Journal of Theoretical and Applied Information Technology
10th April 2013. Vol. 50 No.1

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

DESIGN OF NICHE PSO REDUCE ALGORITHM FOR HYBRID ATTRIBUTES BASED ON NEIGHBORHOOD ROUGH SETS

[1]BAITING ZHAO, [2]XIAOFEN JIA

[1,2]Electrical and Information Engineering College, Anhui University of Science and Technology, Huainan, China, 232001

E-mail: [1]zhaobaiting@126.com, [2]jxfzbt2008@163.com

ABSTRACT

A reduce algorithm based on the neighborhood granulation and niche Particle Swarm Optimization (PSO) algorithm is proposed for the reduction of the real decision system with numerical attributes. In this scheme, a rough model is used based on the neighborhood equivalence. The indiscernibility relation is measured by the neighborhood relation, and the universe spaces are approximated by the neighborhood information granules. The use of the niche technology can avoid the preconvergence of the PSO. The select of fitness function and the adaptive across probability are designed, and the reduction algorithm is presented as well. Furthermore, the dependency function is used to evaluate the significance of the numerical attributes. Experimental results demonstrate the validity and feasibility of the proposed algorithm, in application to a classical data set and four University of California at Irvine (UCI) machine learning databases.

Keywords: *Rough Sets(RS); Reduce; Neighborhood Granulation(NT); Niche Technology(NT); Partical Swarm Optimization(PSO)*

1. INTRODUCTION

Rough set is proposed by Pawlak [1] in 1982. Pawlak rough set indiscernibility relation is based on the strict equivalence relations and equivalence classes, which only suited to deal with the symbolic variable. When deal with hybrid data including numerical attributes, commonly discrete the numerical attributes into symbolic attributes. The conversion will inevitably cause loss of information [2], and calculate results dependent on the effect of the discrete. In order to solve this problem many models be proposed including the fuzzy rough set model [3], similar relationship rough set model [4] and neighborhood relationship model [5]. The neighborhood rough set model granulate the universe by neighborhood relationship, which is intuitive, easy to understand, and can deal with the numeric attributes directly without discretization. Comparing with classical Pawlak rough set, neighborhood rough set model can analysis the numerical data directly without discretization, and also better than fuzzy information entropy method and similar relationship rough set, which expand the application scope of the classical rough set theory.

Knowledge reduction has proven to be NP-hard problem [6]. The existing reduction algorithm is the minimum reduction, which starting from the rough set of nuclear, and constructed reduction using the heuristic search. It is not applicable to all of the knowledge representation system [7]. The greedy reduction algorithm proposed in [5] in most cases is not the minimum reduction. In order to reliably solve global optimization problems, people analog the natural phenomenon, and developed a series of biomimetic intelligent optimization algorithms such as simulated annealing, genetic algorithm [8], particle swarm optimization [9], and so on. But the reduction of the genetic algorithm in [8] has slow rate of convergence, and easily fall into local optimum. PSO proposed in [9] introduces mutation simulated genetic algorithm and formally strengthen the diversity of the particles but upset the balance of the initial population distribution on the feasible region, which eventually led to the convergence of a single model. The niching PSO reduction algorithm based on neighborhood granulation is proposed in this paper. The introduction of niche technology can solve the premature convergence problem of PSO, and avoid the sensitivity and single convergence of

Journal of Theoretical and Applied Information Technology
10th April 2013. Vol. 50 No.1

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

algorithm's parameters on. The simulation results show that the algorithm better than the literature [3,8-11] on the overall performance.

2. THE ROUGH SETS MODEL OF NEIGHBORHOOD GRANULATION

Rough sets is based on equivalence and division. The discourse is divided into several disjoint subset. It causes some samples have the same characteristic but the different decision. And it is difficult to get exactly the same value of the sample to the classification described by numerical value. The neighborhood rough sets using neighborhood relation to substitute the equivalence, and it can make the samples which have the similar value have the same decisions.

Definition 1. Let $NDT = <U, A, D>$ be a given neighborhood decision table. The dependence of the decision attributes D to the condition attributes B is $\gamma_B(D) = Card(\underline{N_B}D)/Card(U)$, and it can be marked as $B \Rightarrow \gamma D$.

In neighborhood decision system $NDT = <U, A, D>$, let $a \in B \subseteq A$, if $\gamma_B(D) > \gamma_{B-a}(D)$, we define a is the necessary attribute in the B which relative to the classification decision D, else a is the unnecessary attribute. If $\forall a \in B$ are necessary attributes, then B is independent.

