

An Improved Firefly Algorithm for Feature Selection in Classification

Huali Xu¹ · Shuhao Yu¹ · Jiajun Chen¹ · Xukun Zuo¹

Published online: 24 January 2018

© Springer Science+Business Media, LLC, part of Springer Nature 2018

Abstract Feature selection functions as an important method of receiving data so as to make the amount of features decrease. While solving the issue of classifying there exists numerous features having no relevance and being unnecessary which have the potential of making classification performance decrease. Firefly algorithm (FA) functions as an efficient method to make computation which is efficient and progressive. Nevertheless, the conventional FA is easily fallen into the local optima which imposes unsatisfactory practice on feature selection. In this research, one proposal was put forward, the firefly algorithm that combines the binary firefly algorithm with opposition-based learning to select features in classification. Experiment outcomes indicate the fact that the means put forward surpasses PSO and the conventional firefly algorithm.

Keywords Firefly algorithm · Feature selection · Opposition-based learning · Classification · Evolutionary computation

1 Introduction

Feature selection is the combinational optimization problem [1]. The feature selection aims to choose a subset of variables that have the ability to portray the input statistic. Meanwhile, they can also reduce the noise or other irrelevant variables impact and accurate prediction outcomes can be offered [2].

In classification, data sets tend to include numerous features, usually involving unrelated and redundant characteristics. Nevertheless, in the light of substantial search space, unrelated and unnecessary features have no efficient usage for categorization. Moreover,

✉ Huali Xu
hl_xu@126.com

¹ School of Electronics and Information Engineering, West Anhui University, Lu'an 237012, Anhui, China

they can weaken the classification's achievement and increase the computation time, which is called "the curse of dimensionality" [3].

For the sake of dealing with this issue, a variety of feature selection techniques have got introduced. The main aim of feature selection consists in eliminating the irrelevant and unnecessary characteristics and choose relevant features from the large feature set. Generally, these methods fall into three categories: filter, wrapper, and embedded. The first one method has nothing to do with particular learning algorithms which make the data reasoning of the feature set so as to choose a differential subset of the features with no thought of internal relationship with the learning algorithm. Techniques of this kind involve the information gaining [4], documents recurrence [5], term strength information [6], Chi square [7], and odd ratio [8]. Widespread application was made so as to make feature selection lessen computational complexity, particularly under the circumstance of the extremely large spatial characteristic space, such as text. Wrapper methods which include study methods being one part of the assessment process. Some typical methods of wrapping methods involve sequential floating selection [9], sequential forward floating selection (SFFS) [10] and sparse logistic regression based methods [11]. The second one refers to the embedded method. Embedded techniques [12–14] involve fluctuating choosing, one part of the training phase, with the condition of no separation of statistic into test and training sets.

Considering the fact that meta-heuristic techniques are able to figure out problem-solving answer rapidly with enough research room with the help of several search methods around the world, recently lots of researchers have made efforts to apply meta-heuristic technique for the sake of coping with the feature selection issue. For example, Huang et al.(2008) [15] put forward a brand-new PSO-SVM modeling which mixed the PSO and served as a backbone for vector machines(SVMs) for the sake of enhancing its categorization correctness and meanwhile making the input feature sub-group selection become better. Neshatian et al. (2009) [16] applied Genetic Programming for Feature Subset Ranking in Binary Classification Problems. Chen et al. (2010) [17] have proposed a novel feature selection that hybridize ant colony optimization (ACO) with rough set theory which can get a higher accuracy. Xue et al. (2016) [18] have presented a PSO which has multiple goals for feature selection.

Firefly algorithm (FA) is a meta-heuristic search on the basis of swarm-intelligence and upgrade method. It encourages the flashing and communication act of fireflies. Because it's very simple, able to share information, quickly come together and it's on the basis of population, at the beginning, some modified variants were proposed that have been successfully explored in various fields such as continuous optimization [19], multimodal [20], constrained optimization [21], and later in real-world problems such as non-convex economic dispatch problems [22], clustering [23], combinatorial optimization [24], image compression [25]. Up to now, FA has got successful usage for lots of challenging upgradation issues as well as NP-hard issues (Yang 2008). In this paper, FA gets applied for feature selection issue. Nevertheless, several shortcomings of FA for feature selection task do exist. At the beginning, different initialization strategies in FA perform differently in different problems. Secondly, in the case that the value of the gbest shows no variations of a defined amount of repetition, it will suffer from being fastened in regional fitness at early stages, then swarm diversity is decreased, Position mutation is needed for the sack of enhancing its search ability and diversity.

