An Information System Integration Method Based on Controllable Genetic Algorithm

Geying Liang and Zongjian Tang

College of Mathematics and Information Science, Guangxi University, Nanning 530004, P.R. China lgy2680@163.com tang2680@tom.com

Abstract. The essential for an enterprise to construct an information system is to integrate the information system with the legacy system. We introduce a biological genetic controllable method into the process of information system integration, and apply self-adaptive genetic algorithm to simulate the process of information system integration. By controlling the original population and the transformation of gene, we can find an optimal integrating effect and an information system integration method.

Keywords: *Information system integration, Genetic algorithm, Genetic controllable, Self-adaptive*

1. INTRODUCTION

When a modern enterprise is to construct an information system, its essential is to integrate the information system with the original system of the enterprise. The objective is to introduce information technology and information system to the enterprise operation process, to raise its efficiency and effectiveness, gain the competition advantage. Because there are many factors influencing the information system integration process, and there are many uncertain factors in the process, the enterprise often can not predict the integration effect. Specially, the enterprise can't know what information system match with it in selection model of information system. Therefore, this paper applies the biological genetic controllable method, from the point of information system integrating with the original system of an enterprise, searches a way to integrate information system.

2. DESCRIPTION OF THE OBJECTIVE OF INFORMATION SYSTEM INTEGRATION

To integrate information system is a complex social-technology system project. In the process of implementing information system integration, people will consider many factors. (1)The objectives of organization, organization structure, business process, human resource, information technology. (2)Information system functions,

performance, structure, and management ideas contain in the information system. (3) How to fully utilize various resources existing in the organization? (4) How to coordinate resources, so that the organization can raise the efficiency and benefit? Except these, the enterprise hopes the cost of information system integration is the lowest. So the objectives of information system integration can be expressed as below.

Suppose in the process of information system integration, the original state of the system is $X = (x_1, x_2, \cdots, x_m)$, m is the number of factors that people will consider in the process of information system integration. $x_i (i = 1, 2, \cdots m)$ is the state of the i th factor in the original system, the range of value is the integer in [0,7], its value expresses the degree of the factor realizing the expected system objective, 0 expresses not realizing the expected objective of the system, 7 expresses realizing the expected objective of the system completely. Different factors have different effect degrees to the information system integration, we can adopt expert evaluation method to give the value of weight $a_i (i = 1, 2, \cdots, m)$, a_i expresses the factor x_i effect degree to original system status, $a_i > 0$ and $a_1 + a_2 + \cdots + a_m = 1$.

Suppose the state of realized information system, for example the commercial information system, can be expressed by $Y = (y_1, y_2, \cdots, y_m)$, $y_i (i = 1, 2, \cdots m)$ coincides with x_i factor of original system, the range of value is the integer in [0,7], its value expresses the degree of the factor realizing the expected system objective, 0 expresses not realizing the expected objective of the system, 7 expresses realizing the expected objective of the system completely. Different factors have different effect degrees to the information system integration, we can adopt expert evaluation method to give the value of weight $b_i (i = 1, 2, \cdots, m)$, b_i expresses the factor y_i effect degree to information system state, $b_i > 0$ and $b_1 + b_2 + \cdots + b_m = 1$.

The process of information system integration is to coordinate information system with original system, and produce a new system $Z = (z_1, z_2, \cdots, z_m)$, $z_i = (a_i x_i + b_i y_i)/2$. The effect degree to the new system of z_i is $\theta_i (i = 1, 2, \cdots m)$, and we use expert evaluation method to set the value of θ_i, $\theta_i > 0$ and $\theta_1 + \theta_2 + \cdots + \theta_m = 1$. In the case of not considering interaction of factors, the benefit objective of the information system integration in the organization is

$$MaxZ, Z = \sum_{i=1}^{m} \theta_i (a_i x_i + b_i y_i) \qquad (1)$$

In order to realize the objective of informization, it needs to reconstruct the original system, and adjust the status of information system in the process of an organization implementing information system integration. To get z_i, the cost coefficient is $c_i (i = 1, 2, \cdots m)$, $c_i \geq 0$ and $c_1 + c_2 + \cdots + c_m = 1$. Then, the cost objective of the information system integration in the enterprise is:

$$MinC, C = \sum_{i=1}^{m} c_i z_i \qquad (2)$$

This is a multi-objective problem [2]. Considering that different organization have different preference to benefit and cost, set benefit preference is $\alpha(\alpha > 0)$, cost

preference is $\beta(\beta > 0)$, and $\alpha + \beta = 1$, synthesize (1) and (2) to construct an object function,

$$MaxQ, Q = \alpha Z - \beta C \qquad\qquad (3)$$

3. THE PROCESS OF USING GENETIC ALGORITHM TO IMPLEMENTING INFORMATION SYSTEM INTEGRATION

In the process of information system integration, original system and information system will adjust continually in the process of informization, matching with each other. According to this characteristic, we adopt self-adaptive genetic algorithm to solve the problems in the process of evolution. We don't use random approach to define the original population, apply expert assess approach to produce the population of original system and the population of information system, and apply XOR algorithm to make a match randomly between the two populations and get the prime population at the beginning of genetic algorithm. This can acquire the prime population corresponding with actuality, and enhance the efficiency of evolution.

