Research Journal of Applied Sciences, Engineering and Technology 6(10): 1841-1845, 2013
DOI:10.19026/rjaset.6.3912
ISSN: 2040-7459; e-ISSN: 2040-7467

Submitted: November 19, 2012 Accepted: January 14, 2013 Published: July 20, 2013

Research Article

Improved Estimation of Distribution Algorithms Based on Gaussian Distribution

[1]Shang Gao, Ling Qiu and [3]Cungen Cao
[1]School of Computer Science and Technology, Jiangsu University of Science and Technology, Zhenjiang 212003, China
[2]Artificial Intelligence of Key Laboratory of Sichuan Province, Sichuan University of Science and Engineering, Zigong 643000, China
[3]Key Laboratory of Intelligent Information Processing, Institute of Computing Technology, Chinese Academy of Sciences, Beijing 100080, China

Abstract: Estimation of Distribution Algorithms (EDAs) is a new kind of evolution algorithm. In EDAs, through the statistics of the information of selected individuals in current group, the probability of the individual distribution in next generation is given and the next generation of group is formed by random sampling. An improved estimation of distribution algorithms based on normal distribution is presented for function optimization in continuous space. The algorithm regarded the selected individual as a normal distribution and the random new populations of normal distribution were generated and some selection of individual are crossed with the best solution. Compared with estimation of distribution algorithms based on uniform distribution and estimation distribution algorithm based on normal distribution, the improved estimation distribution algorithm based on normal distribution is more effective through result. At last, better population selection proportions are analyzed.

Keywords: Continuous space optimization, estimation distribution algorithm, normal distribution, uniform distribution

INTRODUCTION

Since Estimation of Distribution Algorithms (EDAs) were proposed by Baluja in 1994 (Shumeet, 1994), EDAs quickly become an important branch of evolutionary algorithms because they have better mathematical foundation than other evolutionary algorithms. Estimation of Distribution Algorithms (EDAs), sometimes called Probabilistic Model-Building Genetic Algorithms (PMBGAs), are stochastic optimization methods that guide the search for the optimum by building and sampling explicit probabilistic models of promising candidate solutions. On the basis of statistical learning theory, EDAs use some individuals selected from the population at the current evolutionary generation to build a probability model and then produces offspring for the next generation by sampling the probability model in a probabilistic way. A lot of investigations in Guolin (2012), Muhliebe and Paass (1996), Paul and Iba (2002), Pelikan *et al*. (2000), Rong and Yuquan (2012), Shumeet (1994), Zhang (2011) and Zhou and Zengqi (2007) show that EDAs have good optimization performance in both combinatorial problems and numeric optimization problems. Until now there are many studies about EDAs, but EDAs mainly consist of several types: Population Based Incremental Learning (PBIL) (Shumeet, 1994), Univariate Marginal Distribution Algorithm (UMDA), Compact Genetic Algorithm (CGA), mutual-information-Maximizing Input Clustering Algorithm (MIMIC), bivariate marginal distribution algorithm (BMDA), factorized distribution algorithm (FDA), Bayesian Optimization Algorithm (BOA), Extended Compact Genetic Algorithm (ECGA) and Estimation Of Bayesian Network Algorithm (EBNA). UMDA works well only in the solution of linear problems with independent variables, so it requires extension as well as application of local heuristics for combinatorial optimizations. PBIL uses vector probabilities instead of population and has good performance for solving problems with independent variables in binary search space. CGA independently deals with each variable and needs less memory than simple genetic algorithm. MIMIC searches the best permutation of the variables at each generation to find the probability distribution through using Kullback-Leibler distance. BMDA is mainly based on the construction of a dependency graph, which is acyclic but does not necessarily have to be a connected graph. FDA integrates evolutionary

Corresponding Author: Shang Gao, School of Computer Science and Technology, Jiangsu University of Science and Technology, Zhenjiang 212003, China, Tel.: +86-511-84408968

algorithms with simulated annealing. This method requires additively decomposed function and the factorization of the joint probability distribution remains same for all iterations. BOA applies Bayesian network and Bayesian Dirichlet metric to estimate joint probability distributions, thus, it can take advantage of the prior information about the problem. ECGA factorizes the joint probability distribution as a product of marginal distributions of variable size. EBNA employs Bayesian network for the factorization of the joint probability distribution and BIC score. In this study, an improved estimation of distribution algorithms based on normal distribution is presented for function optimization in continuous space. The algorithm regarded the selected individual as a normal distribution and the random new populations of normal distribution were generated and some selection of individual are crossed with the best solution.