1.1 Goals

In this paper, we use several adjustments to avoid the swarm stagnation into local optimal and the premature: (1) begin with a population of excellent methods we apply opposition-based strategy for population initialization, (2) in case that the value of gbest show no variations for a fixed amount of iterations, we use opposition position of the gbest firefly to take the place of the most unfavorable fit particle.

1.2 Organisation

Other part of this research is constructed as we can see: at the very beginning, we present the current research. A related work follows in Sect. 2. Detailed definition of this methodology follows in Sect. 3. Section 4 explains experiment outcomes and comparison of differential models. At last, conclusions are arrived at in Sect. 5.

2 Related Work

2.1 Firefly Algorithm

FA, an algorithm, serves as a target of vital significance formulated through Yang (2008) which gets motivation from social insects called fireflies. Fireflies belong to insects, the primary features of which involves admirable flashing lights. The fireflies flashing patterns which is produced by a bioluminescence procedure, enjoy a special place for each of 2000 current living fireflies species. Two main purposes of these flashing are to attract the potential prey and to mate partners.

To simplify the FA, there exist the following 3 idealized rules (1) The total amount of fireflies have one sex. As a result, sexual attraction can happen to them. (2) The level of attractiveness has positive relationship to the brightness, meaning that the one with less bright tends to be attracted by the one having more light. Under the circumstance that there is no one with the largest amount of light, random attractiveness will happen. (3) The light intensity of a firefly gets influenced by the outlook of the unbiased function. The pseudocode is given in Algorithm 1.

The foundation of the attractiveness and light intensity function as significant matters. Light intensity can be formulated as follows:

$$I = I_0 e^{-\gamma r_{ij}^2} \quad (1)$$

In this equation, I_0 refers to the light intensity at the beginning.

The attractiveness of a firefly results from the light intensity. The attractiveness can be approached as follows:

$$\beta = \beta_0 e^{-\gamma r_{ij}^2} \quad (2)$$

where β_0 is a constant of attractiveness at $r = 0$. γ is light absorption coefficient, which is fixed as 1.0 in FA.

The distance between any two fireflies i and j at x_i and x_j , is the Cartesian distance as following:

$$r_{ij} = \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2} \quad (3)$$

The act of a firefly i is admired by another more appealing (with more light) firefly j which gets impacted

$$x_i = x_i + \beta e^{\gamma r_{ij}^2} (x_j - x_i) + \alpha \left(\text{rand} - \frac{1}{2} \right) \quad (4)$$

From what we can see above, the second one results from attraction. The third one is the disordered parameter, rand is a number producer in disorder having even distribution in $[0, 1]$. Under most circumstances of completion, value is taken as $\beta_0 = 1$ and $\alpha \in [0, 1]$. For more details of the firefly algorithm it can be seen in Yang (2009) and Gandomi et al. (2013).

2.2 Opposition-Based Learning (OBL)

Opposition-based learning (OBL) is originally put forward by Tizhoosh (2006) which gets used to get the optimum solution of a given problem by taking the corresponding method and the anti-solution at the same time.

In general, meta-heuristic algorithms begin with several original solutions (original population) with the purpose of enhancing the group to end up with the global optimal solution(s). Procedure of the searching comes to the end when several previously fixed standards got met. Without previous knowledge about the method, the commonplace initialization will begin with random sampling distribution on the whole range. Under the worst circumstance, in case that the best method is too far away from the random sampling, the computation period will last long. So, complexity of time is on the rise. Imagine that we simultaneously test a method and the totally opposite one, the fitter one is able of being chosen as the initial method. Actually, as the previous research of Tizhoosh indicates that fifty percent of the time a rough thought has long distance from being called ideal solution than the adverse thoughts. Consequently, beginning with the original groups that involve the best of the two guesses tends to be much better. In this research, first, a brand-new form of the OBL strategy was adopted to begin with an excellent likely methods, second, we apply it to make the search methods more various under the circumstance of backwater of the best firefly. The concept of adverse amount, opposition-based initialization is given the following explanation:

Definition 1 Let $x \in [m, n]$ be a real number. The opposition number \tilde{x} is defined by

$$\tilde{x} = m + n - x \quad (5)$$

Correspondingly, the adverse point in D-dimensional space are described as bellowing.

