IEEE Access
Multidisciplinary ⁝ Rapid Review ⁝ Open Access Journal

Received June 18, 2019, accepted July 1, 2019, date of publication July 4, 2019, date of current version July 24, 2019.

Digital Object Identifier 10.1109/ACCESS.2019.2926799

An Improved Fish Swarm Algorithm for Neighborhood Rough Set Reduction and Its Application

LI ZOU [1,2,3], HONGXIN LI [1], WEI JIANG [1], AND XINHUA YANG [2]
[1] Software Technology Institute, Dalian Jiaotong University, Dalian 116028, China
[2] Dalian Key Laboratory of Welded Structures and Its Intelligent Manufacturing Technology (IMT) of Rail Transportation Equipment, Dalian Jiaotong University, Dalian 116028, China
[3] Sichuan Provincial Key Laboratory of Process Equipment and Control, Zigong 643000, China

Corresponding author: Xinhua Yang (yangxhdl@foxmail.com)

This work was supported in part by the National Science Foundation of China under Grant 51875072, in part by the Liaoning Provincial Education Department Project under Grant JDL2017025, and in part by the Open Project Program of the Sichuan Provincial Key Laboratory of Process Equipment and Control under Grant GK201815.

ABSTRACT In this paper, an improved fish swarm algorithm for neighborhood rough set reduction (IFSANRSR) is proposed. In IFSANRSR, by introducing an adaptive function to control the *visual* and *step* size of artificial fish, the problem of inconsistent convergence speed existed in a traditional artificial fish swarm algorithm (FSA) is avoided. The movement of artificial fish in the swarming and following behavior is improved to shorten the running time of the algorithm. The searching behavior is improved to enhance the local search ability without changing the global searching ability of the algorithm. By introducing the mechanism of extinction and rebirth, the worst solution is eliminated and rebirth takes place after each iteration, which ensures a high level of the overall fitness. The experimental results on three datasets from the University of California at Irvine (UCI) show that the attributes reduction by using the IFSANRSR has higher reduction rate and classification accuracy in most cases. It could better deal with real-valued data attributes and ensure the optimal attribute reduction set to be found. The experimental result on the decision system of aluminum alloy welded joints show that by using the IFSANRSR, the key influencing factors of fatigue life of aluminum alloy welded joints could be obtained, and the weight of each influencing factor could be calculated quantitatively.

INDEX TERMS Fish swarm algorithm, neighborhood rough set, welded joints, fatigue life, influencing factor.

I. INTRODUCTION

Rough set theory (RST) is one of the most important artificial intelligence technologies. No prior knowledge is needed when using RST. It can start from the data itself and realize the tasks of data classification and decision rules acquisition. As an efficient tool to feature reduction, RST can reduce the abundant information while preserving the meaning of the features. By now, it has been successfully applied to machine learning, data mining, decision analysis and many other fields [1]–[3]. Traditional RST is based on the equivalent relation and is suitable for category data. When dealing with continuous data, discretization must be done beforehand, which inevitably leads to information lost [4]. To overcome this disadvantage, many extension of RST such as fuzzy rough sets [5], tolerance approximation model [6], covering approximation mode [7] and neighborhood rough set model [8], [9] have been proposed. Among which, the neighborhood rough set model can process both numerical and categorical dataset via the δ-neighborhood set. Attribute reduction is one of the most important content in neighborhood rough set theory, and it aims to obtain a minimal attribute subset from a decision system while maintaining the same classification accuracy by deleting the noisy, irrelevant or misleading features. Generally speaking, attribute reduction algorithm involves heuristic and random

The associate editor coordinating the review of this manuscript and approving it for publication was Zhipeng Cai.

search strategies. In the two strategies, heuristic search often leads to local solution of attribute reduction. To avoid this problem, many researchers have turn to random search strategies such as swarm intelligence algorithm.

The basic idea of swarm intelligence algorithm is to construct stochastic optimization algorithm by imitating the population behavior of organisms in nature. The algorithm mainly simulates the optimization and search process as the searching behavior or evolutionary process of individuals in the population. It uses the points in the searching space to imitate the individuals of organisms in nature. It transforms the objective function of the problem into the adaptability of individuals to the environment in the population. It makes an analogy between the search iteration process of looking for the better feasible solution and the process of retaining the fittest or searching behavior of individuals in a population. Swarm intelligence algorithm includes genetic algorithm [10], ant colony algorithm [11], [12], particle swarm optimization [13], [14], artificial fish swarm algorithm [15], artificial bee colony algorithm [16], glowworm swarm optimization algorithm [17] and bat algorithm [18]. As a new kind of evolutionary algorithm, swarm intelligence algorithm has been successfully applied in the field of traffic flow model validation [19], distributed efficient positioning [20], fault diagnosis [21]–[23] and control of vehicle roll behavior performance [24] for its advantages of distribution, self-organization and strong robustness.

The fish swarm algorithm (FSA) is a new swarm intelligent technique proposed by Li *et al.* [25] that was inspired by the natural schooling behavior of fish. FSA presents a strong ability to avoid local minimums in order to achieve global optimization [26], [27]. It has been applied in function optimization [28] and least squares support vector machine problems [29].

In this work, based on the traditional swarm intelligence algorithm-FSA, a novel IFSANRSR is proposed to deal with continuous numerical data. In order to improve the convergence rate and precision of traditional FSA in IFSANRSR, adaptive function is introduced to control the visual and step size of artificial fish, three behaviors of the fish are improved and mechanism of extinction and rebirth are suggested. Using the strong ability of the global optimization of improved FSA, a neighborhood rough set model is constructed for feature reduction. UCI experimental results show that the proposed techniques have great power in feature reduction. Experimental result on the decision system of aluminum alloy welded joints show that by using IFSANRSR, the key influencing factors of fatigue life of aluminum alloy welded joints could be obtained, and the weight of each influencing factor could be calculated quantitatively. The main contributions of our work are two aspects: One aspect, improved FSA and neighborhood rough set are combined, a novel IFSANRSR is proposed for searching an effective feature subset. The other aspect, the IFSANRSR is used for analysis of the fatigue life influencing factors of the aluminum alloy welded joints. The key fatigue life influencing factors and the quantitatively computation of the weights of each factor are obtained.

This paper is organized as follows. Section II. discusses the improved FSA. Section III. proposes the IFSANRSR and discusses the detail steps and processes of IFSANRSR, some experiments on the standard UCI datasets are carried out and the results of the experiments are analyzed. Section IV. constructs the evaluation model for fatigue life influencing factors of aluminum alloy welded joints by using IFSANRSR. Section V. summarizes the whole paper and presents future work in this area.

II. IMPROVED FSA

FSA is a specific application of swarm intelligent technique to optimize the search domain by simulating the behavior of fish including searching, swarming, and following. It has the abilities of self-organized and converging rapidly on a solution. In a search space, the FSA initially produces a swarm of artificial fish, and each artificial fish performs three behaviors to construct its own local solution then exchanges information in its self-organized system and finally achieves the global solution.

A. BASIC FSA

In basic FSA, the parameters are randomly initialized with a swarm of artificial fish. Suppose $X = (X_1, X_2, \ldots, Xn)$ represents individual artificial fish, $F(X)$ is a fitness function which indicate the food concentration of a fish position X_i, $D_{\mathrm{ij}} = |X_i\text{-}X_j|$ expresses the distance between individual artificial fish, *visual* represents a perceived distance of an artificial fish, artificial fish can only search within the *visual*, *step* is a moving step of an artificial fish, δ is a crowding factor. All the fish try to find and locate the place that can satisfy their food needs by three distinct behaviors.

1) SEARCHING BEHAVIOR

Searching is one of the most basic behaviors of organisms. When an artificial fish searches for food, it always walks a step toward a high concentration of its own visible neighborhood range. Artificial fish can perceive the concentration of food in the environment and tend to swim to the position where the concentration of food is high within the scope of visual. Searching is a random search strategy, with a tendency toward the food rich place. It is expressed as follows:

$$X_{next} = X_i + Rand() \cdot step \cdot \frac{X_j - X_i}{\|X_j - X_i\|}, F(Xj) > F(Xi) \tag{1}$$

$$X_{next} = X_i + Rand() \cdot step, F(Xj) \leq F(Xi) \tag{2}$$

where *Xi* is the current position of an artificial fish, $F(Xi)$ is its food concentration, and *Rand()* is a random value between 0 and 1.A new position *Xj* is selected randomly in the fish's *visual* range. If the formula $F(Xj) > F(Xi)$ is satisfied, the fish moves a step toward the direction of *Xj* to the next position *X next*. If the formula $F(Xj) \leq F(Xi)$ is satisfied, a random step within the *visual* range will be employed.