Definition.2 Let $NDT = <U, A, D>$ be a given neighborhood decision table and $B \subseteq A$, B is a reduction of A, if B satisfy the condition sufficient condition $\gamma_B(D) = \gamma_A(D)$ and necessary condition $\forall a \in B$, $\gamma_{B-a}(D) < \gamma_B(D)$.

The first condition can ensure the positive region of B equal he positive region of A, so B can keep sufficient system information. The second condition can ensure every attribute is Necessary and there's no free attribute in B.

3. THE NICHE PARTICLE SWARM OPTIMIZATION REDUCE ALGORITHM

The particle swarm algorithm is proposed and improved by Kennedy and Eberhart [12] based on swarm intelligence. In order to improve the global search performance of PSO and to improve the effectiveness of search and optimization efficiency, niche technology is introduced into the PSO to form a niche PSO algorithm. The most basic ideas of the niche PSO algorithm is between the different species in isolated geographic niche in the separation of biology that is not competition, and so the exchange of information and independent evolution of the concept of transplantation to the PSO in order to maintain the diversity of the PSO algorithm in the iterative process.

3.1 Binary coding of particle swarm

PSO can not deal with data directly, must be coded into the particle string structure data in the search space. Fixed length string of binary symbols are used to represent the particle groups of individuals. Randomly generated the m length n binary symbol string x_{id} composed by{0,1} binary set. x_{id} Is initial position for particle swarm. Randomly generated v_{id} as the initial speed.

The binary string corresponding to a conditional attribute. If a certain value is 1, then select its corresponding condition attribute, else a value is 0, means they did not select its corresponding conditional attribute. Each microparticle location of the individual corresponds to the conditional attribute space in a subset of attributes, the attributes selected as a possible solution. In order to ensure convergence of the algorithm, to accelerate the search speed, and nuclear properties corresponding bit is always 1, and other random value.

3.2 Selection of fitness function

The fitness function is the only certainty indicators to evaluate the adaptability of the particle position, so the form of the fitness function directly determines the evolutionary behavior of the group. According to the actual requirements of the knowledge reduction problem, we define the fitness function as follows [10]:

$$F(x) = \beta * \frac{n - card(x)}{n} * \frac{1}{1 + e^{\alpha(\gamma_0 - \gamma_x(D))}}$$
$$= \beta * f(x) * P(x)$$

f(x) is objective function; card(x)is the number of 1 in the particles position, that the particles contained condition attributes number ; n is the length of the particle position, that the number of condition attributes; β is penalty factor; P(x) is penalty function; $\gamma_x(D)$ is the dependence of the condition attributes to decision attributes D; γ_0 is preset threshold. Three-part structure of the fitness function, we can keep the decision attribute on the overall condition of property dependence on the

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

same case, and get the best search performance in knowledge reduction.

3.3 Speed and position updates

Calculate the new velocity and position of particle. To each generation, the d-dimensional changes according to the following equation:

$$v_{id} = \omega v_{id} + c_1 rand()(p_{id} - x_{id}) + c_2 Rand()(p_{gd} - x_{id}),$$

$$x_{id} = \begin{cases} 1 & \rho_{id} < sig(v_{id}) \\ 0 & \rho_{id} \geq sig(v_{id}) \end{cases}$$

ω is weighting function; c_1 and c_2 are acceleration constant; rand( )and Rand( ) are random numbers changes in the range of [0,1]; p_{id} is the best position(fitness maximum) of particle i; p_{gd} is the best location for groups of all particles; each component of ρ_{id} are random numbers in [0,1] range values; sig( ) is sigmoid function.

The first part of formula is the speed of the previous generation in the particle. The particle speed is limited to the maximum speed $v_{\max}$. the second part of cognition means the the thinking of particulate. The third part is the social part, which means information sharing and mutual cooperation between the particles.

3.4 Crossover operator

Randomly select $m * P_c$ pairs particles according to the pre-set probability Pc. Cross each position of each pairs of particles on the same probability. Let the position of particles after cross is $x_1' = a_{11}' a_{12}' \cdots a_{1n}'$ and $x_2' = a_{21}' a_{22}' \cdots a_{2n}'$, the operation is described as follows:

$$O(P_c, rand): a_{1i}' = \begin{cases} a_{1i}, rand > 0.5 \\ a_{2i}, rand \leq 0.5 \end{cases},$$

$$a_{2i}' = \begin{cases} a_{2i}, rand > 0.5 \\ a_{1i}, rand \leq 0.5 \end{cases}$$

To calculate the new generated particle fitness value, if one or two fitness value better than the corresponding particles before cross, then accept them as a new particle. Otherwise, return to the cross before the state. If the fitness of the new particles is increased, then the best position to update the individual best position and groups.