Definition 2 Let $X = (x_1, x_2, \dots, x_d)$ be a point in D-dimensional space, in which $x_1, x_2, \dots, x_d \in R$ and $x_i \in [m_i, n_i], \forall i \in \{1, 2, \dots, D\}$ The opposition point

$$\tilde{x} = m_i + n_i - x_i \quad (6)$$

Definition 3 Assume that $X = (x_1, x_2, \dots, x_d)$, a point in D-dimensional space, be a candidate solution. Imagine $f(\cdot)$ is a fitness function which gets applied to assess the candidate's fitness. In accordance with the definition of the opposite point, $\tilde{X} =$

$(\tilde{x}_1, \tilde{x}_2, \dots, \tilde{x}_d)$ is the opposite $X = (x_1, x_2, \dots, x_d)$. Now, in case that $f(\tilde{X}) \geq f(X)$, the candidate solution X can be taken place by \tilde{X} , or else, keep moving with X . Therefore, the point and its opposite point are assessed to keep moving with the fitter one in the same time.

Algorithm1. Opposition-based population initialization

Input: initialize $\{X\}$ =Randomly generated N fireflies.

Output: X : a set of N fireflies.

Algorithm:

- 1: for $i=1$ to N do
 - 2: for $j=1$ to d do
 - 3: $\tilde{X}_{ij} = u_j + l_j - x_{ij}$
 - 4: end for
 - 5: end for
 - 6: $\{OX\} = \{X\} \cup \{\tilde{X}\}$ /* $\{OX\}$ consists of 2S fireflies */
 - 7: Compute fitness of fireflies in $\{OX\}$ using Eq.(1).
 - 8: Sort $\{OX\}$ with fitness values.
 9. $\{X\} = \text{top}(\{OX\}/2)$ fireflies
 10. Return $\{X\}$
-

3 Description of the Proposed Algorithm (MFA)

In this part, the main content is the description of the raised FS algorithm. The main goal involves establishing a global search method. This method has both excellent behavior of coping with feature selection issue and easily implementation.

3.1 Encoding of Fireflies

Unlike the existing studies adopting the binary string, in this paper, we use a probability strategy (Algorithm 2) which can represent a feature selected into the feature subset and be applied as an encoded element. In this techniques various elements fashion a firefly which stands for a candidate method of this issue. Taking a data set with D features as one instance, the i th firefly in the swarm is symbolized with a D -bit real String as below:

$$X = (x_{i1}, x_{i2}, x_{i3}, \dots, x_{id}), \quad i = 1, 2, \dots, m \quad (7)$$

In this equation, m is the swarm size, and $x_{i,j} \in [0, 1]$ means the possibility of the j th feature being selected in the next subset.

For a firefly X_i , it can be decoding to a solution Z_i which is be established as follows:

$$Z_{i,j} = \begin{cases} 1, & x_{ij} \geq \text{rand} \\ 0, & \text{otherwise} \end{cases} \quad (8)$$

where $z_{ij} = 1$ stands for that the j -th feature is selected into the feature subset Z_i . For example, the firefly 10100010 with 8 features indicates that the 1st, 3rd and 7th features are chosen.

3.2 Fitness Function

The fitness function gets applied so as to assess the efficiency of categorization practice and the amount of features, in which the weight for the number of features is extremely tiny. The fitness function can be seen by Eq. (9).

$$\text{Fitness} = \text{ErrorRate} + \alpha \times \# \text{Features} \quad (9)$$

$$\text{ErrorRate} = \frac{FP + FN}{TP + TN + FP + FN} \quad (10)$$

Error Rate stands for the training classification error of the chosen features, Error Rate can be calculated by FP, FN, TP and TN. FP, FN, TP and TN represents false positives, false negatives, true positives, and true negatives, respectively. α indicates the related significance of the amount of the features. In this paper, α is made to serve an extremely small value to make sure that numbers of features are always smaller than Error Rate. So, classification performance can be calculated by Eq. (9), which can find out the feature subgroup with small categorization error ratio.

Algorithm 2 Converting continuous values to discrete(binary) values for feature selection

1: $S \leftarrow \emptyset$, S is the selected feature subset;
 2: for all $d_i \in x, i = 1, \dots, N$ do, x is the particle, d_i is the i th dimension of x ;
 3: if $d_i > \text{rand}$ then
 4: $s \leftarrow s \cup \{i\}$;
 5: end if
 6: end for

Algorithm 3 Opposition strategy

Input: x_b : Best firefly in the swarm.

x_w : Worst firefly in the swarm.

Output: x_w : Update worst fit firefly(opposite of the x_b).