2) SWARMING BEHAVIOR

Swarming is a common activity of fish. A large or small number of fish gather in a place rich in food concentration during the evolution of fishes in nature. For each artificial fish, it is necessary to be as close as possible to the center of the fish swarm and avoid overcrowding around it. The swarming behavior is expressed as follows:

$$X_{next} = X_i + Rand() \cdot step \cdot \frac{X_C - X_i}{\|X_C - X_i\|}, \quad F(Xc) > F(Xi) \text{ and } \frac{ns}{n} < \delta \tag{3}$$

where X_C is the center position of a fish swarm, $F(X_C)$ is it's food concentration and n_s is the number of individuals within the X_C*'s visual*. The fish moves to the center position X_C if $F(X_C) > F(X_i)$ and $ns/n < \delta$ because the center is not overly crowded and has a greater concentration of food than the current position X_i. Otherwise, a searching behavior is employed to find a next position for the fish.

3) FOLLOWING BEHAVIOR

The following behavior is a directional behavior. When a fish is in a high food concentration position, its neighboring partners will follow it quickly. For each artificial fish, it is necessary to move to a place with higher food concentration while avoiding overcrowding.

$$X_{next} = X_i + Rand() \cdot step \cdot \frac{Xmax - X_i}{\|Xmax - X_i\|}, F(X\max) > F(Xi) and \frac{ns}{n} < \delta \tag{4}$$

where X_{max} is the best position of a neighboring fish swarm, $F(X_{max})$ is it's food concentration. $F(X\max) > F(Xi)$ and $ns/n < \delta$ represents X_{max} is in a good condition with higher food concentration and lower crowding, thus the fish will move one step toward X_{max}. Otherwise, searching behavior commences to determine a fish's next position.

Besides, FSA should provide a bulletin that records the optimal place and fitness of artificial fishes during each iteration. After making movements, each fish updates and compares its own state with the bulletin. The value on the bulletin will be replaced if its current state of fish is better. At the end of the algorithm, the best state and the fitness on the bulletin should be output as the final solution of the optimal problem task. The process of the basic FSA is shown as Fig. 1.

Generally speaking, FSA has fast convergence speed, strong robustness, small dependence on initial parameters and fast search of feasible solution range. It is suitable for solving optimization problems with high accuracy requirements. Meanwhile, in the later period of FSA, it may fall into a slow convergence rate and be caught in a limited precision. An improved FSA algorithm is proposed here to solve these problems in this work.

B. IMPROVED FSA

Aiming to improve the convergence rate and precision of traditional FSA, an improved FSA is proposed here. The core idea of improved FSA is: Firstly, in order to give consideration to both convergence speed and accuracy, adaptive function is introduced to control the visual field and step size of the artificial fish. It makes the size of moving step and the vision of artificial fish decrease with the increase of iteration times. Secondly, to increase the convergence speed, the moving strategy of artificial fish in swarming and following behavior is improved. The crowding factor is ignored, and the default searching behavior is cancelled. Thirdly, in order to improve the efficiency of the algorithm, the searching behavior is improved. If no better solution is found, the artificial fish tryes again with the new vision and step size. Lastly, the mechanism of extinction and regeneration is proposed. After each iteration, the worst artificial fish will extinct and a more adaptable artificial fish would be regenerated to ensure the overall fitness.

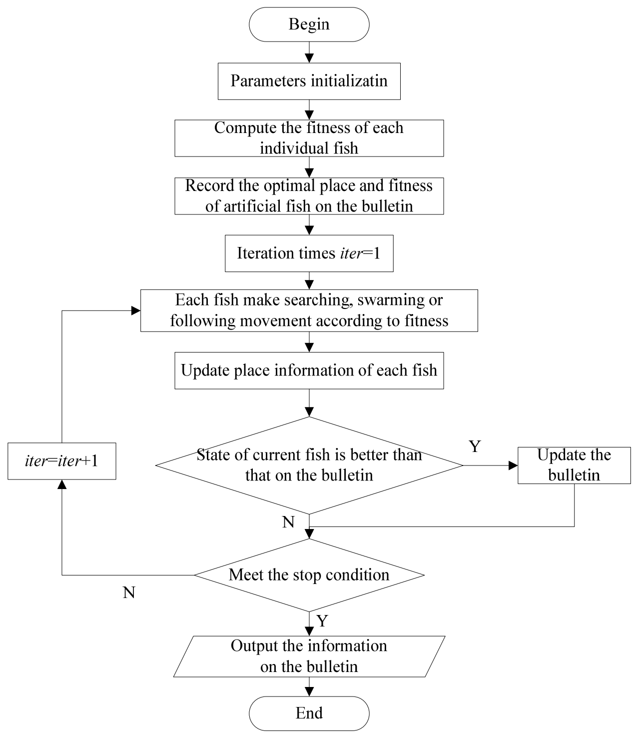


FIGURE 1. Process of basic FSA. Two common stop conditions for FSA are firstly, when the iteration times reach the maximum number of iterations the algorithm stops. Secondly, when the record of bulletin board does not change in several successive iterations, the algorithm stops.

1) ADAPTIVE STEP AND VISUAL

From the analysis of the basic FSA, it could be seen that *visual* and the maximum *step* size of artificial fish have significant influence on the convergence of the algorithm. When *visual* and *step* are set to a fixed constant, it may be too big or too small. In the early period of FSA, bigger *visual* and *step* can accelerate the fast convergence of swarming and following behavior, so that artificial fish can converge to the local and global optimal position quickly after several iterations. In the later period, smaller *visual* and *step* size make searching behavior dominant. The artificial fish around the extreme value could find the location of the adjacent extreme value accurately, so as to find the optimal value.

To assure the convergence speed and the accuracy of the solution, *visual* and *step* size of artificial fish should be reduced gradually. We introduce an improved method of changing of artificial fish *visual* and *step* size based on piecewise adaptive function, so that the *step* size and *visual* could decrease with the number of iterations. The formulas of adaptive *step* size and *visual* are as follows.

$$visual = Max\_V \cdot iter^{\frac{\log(Min\_V/Max\_V)}{\log(Max\_gen)}} \tag{5}$$

$$step = Max\_S \cdot iter^{\frac{\log(Min\_S/Max\_S)}{\log(Max\_gen)}} \tag{6}$$

where Max_V and Min_V indicate the maximum and minimum value of *visual*, Max_S and Min_S are the maximum and minimum *step* size respectively. The adaptive function decays as a power function, and the parameters decay quickly at the beginning of the iteration, so the value of Max_V and Min_V could be increased appropriately to enhance the search ability of the artificial fish.

In the attribute reduction algorithm of neighborhood rough set based on fish swarm, artificial fish is a string of binary codes. The distance between artificial fish is a weighted average distance and the value of the distance is multiple of 1. Therefore, the *step* size and *visual* should be integer values in the application. In order to apply FSA to attribute reduction of neighborhood rough set, the minimum *step* size and *visual* are set to be 1. The formulas are as follows:

$$visual = int(Max\_V \cdot iter^{\frac{\log(1/Max\_V)}{\log(Max\_gen)}}) \tag{7}$$

$$step = int(Max\_S \cdot iter^{\frac{\log(1/Max\_S)}{\log(Max\_gen)}}) \tag{8}$$

2) IMPROVED SWARMING AND FOLLOWING BEHAVIOR

In artificial fish swarm algorithm, the convergence speed is determined by the behavior of swarming and following behavior. Artificial fishes dispersed in the search space perform swarming and following behavior. They move quickly to the center of the fish group and to the location of neighboring partners with better fitness values, which enables the artificial fish to near to the extreme position in a relatively short time and ensures the convergence speed of the artificial fish swarm algorithm.

If the artificial fish finds that the center of the fish group within *visual* field is better than the current position, and the center of the fish group is not too crowded, swarming behavior is performed and the artificial fish moves to that position.

