A Genetic Algorithm with Local Searching Strategy and its Application

Yingchun Liu [1, a], Aitao Zhao [2,b]

[1]Shijiazhuang Vocational Technology Institute, Modern Education Technology Center Shijiazhuang, China

[2]Shijiazhuang Vocational Technology Institute, Department of Information Technology Shijiazhuang, China

[a]yingchunliu@139.com, [b]ait_z@163.com

Keywords: genetic algorithm, Markovain chain, convergence

Abstract. A new genetic algorithm wit h local searching strategy is proposed in t his paper , which over comes the disadvantage of t he GA. It combines GA wit h BP algorithm. The global convergence of the new genetic algorithm is p roved by using Markovian chain theory. And the new GA proves that it is efficient to model in nonlinear system of oil fields.

Introduction

Genetic algorithm (GA) is one of the important algorithm, in many fields has been applied [1]. GA is one kind has the global search ability of evolutionary algorithm, but its local search ability is bad [2]. [3] are given to solve this problem, the research results; Ge Hong [4] proposed GA and simulated annealing algorithm (SA) combined with solutions; Fogel D B [5] gives the GA and evolutionary programming (EP) fusion algorithm for neural network (NN) training problems, the author puts forward a local search strategy with GA, it basically is the NN training of back propagation (BP) algorithm and GA to form a kind of new hybrid algorithm, the application MaErKe markov chain theory proved the global convergence of the new algorithm, and gives examples of application.

Algorithm description

Because with back propagation (BP) network to solve nonlinear modeling of the system of the oil field, and BP algorithm easy to fall into local optimal solution, namely its global search ability is poor. The author mainly considers taking save optimal individual simple genetic algorithm and the BP algorithm are shirt-sleeve, constitute a new hybrid algorithm, the iterative procedure is as follows.

Initialization: buy small group scale Ns, total population size $Nt = (m + 1)\ Ns$, and gives the exchange and mutation probability Pc and Pm, the control precision, read in training data, select the true value coding way. (1) random generation small groups. (2) to small group each individual for initial value, using BP algorithm training k_{BP} step; (3) if at least one individual meets the demand of precision, which is over, or turn to (4); (4) with small body each chromosome gene in the maximum and the minimum of the parameters are regarded as the upper limit and lower limit. Base this, randomly generated m Ns a other individuals, so get a Nt individual total groups. (5) to the total group for reproduction, exchange and variation and calculation fitness value, save optimal individual, repeat k_{GA} step; (6) if the (5) step at least one individual meets the demand of precision, which is over, otherwise, the total group from select the Ns a good solution (including optimal individual) as a new small group, turn to (2).

Algorithm convergence analysis

Use homogeneous markovain chain for the tool to analysis the convergence of GA. Using homogeneous markovain chain to describe the simple genetic algorithm (SGA), state space is the group space, recorded as Λ, space dimension is $|\Lambda| = 2^{nl}$, the elements in the Λ is one ofλ_i, including n long string for l bit string, λ_i,= s_{ij}, individual number j $\in [1, n]$, group space transition probability change by reproduction, exchange and variation caused by three kinds of gene manipulation, three functions respectively by probability matrix R, C, M to describe, namely SGA of markovain chain transition probability matrix with P, P = RCM.

No matter how group initial distribution to SGA, of any state markovain chains are greater than 0 to the only limit distribution, can start from any state, in limited time to any state j, natural k - >∞, SGA can traverse the entire state space [6]. Namely,

Lemma: SGA composition of markovain chain is the ergodic.

Although the SGA with ergodic, but this doesn't mean that it must be converge to the global optimal solution. The following is GA global convergence concept.

Definition: Make F_k is moment k, F_k = max {$f(s_{ikj})$ |, j ∈[1, n]}, status to (group) when λ_i the biggest group fitness, make the F3 is for GA global optimal fitness, F3 = Max {f (sj) |, j ∈ [1, 2 nl]}, if and only if meet the

$\lim_{k\to\infty} p(F_k = F^*) = 1$ (According to the probability, it was established), (1)

GA is global convergence.

Theorem: Optimal save simple genetic algorithm (OMSGA) is the global convergence.