1: $\text{temp} = x_b$
 2: for $j=1$ to d do
 3: $\text{temp}_j = u_j - l_j + \text{temp}_j$
 4: end for
 5: $x_w = \text{temp}$;
 6: return updated worst fit firefly.

 Algorithm 4 Proposed FA-based method for feature selection

Begin

Input: initialize all parameters

Output: feature subsets

1: Divide Dataset into a Training set and a Test set;

2: initializes the position and light of each firefly in the swarm;

Compute light intensity I_i at x_i

Calculate attractiveness(Eq.(1)).from the i th firefly to all the others.

3: while Maximum Iterations or stopping criterion is not met do

4: Evaluate the fitness of each firefly on the Training set;

5: For $i=1$ to Population Size do

6: For $j=1$ to Population Size do

7: Calculate attractiveness(Eq.(1))between i th and j th fireflies

8: If($I_j > I_i$)

9: Move firefly i using Eq.(4).

10: end if

11: Evaluate new solutions and update light intensity.

12: End for

13: End for

14: Rank the fireflies and find the current best.

15: end while

16: Calculate the classification accuracy of the selected feature on the Test set;

17: Return the position of the best firefly(the selected feature subset);

18: Return the training and test classification accuracies;

19: end

3.3 Proposed Method for Feature Selection

This part gives analysis of the raised technique in great detail. Its aim involves applying the fortified algorithm MFA for the feature selection in classification. The raised methodology can be seen below and an algorithmic flow can be seen in Algorithm 4.

For the sake of increasing the FA search capability and decreasing the local optima trapping probabilities, a brand-new adjustment techniques is presented (MFA). Mainly two thoughts exist in the algorithm. One is the opposition-based population initialization method which being used to improve the population diversity. The other is the opposition

strategy (Algorithm 3) which can encourage the whole firefly population to come straight to the best potential local or global individual.

4 Experimental Results and Analysis

4.1 Datasets and Parameter Settings

All experiments are conducted on the ten datasets (Table 1) selected from the UCI. These datasets appear various amount of features, classes and examples that are chosen to check on the raised algorithm. For each datasets, the examples fall into two sets in disorder: seventy percent being the training set and thirty percent being the check set.

Feature selection is binary problem, so representation of the firefly is an “n-bit” string, in which “n” is the whole amount of features in the dataset.

For each dataset, the experiments to examine the feature selection performance of each algorithm has been established for thirty independent times. The parameters are defined: $\beta_0 = 1$, $\gamma = 0.2$, $\alpha = 0.1$, population number is 30, and the largest iteration is 100.

As a wrapper method, the raised algorithms needs the learning algorithm. KNN is acknowledged as being easy and often applied learning algorithm gets adopted in the research and $K = 5(5 \text{ NN})$.

4.2 Comparison Results of MFA and Other Methods

Table 2 indicates the research outcomes by using FA, MFA and PSO to test the accuracy of the raised algorithm on 30 runs. In Tables 2 and 3, the first one, the second one, the third one represents the mean, the excellent and the standard deviation classification accuracy received from the thirty runs on each test set.

According to Table 2, we can see that MFA achieved the most excellent classification performance in the three algorithm in majority of the datasets. The classification performance of MFA was similar to FA on one dataset, better than FA on seven datasets, but worse than FA on three datasets, which proves that the MFA has the advantage of the finding out the feature space adaptively more excellent in comparison with the other techniques. As show in Table 2, we can also see that the stand deviation is minimum for

Table 1 Datasets

Dataset	No. features	No. classes	No. instances
Wine	13	3	178
Vehicle	18	4	846
WBCD	30	2	569
Lung	56	3	32
Movement	90	15	360
Musk1	166	2	476
Madelon	500	2	4400
Zoo	17	7	101
German	24	2	1000
Ionosphere	34	2	351

Table 2 Comparisons between FA, MFA and PSO

Dataset	FA		MFA		PSO	
	Ave (Best) %	SD	Ave (Best) %	SD	Ave (Best) %	SD
Wine	96.10 (100)	2.46	95.26 (98.77)	1.75	95.95 (100)	2.20
Vehicle	81.70 (85.24)	1.84	85.30 (87.01)	0.95	85.07 (87.01)	0.87
WBCD	93.05 (94.74)	2.43	93.98 (94.97)	0.85	93.34 (94.74)	0.52
Lung	79.00 (90)	8.77	78.40 (90)	6.44	74.00 (90)	7.48
Movement	94.28 (95.19)	0.44	94.58 (95.19)	0.38	94.51 (95.31)	0.33
Musk1	81.57 (89.51)	3.12	84.94 (89.51)	2.52	84.60 (91.61)	2.23
Madelon	77.22 (89.23)	9.01	78.86 (84.23)	3.15	76.93 (79.87)	1.54
Zoo	94.08 (97.14)	1.62	95.52 (97.14)	0.72	95.47 (97.14)	0.77
German	67.35 (72)	6.46	68.47 (70.67)	1.45	68.47 (72)	1.89
Ionosphere	88.06 (92.38)	2.15	87.27 (91.43)	1.84	88.86 (93.33)	2.04

