Application of Genetic Algorithm to Spatial Distribution in Urban Planning

Hu Xin
School of Civil Engineering
Xi'an Univ. of Arch. & Tech.
Xi'an, China
huxin197@126.com

Zhang Zhi-xia
School of Management
Xi'an Univ. of Arch. & Tech.
Xi'an, China
zzx1024@126.com

***Abstract*—Urban planning is a important part of urban construction, spatial distribution is subordinate to urban planning, it is difficult to obtain a satisfactory solution by traditional optimal methods. In view of the factors that influence urban spatial distribution are diverse and dynamic and city needs to keep the balance continuously in the development, genetic algorithm method is selected because it just reflects the dynamic idea of spatial planning. Based on the classical optimization method, a genetic search optimization algorithm in continuous space is presented. The selection of the key parameters is discussed. The calculation results of computational examples with the algorithm proposed are given an application example is analyzed to illustrate the use of the model and to demonstrate its capabilities in determining the optimizing scheme of spatial distribution in urban planning. The algorithm is shown to be steadily convergent at global optimum.**

Keywords-*spatial distribution; genetic algorithm;* ***target*** *determination; mutation*

I. INTRODUCTION

Urban planning is the general deployment to the whole social economics development in a certain region, which arranges concretely regional economic development and urban layout, considers comprehensively every aspect of natural resources and social economic base, therefore, it need to analysis many project and the connection between different work. Aimed at these kind of problem, common method can divide the project work into the proper size mission, describe the quality being dependent on between the mission, but it is essential to depict the schedule with Gantt chart or Network chart (PERT/CPM) in actual project work[1]. The size missions are often geographical space, humanistic environment, commercial trade, etc. This paper mainly studies spatial distribution among geographical space, because urban spatial distribution is subordinate to spatial planning, technical study document on it is comparative little. In the course of studying, it usually adopts principal component analysis to analysis quantitatively space on the selected axis, then determines the level and number of developing axis and space[2]. The quantitative analysis method is difficult to obtain a satisfactory solution without mathematic support. In view of spatial distribution of urban planning often includes many non-spatial, dynamic, nonlinear data, so the traditional mathematic methods are not suitable to solve this problem. Genetic algorithm (GA) is proposed to make up the above shortage. Compared to the traditional methods, the method is more efficient and more understandable.

GA is based, in concept, on natural genetic and evolutionary mechanisms working on populations of solutions in contrast to other search techniques that work on a single solution[3]. Searching, not on the real parameter solution space but on a bit string encoding of it, they mimic natural chromosome genetics by applying genetics like operators in search of the global optimum. An important aspect of GA is that, although they do not require any prior knowledge or any space limitations, such as smoothness, convexity or un-modality of the function to be optimized, they exhibit very good performance in the majority of applications[4]. They only require an evaluation function to assign a quality value (fitness value) to every solution produced. Another interesting feature is that they are inherently parallel (solutions are individuals unrelated with each other), and therefore, their implementation on parallel machines reduces significantly the CPU time required.

Compared with other optimization methods, GA is suitable for traversing large search spaces since they can do this relatively rapidly and because the use of mutation diverts the method away from local minima, which will tend to become more common as the search space increases in size. GA gives an excellent trade off between solution quality and computing time and flexibility for taking into account specific constraints in real situations.

This paper puts forward using the optimal algorithm GA to solve two dimensional packing, this can converge rapidly from any initial solution, and obtain an optimal layout scheme.

II. TARGET SYSTEM EVALUATION OF SPATIAL DISTRIBUTION

An urban planning can be evaluated from a single aspect such as industry development, environmental quality, and so on. Such a single-aspect simply doesn't reflect the overall urban development. It will be even less comprehensive if a combined index is used. In order to provide an effective target system evaluation on the overall urban development, it is necessary to establish a systematic, comprehensive index

978-1-4244-3531-9/08/$25.00©2008 IEEE

system[5].The objectives for developing the comprehensive index system are as follows:

• The index system must be able to reflect every aspect of the spatial distribution.

• The data for the indexes must be able to be collected from the reliable sources and be consistent.

• The index system must be able to accommodate the relationship between the evaluation indexes and the evaluation criteria.