From the groups evolutionary process, the crossover probability should be decreased gradually with the evolutionary process, and ultimately tends to a stable value, so as not to impact on the stability of the algorithm. The crossover probability formula as follows:

$$P_c = \frac{1}{1 + e^{\xi * G}} + \psi$$

G is evolution algebra; ξ、 ψ are constant.

3.5 Niche optimization

Niche technology is to maintain the diversity of the population, to find multiple optimal solutions, in accordance with the population range R of niche radius R divided particles, each niche population independent search for the optimal solution, the global optimal solution in the best of all populations generated in the solution. The independent evolution of the various groups to avoid all particles as a population into a local optimum. The basic idea of niche optimization are as follows:

(1) If the particle velocity increment is too small, then generated random number v_{id} in range [0,1] as the speed of the particles, forcing the particle movement and prevent fall into local minima.

(2) If the two particles in close proximity, the poor particles will be restarted , which can ensure the same position has only one fine individual and limiting the distance between the particles to spread the particles over the entire search space.

(3) The poor particle will be restarted in the whole particle swarm and giving up the search in this part space to speed up the search.

4. REDUCTION ALGORITHM

Create a particle swarm X, randomly generated m particles, and initialize the position and velocity of the particle swarm. To the nuclear properties, the corresponding particles take 1, other bits are 0 or 1, which can speed up the convergence of the algorithm. Detail description shown in Algorithm 1.

Algorithm 1: Niche PSO Reduction Algorithm Based on Neighborhood Granulation

Input: $NDT = <U, A, D>$ and the initial population *X;* Output: p_{id} , p_{gd}

1： $\varnothing \rightarrow p_{id}, p_{gd}$, For each $x_i \in X$

2： Calculating v_{id} , calculating x_{id}

3： If the velocity increment is too small, then restart and go to step 2

Journal of Theoretical and Applied Information Technology
10th April 2013. Vol. 50 No.1

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

4: Calculating particle fitness

5: If particles euclidean distance less than the given value, then restart and go to step 2

6: End for

7: Search p_{id} , p_{gd} , calculating $P_c = \frac{1}{1+e^{\alpha G}} + \beta$,cross, calculating particle fitness

8: If $F(x)$ increase, search p_{id} , p_{gd} ; Else back to the state before the cross

9: IF fitness not improve continuous or get max flying algebra, the stop and output p_{id} , p_{gd}

10: Else go to step 2

5. EXPERIMENTS

The data in [13] is used for experiment to verify the validity of the algorithm. Niching PSO reduction algorithm based on neighborhood granulation can obtain the optimal reduction without discretization and previous knowledge. Tab. 1 shows the results of proposed niching PSO reduction algorithm to calculate he minimum reduction. The calculate parameter values shows as follows:m=30, $\alpha = 15$, $\beta = 2$, $\gamma_0 = 0.9$, $\xi = 0.02$, $\psi = 0.4$ and $\delta = 0.15$. The results show the best position of each generation, the best position fitness, flight algebra and time-consuming. In this case, the 4th generation groups get the best position and for more than ten generations does not change, and obtain the minimum reduction {C2,C4,C7}, nuclear properties is C7. Relative to [11], the proposed algorithm do not need preprocess of the data and directly calculate the minimum reduction.

The classical data in [10] are used for experiment, obtained two sets of minimum relative reduction {a1, a4, a5, a9} and {a4, a5,, a6, a9}, which better than the genetic algorithm in [10] which only obtained reduction {a1, a4, a5, a9}. The used of adaptive crossover probability take the better convergence.

Tab.1 Resuls of PSO reduction algorithm

Groups optimal position	Fitness	Flying algebra	Time(s)
1101001	0.7007	1	0.024964
1101001	0.7008	2	0.048611
1101001	0.7008	3	0.072955
0101001	0.9344	4	0.097694
0101001	0.9344	5	0.122093
0101001	0.9344	6	0.146928
0101001	0.9344	7	0.171351
0101001	0.9344	8	0.195159
0101001	0.9344	9	0.219076
0101001	0.9344	10	0.242609

Further test the effectiveness of the reduction algorithm, the database of UCI is used including the "Heart1", "Thyroid", "Vote" and "Wdbc". Support vector machine (SVM) classifier used as evaluation function. The data sample is divided into training and test set. The raw data and reduction data are used to train the classifier. The prediction accuracy is used to evaluate the reduction quality. PSO algorithm is still using the previous parameters in experiment. The SVM of "Vote" used linear kernel function, and other used spline function. The experimental results shown in Table 2.