The number in bold which is on behalf of the optimal value

Table 3 Average numbers of feature selected from the different datasets

Dataset	FA	MFA	PSO
Wine	3.58	6.84	7.24
Vehicle	3.60	10.16	9.06
WBCD	3.04	3.46	11.82
Lung	2.86	6.74	23.58
Movement	11.84	27.46	37.96
Musk1	16.66	76.54	74.26
Madelon	16.32	203.32	235.20
Zoo	3.20	6.46	8.16
German	2.34	12.76	12.42
Ionosphere	3.36	3.26	8.74

the MFA in five datasets than FA and PSO, which proves that MFA outperforms the other algorithms in its stability, and ability to reach optimal.

From Table 3, it's obvious that features subset chosen by MFA were larger than FA on all the ten datasets but smaller than PSO on 7 of 10 datasets. The main reasons are the classification performance that tends to be more significant than the amount of features as considered.

To prove the effectiveness of the MFA algorithm, three existing feature selection algorithms such as ReliefF, sequential forward selection (SFS) and MIM are applied on the same datasets such as the German, Ionosphere, Vehicle and Lung. From the Fig. 1 we can see that the MFA provide higher classification accuracy rates compared to existing feature selection.

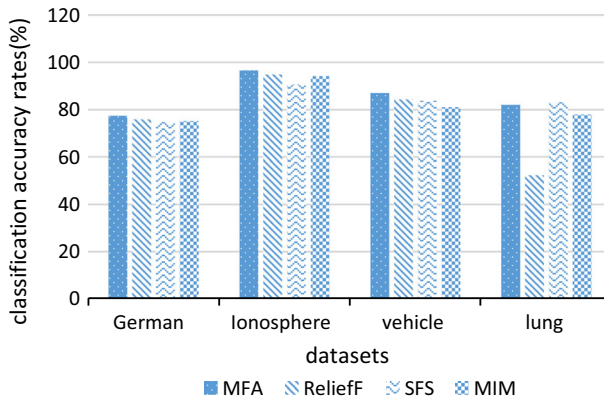


Fig. 1 Classification accuracy rates using different feature selection algorithms on the datasets

5 Conclusion

In this work an improved firefly algorithm method is put forward for feature selection in wrapper mode. The continuous version form of firefly algorithm (FA) is transformed into the binary form by discrete coding. An improved FA employs opposition-based learning in population initialization and opposition strategy in the searching process which fastens the convergence rate to obtain the global optima. These experimental results on datasets indicate that the raised algorithm MFA can obtain better classification accuracy than other methods.

Acknowledgements This research was supported in part by the National Natural Science Foundation of China (No. 61075049), the Universities Natural Science Foundation of Anhui Province (No. KJ2014A277) and the Key Projects for Domestic Visiting and Training of Key Young Backbone Teachers (No. gxfzD2016189), the Universities excellent Young Talents Foundation of Anhui Province (No. gxyq2017056, gxyqZD2016249).

References

- Xu, Z., Jin, R., Lyu, M. R., & King, I. (2009). Non-monotonic feature selection. In *Proceedings of the 26th international conference on machine learning, ICML* (pp. 1145–1152).
- Guyon, I., & Elisseeff, A. (2003). An introduction to variable and feature selection. *Journal of Machine Learning Research*, 3, 1157–1182.
- Dash, M., & Liu, H. (1997). Feature selection for classification. *Intelligent Data Analysis*, 1(3), 131–156.
- Quinlan, J. R. (1986). Induction of decision trees. *Machine Learning*, 1(1), 81–106.
- Liu, L., Kang, J., Yu, J., & Wang, Z. (2005). A comparative study on unsupervised feature selection methods for text clustering. In *Proceedings of 2005 IEEE international conference on natural language processing and knowledge engineering, IEEE NLP-KE'05* (pp. 597–601).
- Yang, Y. (1995). Noise reduction in a statistical approach to text categorization. In *Proceedings of the 18th annual international ACM SIGIR conference on research and development in information retrieval, SIGIR Forum* (pp. 256–263).
- Li, Y., Luo, C., & Chung, S. M. (2008). Text clustering with feature selection by using statistical data. *IEEE Transactions on Knowledge and Data Engineering*, 20(5), 641–652.
- Mengle, S. S., & Goharian, N. (2009). Measure feature-selection algorithm. *Journal of the American Society for Information Science and Technology*, 60(5), 1037–1050.