Proof: Setting Λ_0 is contains the optimal individual x* group (state) set, x* fitness F_0 = F *, due to the optimal save operation, once the state λ_i enter the Λ_0, λ_k (k = i + 1…) will be with probability 1 into Λ_0, namely Λ_0 is closed set. For additional insurance optimal operation, the state space increasing widely $\Lambda + \Lambda_0$, transfer matrix P 'increasing widely:

$$P' = \begin{bmatrix} \mathrm{Q} & 0 \\ \mathrm{T} & \mathrm{P} \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ \mathrm{T} & \mathrm{P} \end{bmatrix} \quad (2)$$

In the above equation: Q is a closed set Λ_0 of the transition probability, because Q = 1, so Λ_0 state is to attract state. For P '> 0, the lemma know, regardless of the initial group (state) λ_i, λ_i through the limited steps k_1, always can reach Λ_0, in after k_1 will meet:

$$\lim_{k\to\infty} p(F_k = F_0 = F^*) = 1 \quad (3)$$

So, it is global convergence of OMSGA. Theorem that realize GA global convergence of the basic strategy is to find the optimal solution preserved. The optimum bit string replication is a kind of commonly used to realize global convergence strategy.

Corollary: Taking local search strategy is global convergence of GA.

Proof: Setting Λ_0 is contains the optimal individual x* of small group (state) set, x* of fitness F_0 = F*. Because of its optimal save operation, similarly can card, with local search strategy is global convergence of GA.

Examples of application

Use a new hybrid algorithm of local search strategy genetic to model and predict the the actual seismic data of sandstone thickness of one oilfield a work area. There are 32 well in the work area. First using the work area the seismic data extraction 50 kinds of characteristic parameter, then use 32 well material, 50 kinds of seismic characteristic parameters correlation analysis, and finally choose maximum amplitude, average amplitude, purpose layer basic frequency, amplitude spectrum energy, instantaneous frequency and instantaneous phase, correlation function maximum, curve length a total of eight parameters as the input of the NN sample. In 29 Wells near and each take eight known as sample, a total of 120 training sample, in addition to 3 Wells as generalization performance test sample. BP network topological structure for 8 - 12-12-1, namely eight input unit, an output unit (sandstone thickness), two hidden layer are 12 unit. The new GA implementation process see figure 1.

For seismic data of large amount of data, and accuracy is high, so in the new GA, the decimal floating-point coding, which is helpful to reduce the amount of calculation and improve the accuracy of modeling. Exchange using heterogeneous arithmetic operation, the variation of random search strategy, random iterative direction to cauchy points with the distribution of the random variable to determine. Modeling of the objective function is absolute error function, and the fitness function is given. Take the new algorithm with the traditional BP network trained NN by the prediction result contrast see table 1. Table 1 shows, hybrid algorithm with a trained NN by the maximum relative prediction error for - 1. 46%, maximum generalization relative error is 8. 21%; BP network with the maximum relative prediction error is 2. 27%, maximum generalization relative error is - 14. 94%; And the new hybrid algorithm is better than the original algorithm to improve the operation speed of about 1. 6 times.