When the distance $d_{i,c} = |X_c\text{-}X_i|$ between the center position of all the partners(X_C) within *visual* and the current position of the artificial fish(X_i) is between *visual* and *step* ($step < d_{i,c} < visual$), the artificial fish need to move many times to reach the center position. This may make the artificial fish fall into the local extreme position in movements, and the increase of the number of moves also reduces the convergence speed of the algorithm. Therefore, the moving *step* size is improved in the swarming behavior as in formula(9).

$$step = Rand() \cdot |XC - Xi| \tag{9}$$

where $Rand()$ is added to prevent the swarming behavior from convergence too fast and falling into the local extreme position. By changing the moving *step* of the swarming behavior, the convergence speed of the algorithm is accelerated.

On the other hand, through parameter analysis of artificial fish swarm algorithm, we find that it is difficult to ensure convergence speed and accuracy at the same time by using crowding factor [25]. In attribute reduction of neighborhood rough set, only the best reduction subset (i.e. the optimal location of artificial fish swarm) is required. Therefore, in attribute reduction of neighborhood rough set based on improved fish swarm, the crowding factor is ignored. In swarming behavior, only the fitness of the partner's central position in the *visual* field is considered, if the fitness of the partner's central position is higher than the current position ($F(X_C) > F(X_i)$), the artificial fish swims one step towards the partner's central position.

The default behavior of swarming is searching behavior. When the center position of fish group searched is inferior to the current position of the artificial fish, the searching behavior is performed. While in FSA, the performance of searching, swarming and following has been compared. Searching behavior may be performed in the swarming behavior once more, which greatly increases the execution time of the algorithm. Therefore, an improvement is made here to cancel searching behavior in the swarming behavior and the artificial fish should remain in its original position.

The improved swarming behavior is shown as formula (10) and (11).

$$X_{next} = X_i + step \cdot \frac{XC - Xi}{\|XC - Xi\|}, F(XC) > F(Xi) \tag{10}$$

$$X_{next} = X_i, F(Xc) \le F(Xi) \tag{11}$$

The principle of following behavior is similar to clustering behavior, the improvement of following behavior is the same as clustering behavior, and it's not described here anymore.

3) IMPROVED SEARCHING BEHAVIOR

In searching behavior, the artificial fish searches the feasible domain space within the number of attempts, and then moves to the next position in the field of visual. Through the analysis of parameters of artificial fish swarm algorithm, we find that the number of attempts has a great influence on searching behavior thus affect the results of searching behavior, which will easily lead to premature and ineffective searching. Therefore, the number of attempts is improved in the algorithm.

If the artificial fish hasn't found the next better position after *try_number* times searching attempts, the artificial fish will try again, but at this time the artificial fish has a higher grasp of the environment in the field of *visual*, and it is easy to result in invalid searching while weakening the execution time of the algorithm. Therefore, we will expand the field of *visual* of the artificial fish, the new field of *visual* is $visual_{\text{new}} = visual + step$. If in this new field of visual, a better position with high fitness is found, the fish would move one step to the better position, the maximum step size is $step_{\text{new}} = 2 \times step$. If the better position is not found after *try_number* times searching attempts in the new field of *visual*, the artificial fish will move randomly in the new field of *visual*.

4) MECHANISM OF EXTINCTION AND REBIRTH

In the biological groups, individuals with higher adaptability can adapt quickly to the changes of the environment, so they have greater survival chances. While individuals with lower adaptability are unable to adapt quickly to the new environment, so they are gradually eliminated in the course of evolution. Inspired by this idea, the artificial fish with high adaptability is defined as elite fish, and the artificial fish with low adaptability is called inferior fish. When FSA is used to solve optimization problems, elite fish can quickly adapt to environmental changes because of their high adaptability. It could find a better location within several iterations. While inferior fish often need more iteration to find a better location to adapt to the environment, which increases the execution time of the algorithm.

In order to shorten the execution time of the algorithm, the extinction mechanism is introduced. In each iteration process of FSA, the artificial fish is sorted according to the fitness level after its position is updated. The artificial fish with the lowest fitness is regarded as inferior fish thus should be extinct. By introducing extinction, the remained artificial fishes have higher adaptability, which improves the overall adaptability level of the fish swarm. However, in the extinction mechanism, the size of the fish swarm shrinks. As the number of iterations increases, the fish swarm becomes smaller and smaller, and the randomness of the algorithm becomes smaller and smaller. It is difficult to ensure that the optimal solution location could be found finally. Therefore, a regeneration mechanism is introduced to regenerate the same number of artificial fish with high adaptability after eliminating inferior fish in each iteration process, so as to ensure that the size of the fish does not change under the premise of improving the overall adaptability level of the fish.

In the FSA based attribute reduction of neighborhood rough set, the fitness values of each artificial fish location are sorted in ascending order after each iteration process. When the fish is reborn, determine whether the fitness of each artificial fish location is equal. If the fitness is equal, any artificial fish will be extinct and reborn to be the artificial fish recorded on the bulletin board. Otherwise an artificial fish with the lowest degree of fitness will be extinct and reborn to artificial fish with greater fitness. By introducing extinction and regeneration, a high overall level of fitness could be assured and the time of each iteration is reduced.

III. IMPROVED FSA FOR NEIGHBORHOOD FEATURE REDUCTION

A. NEIGHBORHOOD FEATURE REDUCTION

According to [30], [31], a neighborhood decision system can be denoted as $ND = (U, C \cup D, \delta)$, where $U = \{x_1, x_2, \ldots, x_n\}$ is a nonempty finite set of objects called the universe, $C = \{a_1, a_2, \ldots, a_m\}$ is the set of condition features, D is the set of decision features, and δ is the neighborhood parameter$(0 \leq \delta \leq 1)$.The neighborhood dependency of D on B is defined as:

$$\gamma_B(D)_\delta = {}^{|POS_B(D)_\delta|}/_{|U|} \tag{12}$$

If the neighborhood dependency $\gamma_{B-\{a\}}(D)_\delta < \gamma_B(D)_\delta$ then the attribute a is indispensably to the set B, otherwise a is redundant to B, that it can be removed from B.

Traditional neighborhood feature reduction algorithm such as forward greedy algorithm is easy to fall into local optimal solution. So improved FSA is introduced to neighborhood feature reduction to solve this problem in this work. Taking advantages of improved FSA's suitable for distributed processing of optimization problem, the search efficiency of attribute reduction and the possibility of finding the best reduction are greatly improved. When we introduce the improved FSA into neighborhood rough set feature reduction, the urgent problem needs to solve is how to measure the distance of discrete feature subsets. In this work, the feature subsets are transformed into binary numbers and the Hamming metric is introduced to measure the distance between two binary numbers as described in reference [32].

B. NEIGHBORHOOD FEATURE REDUCTION PREMISE OF IMPROVED FSA FOR NEIGHBORHOOD FEATURE REDUCTION

1) CODING OF ARTIFICIAL FISH LOCATION

A decision-making system with n attributes has 2^n kinds of combinations of attribute subsets, and the final solution of attribute reduction is expressed as a set of conditional attributes. Therefore, if each attribute is represented by a binary number, the position of each artificial fish is a series of n-bits of binary coding. If the ith feature of the information system is selected as a key feature, then the ith bit binary number is '1' and '0' otherwise. For example, suppose there are six conditional attributes$\{a_1, a_2, \ldots, a_6\}$ in the decision system, if the reduction result is $\{a_2, a_3, a_6\}$ then the corresponding position of artificial fish is "011001" and vice versa, each binary string corresponds to a reduction of the attributes.

2) HAMMING DISTANCE BETWEEN FISH

In this work, the Hamming distance is used to measure the distance between two fishes. Let X_i, X_j be two binary numbers representing positions of two artificial fish. The Hamming distance between X_i and X_j is defined by

$$h(X_i, X_j) = \sum_{i,j=1}^{n} x_i \oplus x_j \tag{13}$$

where $\oplus$ is a *XOR* operation, x_i, $x_j \in \{0, 1\}$. The variable x_i represents a binary bit in X_i.