9. Pudil, P., Novovičová, J., & Kittler, J. (1994). Floating search methods in feature selection. *Pattern Recognition Letters*, 15(11), 1119–1125.
10. Nakariyakul, S., & Casasent, D. P. (2009). An improvement on floating search algorithms for feature subset selection. *Pattern Recognition*, 42(9), 1932–1940.
11. Somol, P., Pudil, P., Novovičová, J., & Pačlık, P. (1999). Adaptive floating search methods in feature selection. *Pattern Recognition Letters*, 20(11), 1157–1163.
12. Guyon, I., & Elisseeff, A. (2003). An introduction to variable and feature selection. *Journal of Machine Learning Research*, 3, 1157–1182.
13. Langley, P. (1994). Selection of relevant features in machine learning. In *Proceedings of the AAAI Fall symposium on relevance*.
14. Blum, A. L., & Langley, P. (1997). Selection of relevant features and examples in machine learning. *Artificial Intelligence*, 97(1–2), 245–271.
15. Lin, S. W., Ying, K. C., Chen, S. C. H., & Lee, Z. J. (2008). Particle swarm optimization for parameter determination and feature selection of support vector machines. *Expert Systems with Applications*, 35(4), 1817–1824.
16. Neshatian, Kh. & Zhang, M. J. (2009). Genetic programming for feature subset ranking in binary classification problems. In *Proceedings of the 12th European conference on genetic programming (EuroGP 2009), Lecture notes in computer science*, 5481, 121–132.
17. Colorni, A., Dorigo, M., Maniezzo, V., et al. (1991). Distributed optimization by ant colonies. *Proceedings of the First European Conference on Artificial Life*, 142, 134–142.
18. Xue, B., Zhang, M. J., & Browne, W. N. (2013). Particle swarm optimization for feature selection in classification: A multi-objective approach. *IEEE Transactions on Cybernetics*, 43(6), 1656–1671.
19. Yang, X. S. (2010). Firefly algorithm stochastic test functions and design optimization. *International Journal of Bio-Inspired Computation*, 2(2), 78–84.
20. Yang, X. S. (2009). Firefly algorithms for multimodal optimization. In *5th symposium on stochastic algorithms, foundations and applications, SAGA 2009* (pp 169–178).
21. Łukasik, S. & Zak, S. (2009). Firefly algorithm for continuous constrained optimization tasks. In *Computational collective intelligence: Semantic web, social networks and multiagent systems, ICCCI 2009* (pp 97–106) Springer.
22. Yang, X. S., Hosseini, S. S. S., & Gandomi, A. H. (2012). Firefly algorithm for solving non-convex economic dispatch problems with valve-loading effect. *Applied Soft Computing*, 12(3), 1180–1186.
23. Senthilnath, J., Omkar, S. N., & Mani, V. (2011). Clustering using firefly algorithm: Performance study. *Swarm and Evolutionary Computation*, 1(3), 164–171.
24. Fister, I. Jr., Yang, X. S., & Brest J. (2012). Memetic firefly algorithm for combinatorial optimization. In *5th international conference on bioinspired optimization methods and their applications, BIOMA 2012*, (pp 75–86).
25. Horng, M. H. (2012). Vector quantization using the firefly algorithm for image compression. *Expert Systems with Applications*, 39(1), 1078–1091.



Huali Xu born in 1977. Master. Associate professor. Her main research interests include data mining, Intelligent computing, and machine learning.



Shuhao Yu is a professor in School of Electronic and Information Engineering, West Anhui University, Lu'an, China. He received the M.S. degree from Hohai University, Nanjing, China, in 2006, and Ph.D. degree at the Hefei University of Technology, Hefei, China, in 2015.



Jiajun Chen is an associate professor, her main research interests include rough set, machine learning and data mining.



Xukun Zuo born in 1978. Master. Associate professor. His main research interests include data mining, Intelligent computing, and machine learning.