Tab.2 Result of reduction

Data	Sampl es	Attri butes	Prediction accuracy (raw)	Prediction accuracy (reduction)
Heart1	270	13	74.01	81.41
Thyroid	9172	29	63.68	67.77
Vote	435	16	95.33	93.63
Wdbc	569	31	91.97	90.31

After reduction the prediction accuracy of "Vote" and "Wdbc" declined compared to the raw data prediction. It is due to the result is not optimal for the limit by flying algebra in complex conditions. In addition data noise can also cause a decline in prediction accuracy. The prediction accuracy is improved because the raw data redundancy attributes affect decision-making. Reduction can reduce the impact on the decision-making to improve the prediction accuracy. Experimental results show that the reduction algorithm used in this paper is effective to reduce the number of attributes, and basically does not change the prediction accuracy, effectively reduce the redundant data of the database to achieve the purpose of reduction.

6. CONCLUSIONS

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

Based on the neighborhood relations the universe space is granulated, and the application range of classical rough set is extended. An niche particle swarm optimization reduction algorithm is proposed for numerical attributes. PSO has the characteristics of global optimization and can solve the difficult problems of the existing algorithms. Experimental analysis shows that this algorithm is an effective method for solving mixed data knowledge reduction. But PSO is time-consuming when dealing with huge amounts of data, so the rapid reduction algorithm is the future research directions. In addition, the choice of the neighborhood operator and the impact on system performance requires further in-depth discussion.

ACKNOWLEDGEMENT

Supported by the Youth Foundation of Anhui University of science & technology of China under Grant No.2012QNZ06 and the Doctor Foundation of Anhui University of science & technology of China under Grant No.11223.

REFERENCE

[1] Pawlak Z,"Rough sets",International Journal of Computer and Information Sciences, 1982, 11(5):341-356.

[2] Jensen R, Shen Q,"Fuzzy-rough attribute reduction with application to web categorization",Fuzzy Sets and Systems, 2004, 141(3):469-485.

[3] Qinghua H, Daren Y, Zongxia X,"Information-preserving hybrid data reduction based on fuzzy-rough techniques", Pattern Recognition Letters,2006,27(5):414-423.

[4] Slowinski R, Vanderpooten D,"A generalized definition of rough approximations based on similarity", IEEE Trans. on Knowledge and Data Engineering,2000,12(2):331−336.

[5] Qinghua H, Daren Y, Zongxia X,"Numerical attribute reduction based on neighborhood granulation and rough approximation", Journal of software,2008,19(3):640-649.

[6] Skowron A, Rauszer C,"The discernibility matrices and functions in information systems", Intelligent Decision Support: Handbook of Applications and Advances of the Rough Sets Theory . Dordrecht, Netherlands: Kluwer Academic Publishers,1992:331-362.

[7] Wang J, Miao D Q,"Analysis on Attribute Reduction Strategies of Rough Set", Journal of Computer Science and Technology, 1998, 13(2):189-192.

[8] Yulig Y, Ye SH,"Rough set reduction for hybrid data based on genetic algorithm", Journal of harbin institute of technology. 2008,40(5):683-687.

[9] Yulig Y, Ye SH,"Research on hybrid decision system reduction algorithm", Journal of system simulation,2007,13(19):2988-3004.

[10] Zhi T, Baodong X, Dingwei W,"Rough set knowledge reduction approach based on GA", Systems engineering,2003,21(4):116-122.

[11] Chao ZH, Chunbao M, Dong S, etc,"Intelligent fault diagnosis method for complex equipment based on rough decision tree model", Acta armamentarii,2008,29(9):3211-8211.

[12] Kennedy J, Eberhart R C,"Particle Swarm Optimization", Proceedings of the IEEE International Conference on Neural Networks. Perth, Australia,1995:1942-1948.

[13] Jiyang Q, Changan Zh,"Research on intelligent fault diagnosis method of the equipments", Chinese journal of scientific instrument, 2006,27(10):1270-1275.