3) CENTER OF FISH

In the improved FSA, it is necessary to get the center position of the fish swarm within *visual* so that the artificial fish can move. The determination of the center position of the fish swarm is very important for the swarming behavior. Let X_1, $X_2, \ldots X_n$ be some binary numbers that represent positions of n fishes, the center position of n fishes is defined as

$$X_C = \{c_i | if \frac{1}{2}\sum_{i=1}^{n} x_i^j > 0.5, then\ ci = 1, else\ ci = 0\} \quad (14)$$

where X_C is the center position of n fishes, x_i^j represents the *ith* bit of fish position X_i.

4) FITNESS FUNCTION AND UPDATING CONDITIONS OF BULLETIN BOARD

In the improved FSA, the design of fitness function is the core step of the whole algorithm, it determines the convergence direction of the algorithm to a great extent. The attribute reduction of neighborhood rough sets is mainly to get the best reduction. Under the condition that the classification ability of decision system is unchanged, the redundant conditional attributes are deleted and the smaller conditional attributes set is obtained.

In attribute reduction of neighborhood rough sets, two main problems are considered: one is the new attribute set keeps the classification ability of the original attribute set or not, the other is whether the new attribute set contains redundant attributes. Therefore, the design of the improved FSA for neighborhood rough set should achieve two goals:

Firstly, the number of conditional attributes in the reduction result should be as small as possible.

Secondly, the classification ability of the reduction set is consistent with that of all conditional attribute sets.

In order to achieve the first goal, the fitness function is set according to formula (15)

$$Fitness = \gamma_B(D)\delta \quad (15)$$

where $\gamma_B(D)\delta$ is the neighborhood classification quality (or neighborhood dependency) of condition feature set B relative to decision D as defined in formula(12).

In order to achieve second goal, the update condition of the bulletin board is that on one hand the fitness of the artificial fish at current location is greater than that of bulletin board recorded location, on the other, the length of reduction set represented by the artificial fish at current location is less than that represented by the recorded position of bulletin board.

5) STOP CONDITION

Generally, there are two common stop conditions for FSA: firstly, when the iteration times reach the maximum number of iterations the algorithm stops. Secondly, when the record of bulletin board does not change in several successive iterations, the algorithm stops. While in attribute reduction, there is no fixed termination condition, the shortest reduction set which guarantees the consistency of classification ability with the original decision table is the final objective. In this work, we design the termination condition of the algorithm as follows:

Firstly, set a maximum number of iterations, when the number of iterations is greater than the maximum number of iterations, the algorithm terminates.

Secondly, the algorithm terminates when the same solution appears in three consecutive iterations.

This ensures that the final solution is the shortest and most suitable reduction set after many iterations, avoids unnecessary iterations and saves the running time of the algorithm.

C. PROCESS OF IFSANRSR

In the proposed IFSANRSR algorithm, improved swarm and following behavior, improved searching behavior and mechanism of extinction and rebirth are introduced to ensure the global stability and to increase the convergence speed of the traditional FSA algorithm. Four aspects of improvement ensure that the reduction results are stable during the simulation process which indicate the stability of the proposed method.

The process of IFSANRSR proposed in this work is shown as the flowchart in Fig. 2.

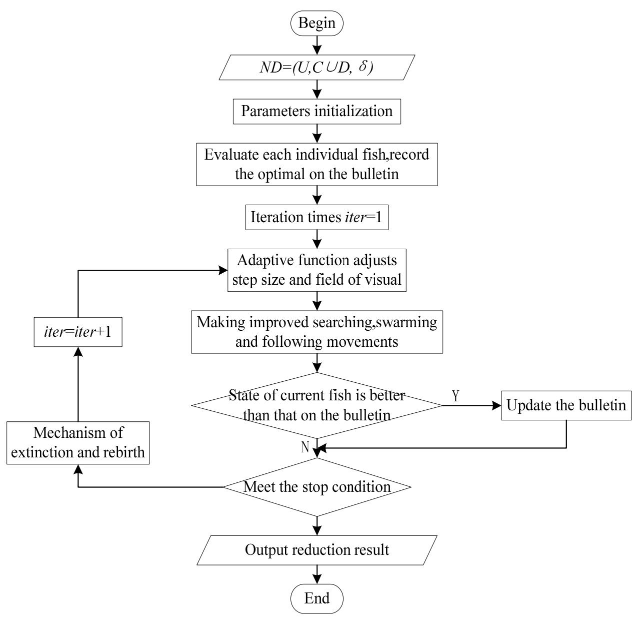


FIGURE 2. Process of basic IFSANRSR. The stop conditions of IFSANRSR is the same as that of FSA.

D. EXPERIMENT AND DISCUSSION

In order to verify the effectiveness of IFSANRSR for neighborhood attribute reduction, experiments are carried out on three data sets from UCI, including Wine, Ionosphere and Lymphography. The name of the dataset, the number of samples, features (conditional attributes), and classes

TABLE 1. Datasets used in experiment.

U	Datasets	Samples	Features	Classes
1	Lymphography	114	18	4
2	Wine	178	13	3
3	Ionosphere	351	34	2

(the category of decision attributes) for each dataset is described in TABLE 1.

Attribute reduction is to find the minimum reduction on condition that the overall classification ability of decision system remains unchanged; the concept of *Rate* is defined to express the efficiency of reduction. The formula of reduction rate is as formula (16).

$$Rate = \frac{|C| - |R|}{|C|} \tag{16}$$

where $|C|$ represents the number of all features in the dataset and $|R|$ is the number of selected features. Reduction rate represents the redundancy of conditional attributes in decision-making system. The higher the reduction rate is, the lower the redundancy of conditional attributes is. In this work, the datasets are normalized before the experiment. The formula for normalization is shown as formula (17).

$$f(x_i) = \frac{x_i - x_{\min}}{x_{\max} - x_{\min}} \tag{17}$$

where x_{max} and x_{min} is the maximum and minimum value of the samples group respectively. The value of the samples all fall in the interval [0,1] after normalization.

1) COMPARISON AND ANALYSIS OF REDUCTION RESULTS

In the reduction experiments, we will use comparable experiments on those three UCI datasets to evaluate the reduction performance of the proposed IFSANRSR with FARNeMF [33], FSANRSR and ACO [34]. Main steps of Hu's forward greedy algorithm is shown as Fig. 3, where ε is the threshold of the significance of attribute and its value is close to 0, further details about FARNeMF could be found in reference [33]:

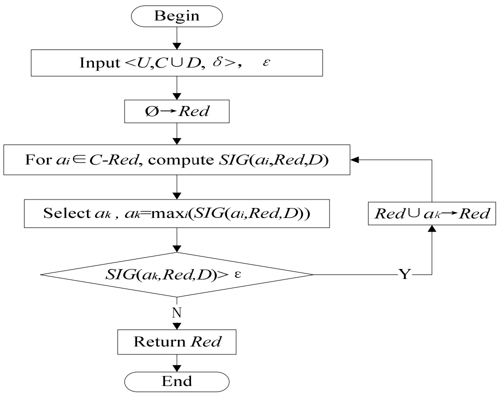


FIGURE 3. Process of FARNeMF where ε is a positive number close to 0 indicate the threshold of the significance of attribute. The stop condition of the algorithm is when no attribute's significance is greater than ε.

These experiments are designed by MATLAB 2014 and tested on a professional workstation running windows 10 with Intel(R) Xeon(R) Bronze 3104 CPU @ 1.70GHz processor and 16GB memory. And those comparable experiments are test on different datasets with the same neighborhood parameter δ, which equals to 0.1. In FSANRSR, the iteration times is set to be 100, the number of artificial fish is set to be half of the number of conditional attributes in the data set, *visual* equals half of the number of the artificial fish, and *step* equals *visual*-1. In IFSANRSR, the iteration times and the number of artificial fish is set to be the same as that in the FSA, the *Max_V* is half of the number of artificial fish, and the *Max_S* is *Max_V* - 1.

The experimental comparison results are summarized in Table 2. The leftmost column indicates the dataset names. The second and third columns are reduction set and reduction rate of corresponding dataset with Hu's FARNeMF algorithm. The fourth column is the running time of algorithm FARNeMF. The following three columns are reduction set, reduction rate and running time of ACO. The next three columns are reduction set, reduction rate and running time of FSANRSR. The eleventh and twelfth columns are reduction set and reduction rate of the proposed algorithm IFSANRST. The last column is the running time of algorithm IFSANRSR.