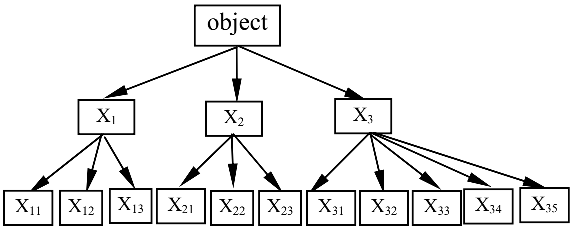

Figure 1. Target evaluation system

According to these objectives, this index system considers the overall urban planning determined by three major indexes[6]: the combined index for social development(X_1), the combined index for economic development(X_2), and the combined index for environmental protection (X_3), optimization objective is cost to be least, then

$$y = \min f(X_1, X_2, X_3) \quad (1)$$

X_1 is further determined by the urban population status (X_{11}), the quality of urban life (X_{12}), and the urban administration (X_{13}). X_{11} is measured by the natural growth rate of population, population density, the enrollment of colleges and universities, and the number of scientists or engineers per ten thousand employees. X_{12} is measured by the per capita income of urban residents, the average salary of urban employees, the per capita annual saving of urban residents, the per capita living space of urban residents, the per capita water consumption of urban residents, X_{13} is measured by the number of traffic accidents per hundred thousand urban residents, and the number of fires per hundred thousand urban residents.

X_2 is further determined by the regional economic activities (X_{21}), the regional combined economic benefits (X_{22}) and the regional international investment (X_{23}). X_{21} can be measured by the regional gross domestic production, the regional per capital income, the regional gross industry and agriculture output, the regional total fixed investment , the regional total retail value of commodities, the regional government revenue, the income generated by tourist industry, and the regional gross export. X_{22} can be measured by the per capital national income, the per capital domestic production, and the per capital industry and agriculture output. X_{23} can be measured by the number of the new business ventures involving international investment, the amount of international capital on the ventures, and the amount of international capital actually used.

X_3 is further measured by the coverage of trees and flowers within urban areas (X_{31}), the per capita coverage of trees and flowers within urban areas (X_{32}), the processing rate of industrial water disposal (X_{33}), the processing rate of industrial gas disposal (X_{34}), and the processing rate of industrial solid disposal (X_{35}).

III. GA APPLIED TO SPATIAL DISTRIBUTION

The GA was first introduced by Holland (1975). It is a global heuristic searching technique whose mechanism is based on the simplifications of evolutionary processes observed in nature. It is an iterative procedure which maintains a population of chromosomes representing different possible solutions to a problem. Each single iteration is called a generation. In each generation, the fitness of each chromosome is evaluated, which is decided by the fitness function, and some chromosomes are selected as the parental chromosomes. Based on the parental chromosomes, new chromosomes, called offspring (also called child chromosomes), are reproduced by three genetic operators, select, crossover and mutation. The offspring are supposed to inherit the excellent genes from their parents, so that the average quality of solutions is better than that in the previous generations. This evolution process is repeated until some termination criterion is met[7]. The following sub-sections describe in detail how the GA is developed to solve the spatial distribution of urban planning.

A. *Initialization*

The first step in constructing the GA is to define an appropriate genetic representation (coding). A good representation is crucial because it significantly affects all the subsequent steps of the GA. In this research, with the limit of urban planning classification standards, code can set simply to industrial area, residential area and non-constructive land, it also increases other class according to actual situation, and then the chromosome gene is determined. Therefore, the length of chromosome n is also determined, but the number of genes in the sub-chromosome, varies. If one sub-chromosome comprises multiple genes, it indicates that the corresponding assembly line performs multiple processes according to the gene sequence in the sub-chromosome.

Based on the length of chromosome n, initial population size m is easy to acquire. in general case, the value of m is 50～100.The gene of chromosome is obtained randomly. According to the relevant standards, Types and code length of land use is determined in Table 1.

