for vessels, forcedly require the vessel, of which voyage exceeds a certain time, and provide the mental health counselor.

(4) The shipping company or the crew management company should authorize the master to coordinate the sailing team, in order to improve the ability of teamwork and enhance the obedience consciousness of the member. Coordinate content should focus on breakthrough teamwork's five obstacles, named "lack of trust," "fear of conflict," "lack of investment," "escape responsibility," "ignore result." The competent authority should evaluate the leadership, organization, and coordination capacity quantitatively and bring the ability into the master competency ability requirement.

(5) In order to enhance the crew's ability of dealing with crisis, maritime education should focus on the situational awareness. Navigation is the dynamic process, and there are so many emergencies which demand the crew to develop the excellent response ability. Therefore, it is necessary to add the operation of navigation simulator to the maritime education and increase the investment of the study of navigation simulator for the better simulation result.

4. Conclusion

In this paper, based on immune genetic algorithm, we propose a multilevel association rule mining for BRM announced by IMO in 2010, named IGA-BRM. The goal of this paper is to mine the association rules among the items of BRM and the vessel accidents.

Because of the indirect data that can be collected which seems useless for the analysis of the relationship between items of BIM and the vessel accidents, a cross level coding scheme for mining the multilevel association rules is proposed. By this way, based on basic maritime investigation data, the relationship between items of BRM and vessel accidents is built. We execute the IGA-BRM with the cross level coding scheme for the research of BRM. As a result, based on the basic maritime investigation reports, many important association rules about the items of BRM are mined. Among these association rules, we find that the item of pressure and fatigue in BRM is the most important factor which leads to the worst or most serious accident, and team work abilities and working attitude are the second important factor that causes the serious accident, while the bad situational awareness and ability of risk assessment are the third one that causes the general accident and so on. At last of the paper, according to the pattern that mined using IGA-BRM algorithm, five suggestions are provided for the work of seafarer training, assessment, and management on the maritime competent authority level.

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

References

[1] Y. L. Liu, "Rating model of resource management competency of seafarers," *Navigation of China*, vol. 36, no. 2, pp. 82–85, 2013.

[2] B. Liu, W. Hsu, and Y. Ma, "Integrating classification and association rule mining," in *Proceedings of the ACM Knowledge Discovery and Data Mining (KDD '98)*, pp. 80–86, 1998.

[3] B. Liu, M. Hu, and W. Hsu, "Intuitive representation of decision trees using general rules and exceptions," in *Proceedings of the 7th National Conference on Artificial Intellgience*, pp. 615–620, July 2000.

[4] J. Hipp, U. Guntezr, and G. Nakhaeizadeh, "Algorithms for association rule mining—a general survey and comparison," in *Proceedings of the 6th ACM International Conference on Knowledge Discovery and Data Mining*, Boston, Mass, USA, 2000.

[5] S. Yin, G. Wang, and H. Karimi, "Data-driven design of robust fault detection system for wind turbines," *Mechatronics*, 2013.

[6] R. Rastogi and K. Shim, "Mining optimized association rules with categorical and numeric attributes," *IEEE Transactions on Knowledge and Data Engineering*, vol. 14, no. 1, pp. 29–50, 2002.

[7] H. Mo and L. Xu, "Immune clone algorithm for mining association rules on dynamic databases," in *Proceedings of the 17th IEEE International Conference on Tools with Artificial Intelligence (ICTAI '05)*, pp. 202–206, Hong Kong, China, November 2005.

[8] T. D. Do, S. C. Hui, A. C. M. Fong, and B. Fong, "Associative classification with artificial immune system," *IEEE Transactions on Evolutionary Computation*, vol. 13, no. 2, pp. 217–228, 2009.

[9] B. Liu, M. Hu, and W. Hsu, "Multi-level organization and summarization of the discovered rules," in *Proceedings of the 6th ACM International Conference on Knowledge Discovery and Data Mining (KDD '01)*, pp. 91–100, Boston, Mass, USA, August 2000.

[10] S. Yin, S. X. Ding, A. H. A. Sari, and H. Hao, "Data-driven monitoring for stochastic systems and its application on batch process," *International Journal of Systems Science*, vol. 44, no. 7, pp. 1366–1376, 2013.

[11] S. Yin, H. Luo, and S. X. Ding, "Real-time implementation of fault-tolerant control systems with performance optimization," *IEEE Transactions on Industrial Electronics*, vol. 64, no. 5, pp. 2402–2411, 2014.

[12] S. Yin, S. Ding, A. Haghani, H. Hao, and P. Zhang, "A comparison study of basic datadriven fault diagnosis and process monitoring methods on the benchmark Tennessee Eastman process," *Journal of Process Control*, vol. 22, no. 9, pp. 1567–1581, 2012.

[13] P. Hajela, J. Yoo, and J. Lee, "GA based simulation of immune networks applications in structural optimization," *Engineering Optimization*, vol. 29, no. 1–4, pp. 131–149, 1997.

[14] L. Wang, J. Pan, and L. Jiao, "Immune algorithm," *Acta Electronica Sinica*, vol. 28, no. 7, pp. 74–78, 2000.

[15] J. Timmis, M. Neal, and J. Hunt, "An artificial immune system for data analysis," *BioSystems*, vol. 55, no. 1–3, pp. 143–150, 2000.

[16] S. M. Garrett, "How do we evaluate artificial immune systems?" *Evolutionary Computation*, vol. 13, no. 2, pp. 145–178, 2005.

[17] V. Cutello, G. Nicosia, M. Pavone, and J. Timmis, "An immune algorithm for protein structure prediction on lattice models," *IEEE Transactions on Evolutionary Computation*, vol. 11, no. 1, pp. 101–117, 2007.

[18] L. Jiao and L. Wang, "A novel genetic algorithm based on immunity," *IEEE Transactions on Systems, Man, and Cybernetics A*, vol. 30, no. 5, pp. 552–561, 2000.