As could be seen from TABLE 2, for Lymphography data set, the minimal reduction results obtained by the four algorithms are the same as {2,13,14,15,16,18}, and the FARNeMF algorithm has shorter running time. For Wine dataset, the reduction result obtained by FARNeMF algorithm is {1,6,7,10,13}, the reduction rate is 0.62 and the running time is 0.2 seconds; the reduction result obtained by ACO is {1,5,7,11}, the reduction rate is 0.69 and the running time is 34 seconds; the minimum reduction set obtained by FSANRSR and IFSANRSR algorithm is the same as {1,7,11,13}, the reduction rate is the same as 0.69 and the running time is 68 and 22 seconds respectively. For Ionosphere datasets, both FSANRSR and IFSANRSR algorithm have a shorter reduction set, and a higher reduction rate.

As is shown in TABLE 2, the FARNeMF algorithm obtains good time performance for its quickly heuristic search, but it is easily trapped in a local solution. The three algorithms of

TABLE 2. Comparison of reduction results.

Datasets	FARNeMF			ACO			FSANRSR			IFSANRSR		
	Reduction	Rate	Time(s)	Reduction	Rate	Time(s)	Reduction	Rate	Time(s)	Reduction	Rate	Time(s)
Lymphography	2,13,14,15,16,18	0.67	0.1	2,13,14,15,16,18	0.67	57	2,13,14,15,16,18	0.67	118	2,13,14,15,16,18	0.67	49
Wine	1,6,7,10,13	0.62	0.2	1,5,7,11	0.69	34	1,7,11,13	0.69	68	1,7,11,13	0.69	22
Ionosphere	1,2,4,8,12,23,33	0.79	0.8	1,2,4,8,12,23,33	0.79	301	2,3,9,23,30,33	0.82	828	2,3,9,23,30,33	0.82	283

ACO, FSANRSR and IFSANRSR could overcome the weakness of the FARNeMF, both of FSANRSR and IFSANRSR have the same reduction result and reduction rate on the three datasets of Lymphography, Wine and Ionosphere. While compared with FSANRSR and ACO, the IFSANRSR has shorter running times. To sum up, the three algorithm of ACO, FSANRSR and IFSANRSR outperform that of the FARNeMF, IFSANRSR has better running time performance than FSANRSR and ACO algorithm.

2) COMPARISON AND ANALYSIS OF REDUCTION RATE

Neighborhood parameter δ could be considered as a tool to control the granularity of data analysis. It plays an important role in neighborhood rough sets. We here change the value of neighborhood parameter in the neighborhood rough set model. The neighborhood parameter values of δ vary from 0.05 to 1 with step 0.05 to get different feature subsets and then evaluate the reduction performance by calculating relevant reduction rate. To graphically illustrate the reduction performance of FARNeMF and IFSANRSR with different neighborhood parameters, take neighborhood parameter δ as the horizontal coordinate and the reduction rate as the vertical coordinate. The experimental results of the three datasets are in Fig. 4.

As could be seen from Fig. 4, in the three datasets, the granule is rougher as the neighborhood parameter δ is bigger. The reduction rates of both IFSANRSR and FARNeMF are decreasing with the increase of neighborhood parameter δ in most situations. For Lymphography and Wine datasets, the reduction rate of IFSANRSR algorithm is higher than that of FARNeMF algorithm. For Ionosphere datasets, only when the parameter of neighborhood is between 0.26 and 0.55, the reduction rate of FARNeMF algorithm is higher than that of IFSANRSR algorithm, and in other cases the reduction rate of FARNeMF is lower than that of IFSANRSR algorithm. Therefore, through the comparative analysis of the three datasets, we could find that the reduction rate of IFSANRSR algorithm is higher than that of FARNeMF algorithm in most cases, which indicates that IFSANRSR algorithm can get shorter reduction result than FARNeMF algorithm.

3) COMPARISON AND ANALYSIS OF CLASSIFICATION ACCURACY

Attribute reduction aims to delete unnecessary and redundant attributes by keeping the classification ability of the decision system unchanged and tries to avoid the influence of redundant attributes on rule extraction. In order to evaluate the classification ability of the reduction sets by using IFSANRSR and FARNeMF, support vector machine is used to extract classification rules from the reduction set. We apply tenfold cross-validation method to estimate the classification. Comparison of classification accuracy of the proposed algorithm IFSANRSR with that of Hu's neighborhood-based feature reduction algorithm FARNeMF is shown in Fig. 5. In the experiment, the value of the neighborhood parameter δ varies from 0.05 to 1 with step of 0.05 to get different feature reduction subsets and then the selected features are evaluated with SVM.

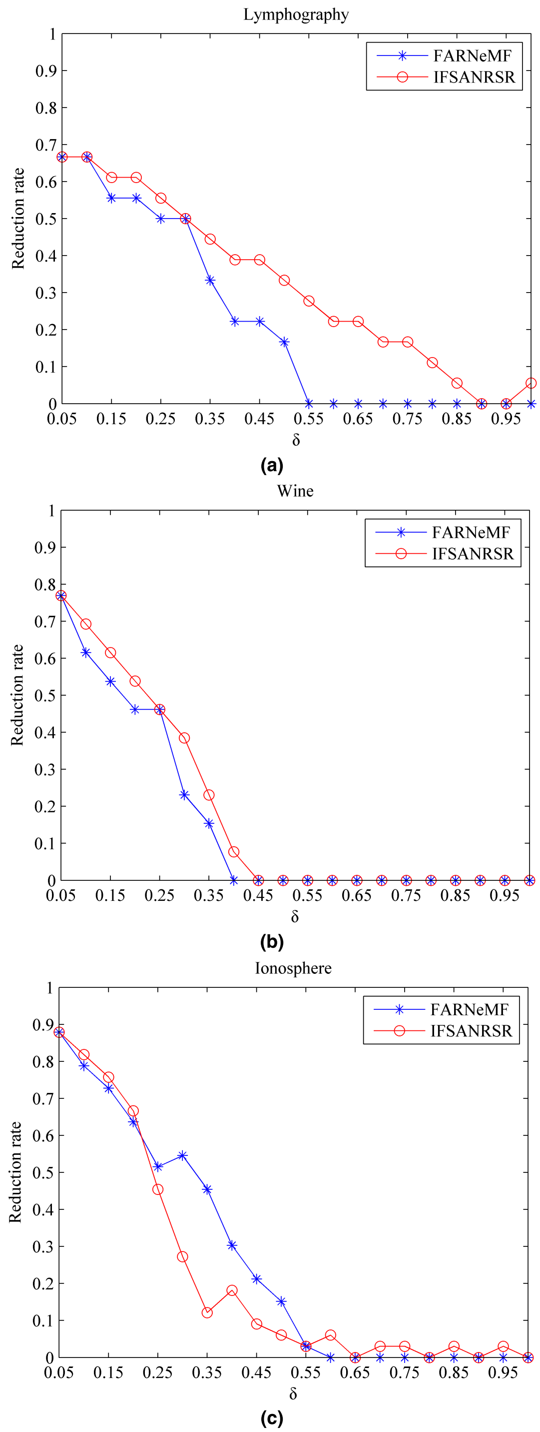


FIGURE 4. **Comparison of reduction rate with different δ. (a) Lymphography. (b) Wine. (c) Ionosphere.**

As could be seen from Fig. 5, for Lymphography dataset, the classification accuracy of the reduction sets obtained by using IFSANRSR algorithm is generally higher than that of FARNeMF algorithm. Only when δ is 0.35, the classification accuracy of IFSANRSR algorithm is slightly lower. We can find that the classification accuracy reaches the best