BASIC ESTIMATION OF DISTRIBUTION ALGORITHMS

Estimation of Distribution Algorithms (EDAs), sometimes called Probabilistic Model-Building Genetic Algorithms (PMBGAs), are stochastic optimization methods that guide the search for the optimum by building and sampling explicit probabilistic models of promising candidate solutions. Optimization is viewed as a series of incremental updates of a probabilistic model, starting with the model encoding the uniform distribution over admissible solutions and ending with the model that generates only the global optima.

EDAs belong to the class of evolutionary algorithms. The main difference between EDAs and most conventional evolutionary algorithms is that evolutionary algorithms generate new candidate solutions using an implicit distribution defined by one or more variation operators, whereas EDAs use an explicit probability distribution encoded by a Bayesian network, a multivariate normal distribution, or another model class. In EDAs the new population of individuals is generated without using neither crossover nor mutation operators. Instead, the new individuals are sampled starting from a probability distribution estimated from the database containing only selected individuals from the previous generation. Figure 1 illustrates the flowchart of EDA.

The general procedure of an EDA is outlined in the following:

Step 1: t = 0
Step 2: Initialize model M (0) to represent uniform distribution over admissible solutions
Step 3: While (termination criteria not met)
Step 4: P = generate N>0 candidate solutions by sampling M (t)
Step 5: F = evaluate all candidate solutions in P
Step 6: M (t+1) = adjust-model (P, F, M (t))
Step 7: t = t + 1

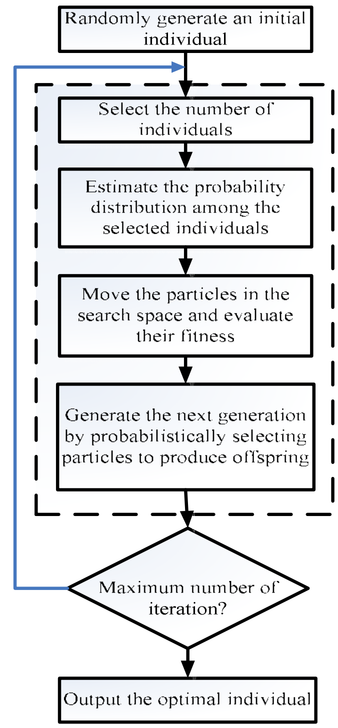

Fig. 1: Illustrates the flowchart of EDA

Using explicit probabilistic models in optimization allowed EDAs to feasibly solve optimization problems that were notoriously difficult for most conventional evolutionary algorithms and traditional optimization techniques, such as problems with high levels of epistasis. Nonetheless, the advantage of EDAs is also that these algorithms provide an optimization practitioner with a series of probabilistic models that reveal a lot of information about the problem being solved. This information can in turn be used to design problem-specific neighborhood operators for local search, to bias future runs of EDAs on a similar problem, or to create an efficient computational model of the problem.

IMPROVED EDA BASED ON GAUSSIAN DISTRIBUTION

In probability theory, the Gaussian (or normal) distribution is a continuous probability distribution, defined on the entire real line that has a bell-shaped probability density function, known as the Gaussian function or informally as the bell curve:

$$f(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}, \quad -\infty < x < \infty \tag{1}$$

The parameter μ is the mean or expectation (location of the peak) and σ^2 is the variance. σ is known as the standard deviation.

The estimator μ is:

$$\hat{\mu} = \bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$$

The estimator σ is:

$$\hat{\sigma} = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2}$$

The Box-Muller transform is one of the earliest exact transformation methods. It produces a pair of Gaussian random numbers from a pair of uniform numbers.

$$\begin{cases} x_1 = u + \sigma(-2\ln u_1)^{1/2}\cos 2\pi\, u_2 \\ x_2 = \mu + \sigma(-2\ln u_1)^{1/2}\sin 2\pi\, u_2 \end{cases} \quad (2)$$

where, u_1 and u_2 are random numbers of uniform distributed in the interval [0, 1].

In the study we are considering the following global optimization problem:

$$\begin{aligned} &\min f(x_1, x_2, \cdots, x_n) \\ &s.t. \quad a_i \le x_i \le b_i (i = 1,2,\cdots,n) \end{aligned} \quad (3)$$

Finding the global minimum (or maximum) of a function is far more difficult: symbolic (analytical) methods are frequently not applicable and the use of numerical solution strategies often leads to very hard challenges.

The framework of solving optimization problem by EDA is as follows:

Step 1: Generate N solutions $x_i(i = 1, 2, \ldots ,n)$ from $[\alpha_i, b_i]$, using the uniform design technique.
Step 2: Assess the fitness of all individuals in the initial population and retain the best solution.
Step 3: Order the population by fitness in descending sorting and choose the optimal m individuals ($m \le N$).
Step 4: Analyze the generated m individuals information and calculate the mean $\hat{u}_\iota$ and variance $\hat{\sigma}_\iota$ of each variable.
Step 5: Sample N new solutions according to formula (2).
Step 6: If the given stopping condition (up to the required number of iterations $n_{\max}$) is not met, go to step 2.