3.1 The Code of Problem

This problem will adopt the method of binary code. According to the basic principal of genetic algorithm, every factor effecting on the process of information system can be expresses as a gene. The sequence of gene is decided by the project preference sequence, a_i and b_i. Every gene can be represented by 3 bits, and can get the any integer value between [0, 7]. Therefore, $X = x_1 x_{2 \cdots} x_{m-1} x_m$, $Y = y_1 y_2 \cdots y_{m-1} y_m$, the length of gene-string is $l = 3m$.

3.2 The Design of Fitness

The formula (3) is a problem of evaluating maximum, so design a fitness function as below:

$$F = \alpha Z - \beta C + \lambda \qquad\qquad (4)$$

In the formula (4), λ is a proper small number, when $\alpha Z - \beta C + \lambda \leq 0$, $F = 0$.

3.3 The Design of Genetic Operator

3.3.1 Selection Operators

Suppose the size of population is N, sort on the adaptive values descending, adopt excellent individual proportion-conservation strategy, that is, set a proper ratio, such as $d = 0.1$, let the front dN individuals as the next generation individuals, choose the left individuals to next generation according to roulette wheel method, until the sum of individuals is N.

3.3.2 Crossover Operators and Mutation Operators

In the self-adaptive genetic algorithm [3], the probability of individual crossover and the probability of mutation are not constant, they are adjusted by the fitness of individual, crossover operator and mutation operator are decided by whether the individual fitness is larger than average fitness. If the larger fitness f is less than the average fitness f_{avg} between two individuals which execute crossover operation, the probability of crossover between the two individual are equal and are constant values. If the larger fitness f is larger than the average fitness f_{avg} between two individuals which execute crossover operation, adjust the probability of crossover as below: the probability of crossover is the function of the max fitness f_{max} and the min fitness f_{min}. And the self-adaptive crossover probability P_c can be represented as below:

$$P_c = \begin{cases} P_{c0}, & f \leq f_{avg} \\ P_{c1} + \dfrac{f_{max} - f}{f_{max} - f_{min}}(P_{c0} - P_{c1}), & f > f_{avg} \end{cases} \tag{5}$$

In formula (5), P_{c0} is a larger probability of crossover, P_{c1} is a less probability of crossover, and $P_{c0} > P_{c1}$. That is, the individual that has a less fitness will adopt a larger probability of crossover, so that it can evolve more rapidly. But the individual that has a larger fitness will adopt a less probability of crossover, so that it can keep on its excellence characteristics.

The self-adaptive probability of mutation P_m can be represented as below:

$$P_m = \begin{cases} P_{m0}, & f' \leq f_{avg} \\ P_{m1} + \dfrac{f_{max} - f'}{f_{max} - f_{min}}(P_{m0} - P_{m1}), & f' > f_{avg} \end{cases} \tag{6}$$

In formula (6), f' is the individual fitness which execute the mutation operation, P_{m0} is a larger probability of mutation, P_{m1} is a less probability of mutation, and $P_{m0} > P_{m1}$. That is, the individual that has less fitness will adopt the larger

probability of mutation, and the individual that has larger fitness will adopt the less probability of mutation, that can ensure multiformity of population, and overcome the premature convergence of the algorithm.

4. THE DESIGN OF GENETIC ALGORITHM

Step 1: N experts will evaluate X and Y respectively, and get 2N individuals. Match individual among X with individual among Y according to the random principal, the two matching individual will adopt simple XOR operation at the same position to get the new value. For example, X=101101010011, Y=100110011010, than the value of new individual is 001011001001. Through this approach, get N individual as original population. Let the iterative number of generation $t = 1$, and the crossover point $k = 1$.

Step 2: If the iterative number of generation t is larger than the max iterative number of generation, go to step 5, otherwise go to step 3. If $k > m$, then $k = 1$.

Step 3: Calculate the fitness of each individual and the average fitness of all individual according to formula (4), and sort individuals on the value of fitness, select dN individuals in front of the population as next generation individual. Use roulette wheel approach to select two individuals among the left individuals to match crossover according to P_c, that is execute crossover operation on gene chromosome in the k th group. Then execute mutation operation to the new individual according to P_m.

Step 4: $t = t + 1$, $k = k + 1$, go to step 2.

Step 5: Take the optimal individual as the optimal solution, the algorithm is over.

5. CONCLUSIONS

Because there are many factors influencing the information system integration process, and there are many uncertain factors in the process, the enterprise often can not predict the integration effect. This paper put forward a controllable self-adaptive genetic algorithm. By controlling the original population and cross point in the genetic process, it simulates the process of information system integration, and attain to the optimal effect of information system integration.

REFERENCES

1. M. Zhou and S. Sun, *Genetic Algorithms: Theory and Applications* (National Defense Industry Press: Beijing, Beijing, 2005).
2. C.A. Coello, An Updated Survey of GA-based Multi-objective Optimization Techniques, *ACM Computing Surveys.* Volume 32, Number 2, pp.109-143, (2000).

3.   L. Tong, Z. Chen, Z. Yuan, and L. An, A Double Population Self-adaptive Genetic
      Algorithm for Partner Selection of Virtual Enterprise, *Computer Engineering*. Volume 32,
      Number 8, pp.192-194, (2006).
4.   G. Chen and X. Kuan, The Information System Security Technology Scheme Based on
      Genetic Algorithm, *Journal of National University of Defense Technology*. Volume 27,
      Number 6, pp.130-134, (2005).