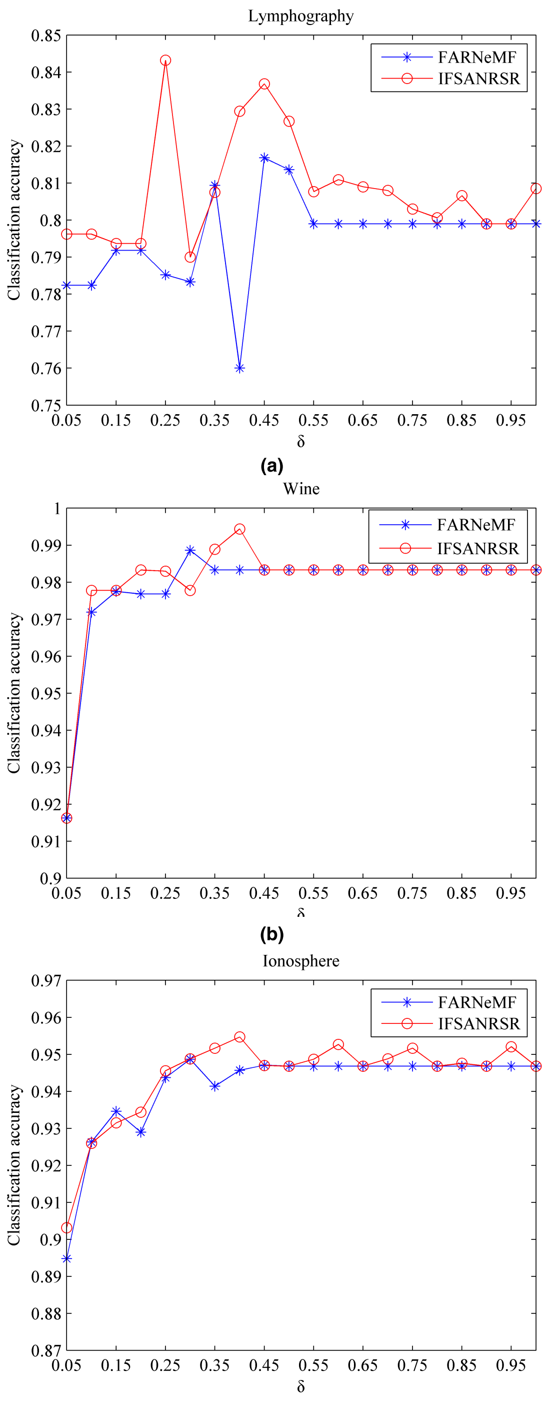


FIGURE 5. Comparison of classification accuracy with different δ. (a) Lymphography. (b) Wine. (c) Ionosphere.

performance when the δ equals 0.4 in the IFSANRSR, while it reaches the best performance with δ equaling 0.3 in the FARNeMF. For Wine dataset, the classification accuracy reaches the best performance when the δ equals 0.25 in the IFSANRSR, while it has the best performance with δ equaling 0.45 in the FARNeMF. The classification accuracy of the reduction sets obtained by using IFSANRSR algorithm is higher than that of FARNeMF algorithm in most cases. For Ionosphere dataset, with the changing of δ from 0.05 to 0.3, the classification accuracy of IFSANRSR and FARNeMF algorithm both increase rapidly. When δ is greater than 0.3, both IFSANRSR and FARNeMF algorithm have a relatively high smooth classification accuracy level. The classification accuracy reaches the best performance when δ equals 0.4 in the IFSANRSR algorithm while it reaches the best performance with δ equaling 0.3 in the FARNeMF algorithm. The classification accuracy of the reduction sets obtained by using IFSANRSR is higher than that of FARNeMF in most cases. It shows that the neighborhood-based feature selection can find some good features for classification and efficiently delete the redundant and irrelevant features from the original data.

IV. FATIGUE LIFE INFLUENCING FACTORS BASED ON IFSANRSR

By using IFSANRSR, evaluation model of fatigue life influencing factors of aluminum alloy welded joints is built in this work. First of all, fatigue data of aluminum alloy welded joints are collected by three-point bending fatigue test of 5083 and 5A06 aluminum alloy T-welded joints [35] and by literature review [36], [37]. Decision system for fatigue life influencing factors are established after preprocesses of the fatigue data. Then IFSANRSR is used for attributes reduction and weights computation of the fatigue life influencing factors of aluminum alloy welded joints. At last, evaluation and analysis of reduction result of influencing factors are obtained.

A. DECISION SYSTEM FOR FATIGUE LIFE INFLUENCING FACTORS

Decision system for fatigue life influencing factors is constructed by three-point bending fatigue test besides by literature review. Altogether, there are 75 sample data and 7 main factors affecting the fatigue life of aluminum alloy welded joints in the decision system, including material type, welding method, plate thickness, stress ratio, load type, joint type and equivalent structural stress range. Among which, there are four influencing factors including material type, welding method, load type and joint type have discrete value. Totally, there are 8 types of materials, 4 kinds of welding method, 2 kinds of load types and 4 kinds of joint types. Let C ={material type(C_1), welding method(C_2), plate thickness(C_3), stress ratio(C_4), load type(C_5), joint type(C_6), equivalent structural stress range(C_7)}, D ={Fatigue Life(number of stress cycles)}, the decision system is constructed and part of the fatigue data collected are shown as Table 3.

B. TYPES OF GRAPHICS ATTRIBUTES REDUCTION AND WEIGHTS COMPUTAION

In the decision system of fatigue life influencing factors of aluminum alloy welded joints, IFSANRSR algorithm is used to obtain the key influence factors which affect the fatigue life of the aluminum alloy welded joints. For comparison,

TABLE 3. Decision system of fatigue life influencing factors(part).

U	C_1	C_2	$C_3(mm)$	C_4	C_5	C_6	$C_7(MP)$	D
1	5083 H11	MIG	10	0.1	4B	TJ:P	187.7	29250
2	5083 H11	MIG	10	0.1	4B	TJ:P	134.07	161650
3	AlMg4MnCr	GMAW	2.5	0.1	T	LJ_SS:p	154.71	17160
4	AlMg4MnCr	GMAW	2.5	0.1	T	LJ_SS:p	108.3	142840
5	AlMgSi1 (6082)	TIG	3	0	T	LJ_DS:p	282.5	3560
6	AlMgSi1 (6082)	TIG	3	0	T	LJ_DS:p	147.98	198930
7	5A06+5083	MIG	10	0.1	4B	T	79.72	2080000
8	5A06	MIG	16	0.1	3B	T	37.26	6368500
9	NP5/6	Manual Arc	4.7625	0	T	SJ_DS:p	257.6	32000
10	HP30	Manual Arc	4.7625	0	T	SJ_DS:p	154.95	188000
				…				

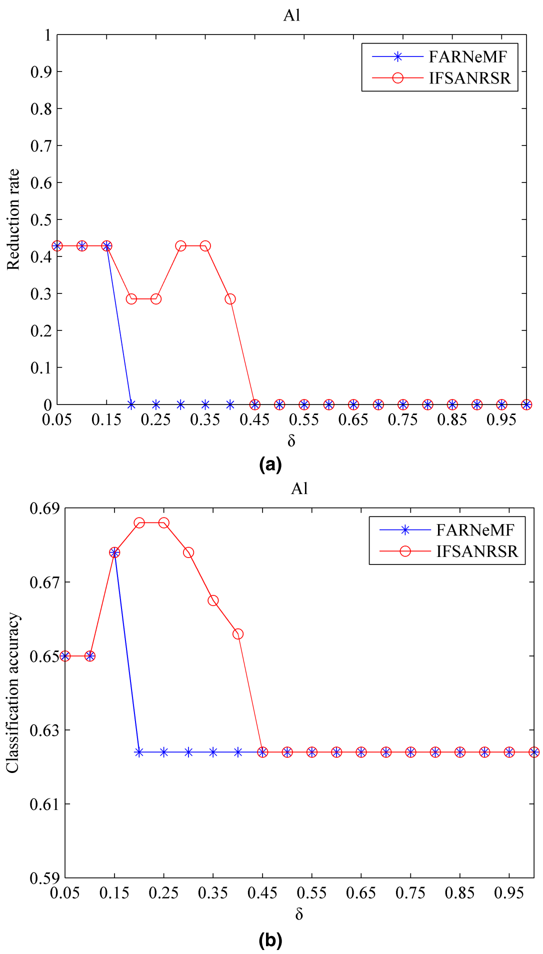


FIGURE 6. Comparison of reduction rate and classification accuracy with different δ. (a) Comparison of reduction rate. (b) Comparison of classification accuracy.

Hu's FARNeMF algorithm is also used for the attributes reduction of the decision system. The maximum iteration of IFSANRSR algorithm are set to be 100 times, the maximum step length is 2, the maximum field of *visual* is 3, and the number of fishes in the swarm is 4. The comparison of reduction rate and classification accuracy of the two algorithms with different neighborhood parameters is shown in Fig. 6.