The ideas for Improved EDA are given as follows. We will make full use of the best solution, which is reserved. According to the crossover probability p_c, select N × p_c individuals randomly and cross them with the best individual cross. The cross method is:

$$x^{new} = ax_{\min} + (1-a)x^{old} \quad (4)$$

where, α is [0, 1] random number.

The improved estimation of distribution algorithm for optimization problem is as follows:

Step 1: Generate N solutions x_i (i = 1, 2, … , n) from $[\alpha_i, b_i]$, , using the uniform design technique.
Step 2: Assess the fitness of all individuals in the initial population and retain the best solution.
Step 3: Order the population by fitness in descending sorting and choose the optimal m individuals ($m \le N$).
Step 4: Analyze the generated m individual's information and calculate the mean $\hat{u}_\iota$ and variance $\hat{u}_\iota$ of each variable.
Step 5: Sample N new solutions according to formula (2).
Step 6: According to the crossover probability p_c, select N × p_c individuals randomly and cross them with the best individual cross according to formula (4).
Step 7: If the given stopping condition (up to the required number of iterations $n_{\max}$) is not met, go to step 2.

NUMERICAL EXPERIMENTS AND RESULTS

To illustrate the advantage of his normal distribution, the estimation of distribution algorithms based on uniform distribution is discussed. The EDA based on uniform distribution is as follows:

Step 1: Generate N solutions x_i (i = 1, 2, … , n) from $[a_i, b_i]$, using the uniform design technique.
Step 2: Assess the fitness of all individuals in the initial population and retain the best solution.
Step 3: Order the population by fitness in descending sorting and choose the optimal m individuals ($m \le N$).
Step 4: Analyze the generated m individual's information and calculate the minimum $x_{min}^{(i)}$ and maximum values $x_{min}^{(i)}$ of each variable.
Step 5: Sample N new uniform random solutions from $[x_{min}^{(i)}, x_{max}^{(i)}]$.
Step 6: If the given stopping condition (up to the required number of iterations $n_{\max}$) is not met, go to step 2.

The following well-known functions are used in our experimental studies:

$$\min F_1 = \sum_{i=1}^{30} x_i^2 \qquad -1 \le x_i \le 1 (i = 1,2,\cdots 30)$$

$$\min F_2 = \sum_{i=1}^{30}|x_i| + \prod_{i=1}^{30}|x_i| \qquad -1 \le x_i \le 1 (i = 1,2,\cdots 30)$$

$$\min F_3 = \max_{1 \le i \le 30}|x_i| \quad -1 \le x_i \le 1 (i = 1,2,\cdots 30)$$