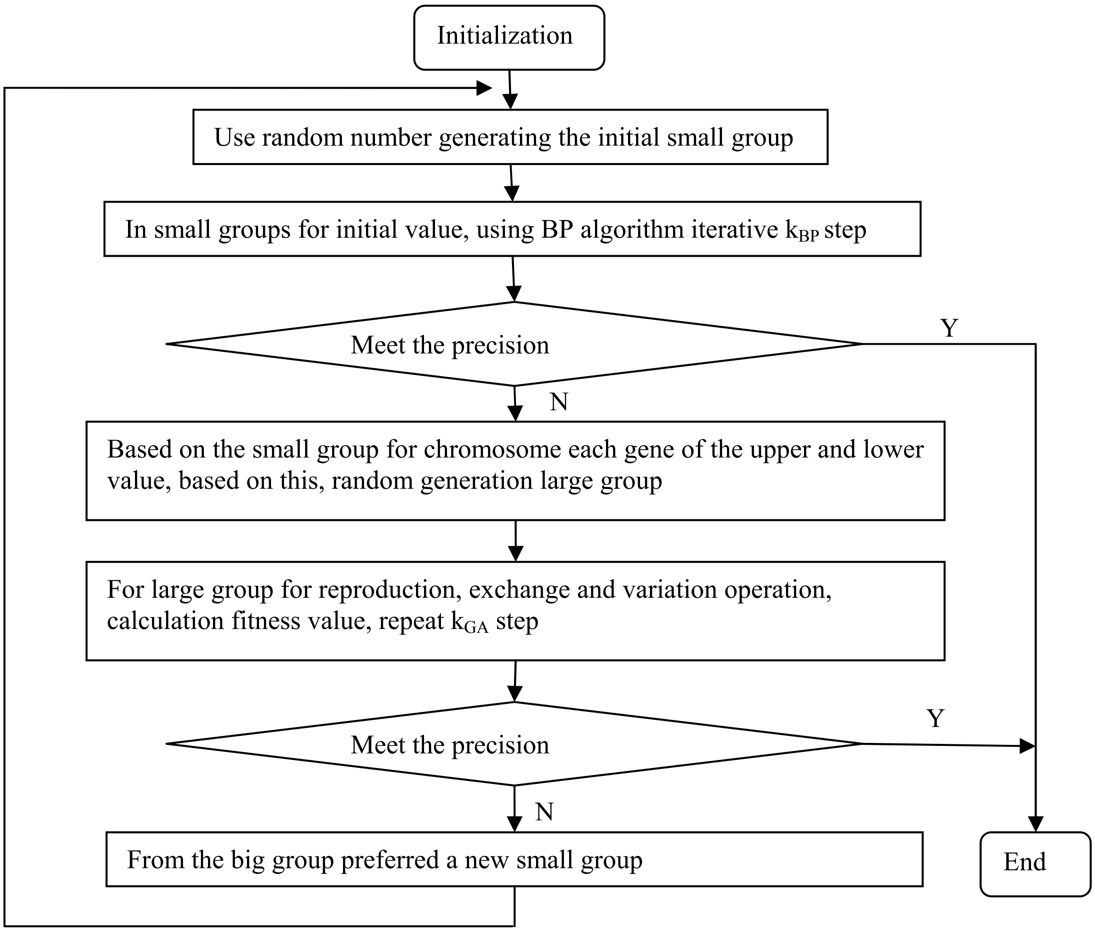

Figure 1. The new GA algorithm process

Table 1. contrast of NN with new algorithm and the original BP network prediction results

Serial number	Sandstone Thickness/m	A new hybrid algorithm NN		BP network.	
		Predicted value/m	Relative error/%	Predicted value/m	Relative error/%
1	6.8	6. 832 236	- 0. 47	6. 710 241	1. 32
2	8.7	8. 691 341	0. 21	8. 716 532	- 0. 19
3	4.6	4. 611 938	- 0. 26	4. 576 540	0. 51
4	2.0	2. 029 382	- 1. 46	1. 009 203	- 0. 46
5	6.5	6. 489 639	0. 16	6. 352 378	2. 27
6	8.3	8. 321 090	- 0. 25	8. 260 890	0. 47
7	3.1	3. 087 496	0. 40	3. 116 778	- 0. 54
8*	5.9	6. 343 739	- 7. 52	5. 286 925	10. 39
9*	7.8	7. 159 621	8. 21	9. 080 346	14. 94
10*	9.9	9. 393 122	5. 12	9. 126 592	7. 80

Note: * is generalization test well.

Conclusions

According to GA to improve the poor local search capability shortcomings, the GA and BP algorithm combined with, this paper presents a new hybrid algorithm is proposed. The application markovain chain theory proves that the new algorithm is globally convergent. Use with the new algorithm and BP network trained NN by modeling and forecasting reservoir system, the results show that the hybrid algorithm has higher accuracy, thus has certain application prospect.

References

[1] M. R. Azimi-Sadjadi, R. J. Lion. Fast Learning Process of Multilayer Neural Networks Using Resursive Least Squares Method. IEEE Trans.On SP, Vol.40(2010), p.446~450.

[2] Kennedy, J. and Eberhart. R. C. Particle swarm optimization. In Proc. IEEE Iternational Conference on Neural Networks, IEEE Service Center, pisataway, NJ, 2010, p. 39~43.

[3] Duta S, Shekhar S. Bond rating: a non-conservative application of neural networks. Proc IEEE Int Conf on Neural Networks[C].San Diego: IEEE, 2011, p. 443~450.

[4] Ge Hong, Mao Zong-yuan. The Analysis of t he Local Search Efficiency of Genetic Neural Networks and t he Improvement of Algorithm [A]. Processing of t he 4t h World Congress on Intelligent Cont rol and Automation[C] . Hefei : Press of East China University of Science and Technolgy ,2002.

[5] Fogel D B. Asymptotic Convergence Properties of Genetic Algorithms and Evolutionary Programming : Analysis and Experiments[J] .Cybernetics and System ,1994 ,25 (6) :389 - 407.

[6] Gambardella L M , Dorigo M. Solving symmetric and asymmetric TSPs by ant colonies[A]. In Proceedings of t he IEEE international conference on evolutionary computation (ICEC ’96) [C]. USA : IEEE Press , 1996. 622 - 627.

[7]Yanmei Li, Jingmin Wang. Risk Assessment Model of Smart Grid Project Based on Variable Pr ecision RS and LSSVM. Advances in Information Sciences and Service Sciences (AISS). Vol.3, No .10, 2011,11 :375-383.