As could be seen from Fig. 6 (a), when δ is within (0.05, 0.2), the decision system of fatigue life influencing

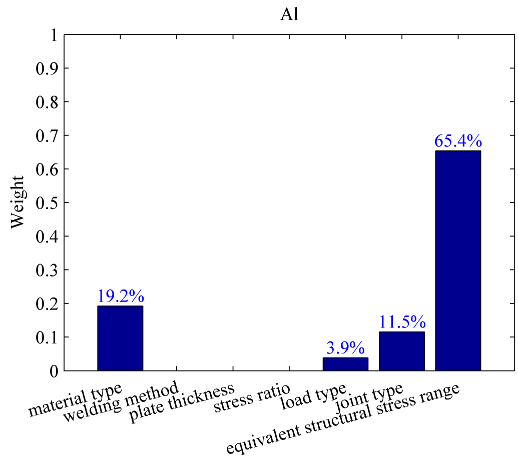


FIGURE 7. Weights of the fatigue life influencing factors of the aluminum alloy welded joints.

factors of aluminum alloy welded joints could be reduced by using FARNeMF algorithm, when δ is greater than 0.2, the decision system could not be reduced. Different from this, by using IFSARNSR algorithm, the decision system could be reduced when δ is between (0.05, 0.45). Both the FARNeMF algorithm and the IFSARNSR algorithm have the same best reduction rate for the decision system of fatigue life influencing factors of aluminum alloy welded joints. As could be seen from Fig. 5 (b), when δ is less than 0.15, both of the two algorithms have the same classification accuracy. When δ is between(0.25, 0.3), the IFSARNSR algorithm has higher reduction rate than the FARNeMF algorithm.

Accordingly, IFSANRSR algorithm is used to obtain the reduction result of the decision system of fatigue life influencing factors of aluminum alloy. To obtain a higher reduction rate, the neighborhood parameter δ is set to be 0.15. The obtained reduction result for the decision system of fatigue life influencing factors of aluminum alloy is $\{C_1$(material type), C_5(load type), C_6(joint type), C_7(equivalent structural stress range)$\}$. The weights of the 7 fatigue life influencing factors of aluminum alloy are $\{0.1923, 0, 0, 0, 0, 0, 0.0385, 0.1154, 0.6538\}$. Quantitative calculation result for weight of each influencing factor by using IFSANRSR algorithm is shown in Fig. 7.

C. DISCUSSION

As could be seen from Fig. 5, both Hu's FARNeMF algorithm and the IFSANRSR algorithm proposed in this work could be used for attributes reduction of the decision system of fatigue life influencing factors of aluminum alloy. By using IFSANRSR algorithm, the key influencing factors of fatigue life could be better obtained when the neighborhood parameter δ is between (0.15, 0.45).

As is shown in Fig. 7, key fatigue life influencing factors including equivalent structural stress range, material type, joint type and load type are the most important in the decision system of aluminum alloy welded joints. Quantitative calculation for weight of each influencing factor by using IFSANRSR algorithm is obtained at the same time. In descending order of attribute importance, they are equivalent structural stress range(0.6538), material type(0.1923), joint type(0.1154) and load type(0.0385). Reduction and quantitative calculation results indicate that besides equivalent structural stress range, other influencing factors such as material type, joint type and load type also have important influence on fatigue life prediction of aluminum alloy welded joints.

V. CONCLUSION

A novel IFSANRSR algorithm is proposed in this work. It could overcome the weakness of the conventional greedy searching approach (Hu's algorithm) to feature reduction. Experiments on three UCI test datasets are carried out. The results show that compared with the traditional FARNeMF algorithm, IFSANRSR algorithm has shorter reduction result and higher reduction rate, thus could better deal with continuous data attributes and assure the optimal attribute reduction set. IFSANRSR can quickly converge, it has a strong search capability in the problem space and could find the minimal reductions efficiently.

Experiments on the decision system for fatigue life influencing factors of aluminum alloy welded joints shows that by using IFSANRSR algorithm, key fatigue life influencing factors could better be found and the quantitative calculation for weight of each influencing factor could be obtained at the same time.

Future work would be concentrated on further validation of the IFSANRSR algorithm and its application for evaluation of fatigue life influencing factors when more fatigue test samples are available.

REFERENCES

[1] S. Chakhar, A. Ishizaka, A. Labib, and I. Saad, "Dominance-based rough set approach for group decisions," *Eur. J. Oper. Res.*, vol. 251, no. 1, pp. 206–224, 2016. doi: 10.1016/j.ejor.2015.10.060.

[2] M. Kadziński, R. Słowiński, and S. Greco, "Multiple criteria ranking and choice with all compatible minimal cover sets of decision rules," *Knowl.-Based Syst.*, vol. 89, no. 11, pp. 569–583, 2015. doi: 10.1016/j.knosys.2015.09.004.

[3] T. Entani and M. Inuiguchi, "Pairwise comparison based interval analysis for group decision aiding with multiple criteria," *Fuzzy Sets Syst.*, vol. 274, no. 9, pp. 79–96, 2015. doi: 10.1016/j.fss.2015.03.001.

[4] R. Jensen and Q. Shen, "Semantics-preserving dimensionality reduction: Rough and fuzzy-rough-based approaches," *IEEE Trans. Knowl. Data Eng.*, vol. 16, no. 12, pp. 1457–1471, Dec. 2004. doi: 10.1109/tkde.2004.96.

[5] Z. Zhang, "Attributes reduction based on intuitionistic fuzzy rough sets," *J. Intell. Fuzzy Syst.*, vol. 30, no. 2, pp. 1127–1137, 2016. doi: 10.3233/IFS-151835.

[6] L. Wang, Y. Jun, and T. Li, "Matrix-based approach for calculation of the approximations of a concept under tolerance probabilistic rough sets model," *Int. J. Advancements Comput. Technol.*, vol. 4, no. 11, pp. 368–377, 2012. doi: 10.4156/ijact.vol4.issue11.40.

[7] M. Restrepo, C. Cornelis, and J. Gómez, "Duality, conjugacy and adjointness of approximation operators in covering-based rough sets," *Int. J. Approx. Reasoning*, vol. 55, no. 1, pp. 469–485, 2014. doi: 10.1016/j.ijar.2013.08.002.

[8] K. Zeng, "Preference mining using neighborhood rough set model on two universes," *Comput. Intell. Neurosci.*, vol. 2016, pp. 1–13, Oct. 2016. doi: 10.1155/2016/6975458.

[9] Q. Hu, D. Yu, J. Liu, and C. Wu, "Neighborhood rough set based heterogeneous feature subset selection," *Inf. Sci.*, vol. 178, no. 18, pp. 3577–3594, 2008. doi: 10.1016/j.ins.2008.05.024.

[10] A. Diabat and R. Deskoores, "A hybrid genetic algrithm based heuristic for an integrated supply chain problem," *J. Manuf. Syst.*, vol. 38, pp. 172–180, Jan. 2016. doi: 10.1016/j.jmsy.2015.04.011.

[11] M. R. Jabbarpour, H. Zarrabi, J. J. Jung, and P. Kim, "A green ant-based method for path planning of unmanned ground vehicles," *IEEE Access*, vol. 5, pp. 1820–1832, 2017. doi: 10.1109/ACCESS.2017.2656999.

[12] W. Deng, J. Xu, and H. Zhao, "An improved ant colony optimization algorithm based on hybrid strategies for scheduling problem," *IEEE Access*, vol. 7, pp. 20281–20292, 2019. doi: 10.1109/ACCESS.2019.2897580.

[13] J. Kennedy and R. Eberhart, "Particle swarm optimization," in *Proc. IEEE Int. Conf. Neural Netw.*, Perth, WA, Australia, Nov./Dec. 1995, pp. 1942–1948.

[14] K. Kalita, U. Ragavendran, M. Ramachandran, and A. K. Bhoi, "Weighted sum multi-objective optimization of skew composite laminates," *Struct. Eng. Mech.*, vol. 69, no. 1, pp. 21–31, 2019. doi: 10.12989/sem.2019.69.1.021.

[15] H. Ge, L. Sun, X. Chen, and Y. Liang, "An Efficient artificial fish swarm model with estimation of distribution for flexible job shop scheduling," *Int. J. Comput. Intell. Syst.*, vol. 9, no. 5, pp. 917–931, 2016. doi: 10.1080/18756891.2016.1237190.