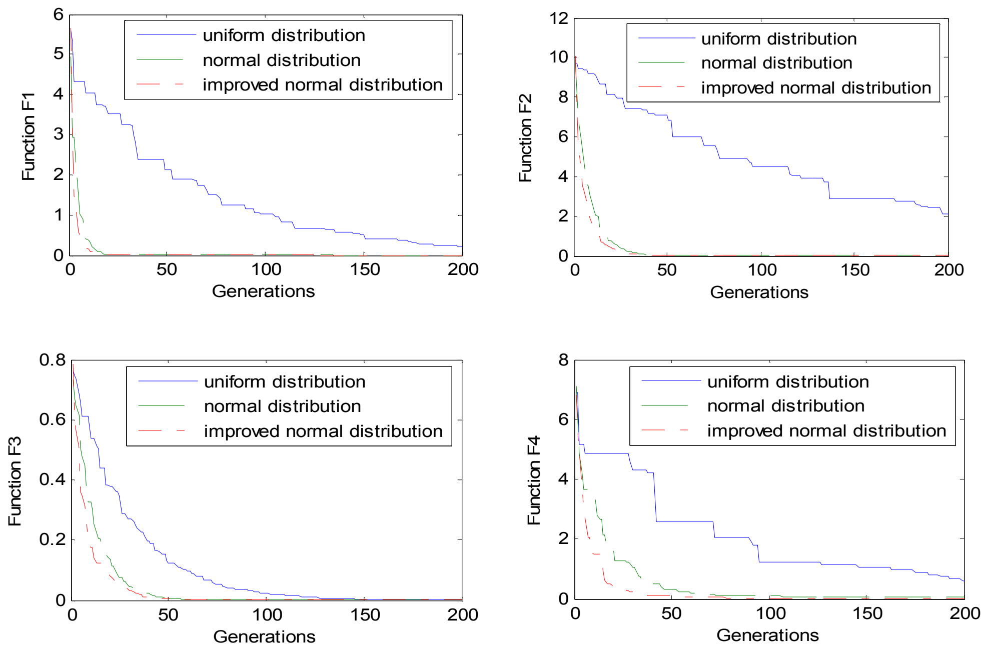

Fig. 2: Comparisons of three algorithms

Table 1: Comparison results of m/N

Algorithm	m/N	Maximum number of iterations	Minimum number of iterations	Average number of iterations
EDA	10%	-	88	-
	20%	128	108	117
	30%	145	132	137.4
	40%	173	154	163.
	50%	215	193	199.1
Improved EDA	10%	-	83	-
	20%	115	107	112
	30%	145	129	133.8
	40%	172	152	162.3
	50%	202	185	195.1

$$\min F_4 = \sum_{i=1}^{30}\left(\sum_{j=1}^{i} x_j\right)^2 \qquad -1 \le x_i \le 1 (i = 1,2,\cdots 30)$$

We set the parameter values as follows: N = 1000, m = 0.4*N, p_c = 0.01 and nmax = 200. From Fig. 2, the convergence speed of improved EDA based on normal distribution is much faster than the EDA based on normal distribution and EDA based on uniform distribution. The improved EDA based on normal distribution take full advantage of the mean and variance of the data information and retain the best solution, so it is fastest of the three algorithms.

The test function is:

$$\min F_5 = \sum_{i=1}^{30}[x_i^2 - 10\cos(2\pi x_i) + 10]$$
$$s.t. \quad -5.12 \le x_i \le 5.12 (i = 1,2,\cdots,30)$$

The parameters which affect performance of EDA are the population N and m. When N = 1000, the results are not the same as different the ratio of m/N. We can act the number of iterations as a comparison parameter, which algorithm reached rand around the best 0.000001. With different proportions, the algorithm is tested 10 times. The results are shown in Table 1. When m/N = 10%, it may not converge.

From Table 1, we can see that the laws of the two algorithms are similar.

If the ratio of m/N is the greater, the effect is the worse. Of course, the ratio m/N is too small; it is easy to fall into local minima. So the ratio of m/N is 20% - 40%, the results were quite good.

CONCLUSION

EDAs belong to the class of evolutionary algorithms. The main difference between EDAs and most conventional evolutionary algorithms is that evolutionary algorithms generate new candidate solutions using an implicit distribution defined by one or more variation operators, whereas EDAs use an explicit probability distribution encoded by a Bayesian network, a multivariate normal distribution, or another model class. This study presents an improved EDA based on normal distribution e for solving continuous space optimization problems. The algorithm has strong ability to search optimization problems. Especially it is suitable for complex function optimization and high dimensional optimization problems. The EDA based on normal distribution can be improved further, such as with the mutation operation, or hybrid other intelligent algorithms. Our further work will focus on the application of EDA to real-world problems.

ACKNOWLEDGMENT

This study was supported by the Open Project Program of Key Laboratory of Intelligent Computing and Information Processing (Xiangtan University), Ministry of Education (No. 2011ICIP05), Artificial Intelligence of Key Laboratory of Sichuan Province, Jiangsu 333 Project, Qing Lan Project. And the National Natural Science Foundation of China under Grant 51008143.

REFERENCES

Guolin, Y., 2012. Multi-objective estimation of estimation of distribution algorithm based on the simulated binary crossover. J. Converg. Inform. Technol., 7(3): 110-116.

Muhliebe, H. and G. Paass, 1996. From Recombination of Genes to the Estimation of Distributions I Binary Parameters: Lecture Notes in Computer Science. Springer Verlag, Berlin, Germany, 1141: 178-187.

Paul, T.K. and H. Iba, 2002. Linear and combinatorial optimizations by estimation of distribution algorithms. Proceeding of the 9th MPS Symposium on Evolutionary Computation. IPSJ, Japan.

Pelikan, M., D.E. Godberg and E.C. Paz, 2000. Linkage problem, distribution estimation and Bayesian networks. Evol. Comput., 8(3): 311-340.

Rong, H. and L. Yuquan, 2012. A novel estimation of distribution algorithm with multiple probability models. Adv. Inform. Sci. Serv. Sci., 4(17): 308-315.

Shumeet, B., 1994. Population based incremental learning: A method for integrating genetic search based function optimization and competitive learning. Technical Report, No. CMU-CS-94-163, Carnegie Mellon University, Pittsburgh, Pennsylvania.

Zhang, R., 2011. A rule-based estimation of distribution algorithm for solving job shop. J. Converg. Inform. Technol., 6(8): 220-227.

Zhou, S. and S. Zengqi, 2007. A survey on estimation of distribution algorithm. Acta Automat. Ainica, 33(2): 113-124. (In Chinese)