[16] S. Sundar, P. N. Suganthan, C. T. Jin, C. T. Xiang, and C. C. Soon, "A hybrid artificial bee colony algorithm for the job-shop scheduling problem with no-wait constraint," *Soft Comput.*, vol. 21, no. 5, pp. 1193–1202, 2016. doi: 10.1007/s00500-015-1852-9.

[17] A. Ray and D. De, "An energy efficient sensor movement approach using multi-parameter reverse glowworm swarm optimization algorithm in mobile wireless sensor network," *Simul. Model. Pract. Theory*, vol. 62, pp. 117–136, Mar. 2016. doi: 10.1016/j.simpat.2016.01.007.

[18] M. R. Sathya and M. M. T. Ansari, "Load frequency control using bat inspired algorithm based dual mode gain scheduling of PI controllers for interconnected power system," *Int. J. Elect. Power Energy Syst.*, vol. 64, pp. 365–374, Jan. 2015. doi: 10.1016/j.ijepes.2014.07.042.

[19] A. Poole and A. Kotsialos, "Swarm intelligence algorithms for macroscopic traffic flow model validation with automatic assignment of fundamental diagrams," *Appl. Soft Comput.*, vol. 38, pp. 134–150, Jan. 2016. doi: 10.1016/j.asoc.2015.09.011.

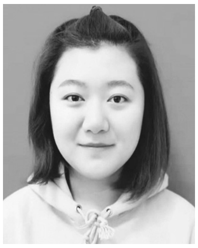
[20] A. O. de Sá, N. Nedjah, and L. de Macedo Mourelle, "Distributed efficient localization in swarm robotic systems using swarm intelligence algorithms," *Neurocomputing*, vol. 172, pp. 322–336, Jan. 2016. doi: 10.1016/j.neucom.2015.03.099.

[21] W. Deng, H. Zhao, X. Yang, J. Xiong, M. Sun, and B. Li, "Study on an improved adaptive PSO algorithm for solving multi-objective gate assignment," *Appl. Soft Comput.*, vol. 59, pp. 288–301, Oct. 2017. doi: 10.1016/j.asoc.2017.06.004.

[22] W. Deng, R. Yao, H. Zhao, X. Yang, and G. Li, "A novel intelligent diagnosis method using optimal LS-SVM with improved PSO algorithm," *Soft Comput.*, vol. 23, no. 7, pp. 2445–2462, 2017. doi: 10.1007/s00500-017-2940-9.

[23] H. Zhao, M. Sun, W. Deng, and X. Yang, "A new feature extraction method based on EEMD and multi-scale fuzzy entropy for motor bearing," *Entropy*, vol. 19, no. 1, p. 14, 2017. doi: 10.3390/e19010014.
[24] Z. Wang, Y. Qin, C. Hu, M. Dong, and F. Li, "Fuzzy observer-based prescribed performance control of vehicle roll behavior via controllable damper," *IEEE Access*, vol. 7, pp. 19471–19487, 2019. doi: 10.1109/ACCESS.2019.2896588.
[25] X.-L. Li, "An optimizing method based on autonomous animats: Fish-swarm algorithm," *Syst. Eng.-Theory Pract.*, vol. 22, no. 11, pp. 32–38, 2002.
[26] J. M. Almendros-Jiménez, A. Becerra-Terón, and G. Moreno, "Fuzzy queries of social networks with FSA-SPARQL," *Expert Syst. Appl.*, vol. 113, pp. 128–146, Dec. 2018. doi: 10.1016/j.eswa.2018.06.051.
[27] D. P. F. Cruz, R. D. Maia, and L. N. De Castro, "A critical discussion into the core of swarm intelligence algorithms," *Evol. Intell.*, vol. 12, no. 2, pp. 189–200, 2019. doi: 10.1007/s12065-019-00209-6.
[28] S. Zhang, X. Zhao, C. Liang, and X. Ding, "Adaptive power allocation schemes based on IAFS algorithm for OFDM-based cognitive radio systems," *Int. J. Electron.*, vol. 104, no. 1, pp. 1–15, 2017. doi: 10.1080/00207217.2016.1178345.
[29] J. Bai, L. Yang, and X. Zhang, "Parameter optimization and application of support vector machine based on parallel artificial fish swarm algorithm," *J. Softw.*, vol. 8, no. 3, pp. 673–679, 2013.
[30] Q. Hu, D. Yu, and Z. Xie, "Neighborhood classifiers," *Expert Syst. Appl.*, vol. 34, no. 2, pp. 866–876, 2008. doi: 10.1016/j.eswa.2006.10.043.
[31] Q. Hu, W. Pedrycz, D. Yu, and J. Lang, "Selecting discrete and continuous features based on neighborhood decision error minimization," *IEEE Trans. Syst., Man, Cybern. B. Cybern.*, vol. 40, no. 1, pp. 137–150, Feb. 2010. doi: 10.1109/TSMCB.2009.2024166.
[32] Y. Chen, Z. Zeng, and J. Lu, "Neighborhood rough set reduction with fish swarm algorithm," *Soft Comput.*, vol. 21, no. 23, pp. 6907–6918, 2017. doi: 10.1007/s00500-016-2393-6.
[33] Q. H. Hu, H. Zhao, and D. R. Yu, "Fast reduction algorithm for symbols and numerical properties based on neighborhood rough sets," *Pattern Recognit. Artif. Intell.*, vol. 21, no. 6, pp. 730–738, 2008. doi: 10.3969/j.issn.1003-6059.2008.06.004.
[34] R. Jensen and Q. Shen, "Fuzzy-rough data reduction with ant colony optimization," *Fuzzy Sets Syst.*, vol. 149, no. 1, pp. 5–20, 2005. doi: 10.1016/j.fss.2004.07.014.
[35] L. Zou, X. Yang, J. Tan, H. Xu, and Y. Sun, "Fatigue life prediction of 5083 and 5A06 aluminum alloy T-welded joints based on the fatigue characteristics domain," *Frattura Integritá Strutturale*, vol. 12, no. 45, pp. 43–66, 2018. doi: 10.3221/IGF-ESIS.45.05.
[36] N. Sidhom, A. Laamouri, R. Fathallah, C. Braham, and H. P. Lieurade, "Fatigue strength improvement of 5083 H11 Al-alloy T-welded joints by shot peening: Experimental characterization and predictive approach," *Int. J. Fatigue*, vol. 27, no. 7, pp. 729–745, 2005. doi: 10.1016/j.ijfatigue.2005.02.001.
[37] S. Beretta and G. Sala, "A model for fatigue strength of welded lap joints," *Fatigue Fract. Eng. Mater. Struct.*, vol. 28, nos. 1–2, pp. 257–264, 2005. doi: 10.1111/j.1460-2695.2004.00849.x.

LI ZOU received the Ph.D. degree in mechanical engineering from Dalian Jiaotong University, in 2018. She is currently an Associate Professor with the Software Technology Institute, Dalian Jiaotong University. She has authored 15 papers in peer-reviewed international journals and conferences, such as *Entropy*, *Material Science and Engineering A*, *Soft Computing*, and the *Transactions of the China Welding Institution*. Her research interests include data mining, intelligent computing, and fatigue analysis.



HONGXIN LI was born in Anshan, Liaoning, China, in 1996. She received the bachelor's degree from the University of Technology and Science Liaoning, in 2018. She is currently pursuing the master's degree in software engineering with Dalian Jiaotong University. Her research interests include data mining and intelligent computing.

WEI JIANG was born in Jiujiang, Jiangxi, China, in 1996. He received the bachelor's degree from the Harbin University of Commerce, in 2017. He is currently pursuing the master's degree in software engineering with Dalian Jiaotong University. His research interests include data mining and intelligent computing.

XINHUA YANG received the Ph.D. degree in mechanical engineering from the Dalian University of Technology, in 2005. He is currently a Professor and a Doctoral Supervisor with the College of Materials Science and Engineering, Dalian Jiaotong University. He has authored over 25 papers in peer-reviewed international journals and conferences, such as *Entropy*, *Material Science and Engineering A*, *Soft Computing*, *Applied Soft Computing* journal, *Kybernetes*, *Symmetry-Basel*, and the *Transactions of the China Welding Institution*. His research interests include welding mechanics, structural reliability analysis, and intelligent technology and application.