TABLE I. TYPES AND CODE LENGTH OF LAND USE

code length	Types of land use	Percentage	Remark
4	non-constructive land	c%	Not belong to urban land
8	residential land	p%	According to national standards
8	industrial land	q%	According to national standards
8	public service land	d%	According to national standards
…	…	…	…

Each solution chromosome is made of k genes, where k is the number of land types. $gene_1$ is used to determine the spatial arrangement of non-constructive land, $gene_k$ represents of arrangement of k land.

$$Chromosome = (gene_1, gene_2, \cdots, gene_k) \tag{2}$$

B. Fitness

The main purpose of this phase is to evaluate the time and cost for each possible solution (chromosome) in generation in order to determine the fitness of the solution. This fitness determines the likelihood of survival and reproduction of each solution in following generations. The greater the fitness of a chromosome is, the greater the probability to survive. As such, this phase evaluates the identified fitness functions for each solution using the following function:

$$f(x) = c_1 g(x_1) + c_2 g(x_2) + c_3(x_3) \tag{3}$$

Where, $f(x)$ is the profit of spatial distribution, c_1, c_2, c_3 are separately proportion of the social, profit relative to whole profit, $g(x_1)$, $g(x_2)$, $g(x_3)$ are respectively profit function expressed by social, economic, environmental protection indexes.

C. Selection

The selection in the GA is the process of selecting chromosomes for the next generation in terms of their fitness. Many selection schemes have been reported. The preferred selection is commonly utilized because it is simple to implement and provides good solution [8]. In this research, this scheme is used and its procedure can be described as follows:

1) Accumulate fitness of every chromosome in order, generate fitness accumulated value S_n:

$$S_n = \sum_i f_i \tag{4}$$

2) Determine Probability of chromosome selected p_i by means of formula:

$$p_i = f_i / S_n \tag{5}$$

3) Accumulate p_i and generate accumulated Probability g_i:

$$g_i = \sum_{j=1}^{i} p_j \tag{6}$$

4) Generate random number r in uniform Distribution at [0，1].

5) Compare r with g_i, if $g_{i-1} < r < g_i$, then select chromosome i as child chromosome.

6) Operate repeated 4) and 5) until the numbers of child chromosome equal to current number.

D. Crossover

The crossover operation is a random process with a probability of crossover, which breeds a pair of child chromosomes from a pair of parental chromosomes. The typical probability of the crossover operator is between 0.2 and 0.6. In this research, two individuals are randomly chosen to act as parents. One of the parents is chosen amongst the best individuals in the population, while the other is randomly chosen from the whole current population at the crossover probability 0.5. An example of a crossover outcome is given in Table 2.

TABLE II. RESULT OF CROSSOVER

Before crossover	Parent1：11011100010101 ¦ 00001100 Parent2：11000100010101 ¦ 00001110
After crossover	Chilid1：11011100010101 ¦ 00001110 Chilid2：11000100010101 ¦ 00001100

E. Crossover

The mutation operation is critical to the success of the GA since it diversifies the search directions and avoids convergence to local optima. It is used to transform the chromosome by the means of randomly changing the ones of genes. Only some offspring take part in the mutation operation. The size is determined by the probability of mutation (the typical value is between 0.0015 and 0.03). In this research, the inversion mutation operator [9] is adopted, which is implemented by simply exchanging the genes between two randomly selected genes of a chromosome. Figure 2 shows an example of this mutation operator.

Original chromosome	R11	R12	R13	R14	R15
Mutated chromosome	R11	R12	R15	R14	R13

Figure 2. Example of inversion mutation operator

F. Termination condition

The GA is repeated iterative process and by using a diversity measure to stop the algorithm. During the iteration, it performs fitness calculating, selection, crossover and mutation until meeting termination condition. In this research, assume that the specified maximal number of generations is 200 and the allowed lowest standard deviation value is 0.1, once the standard deviation is less than 0.2 whichever generation the GA is running at, it will be terminated.

IV. CASE STUDY

A project example is analyzed in order to illustrate the use of the present optimization model and demonstrate its capabilities. The example project is spatial distribution of special urban in Xian. The repetitive activities in the spatial distribution have been identified as locate and clear, excavate, test, and backfill. The model was able to significantly reduce the search space by precluding dominated solutions in the successive generations of the genetic algorithm. The project is divided into three sections. Each activity is performed by a single crew progressing from the first to the third section. The precedence relationships among succeeding activities are finish to start with no lag time. Figure 3 summarizes the change of solution used by GA for the example project.

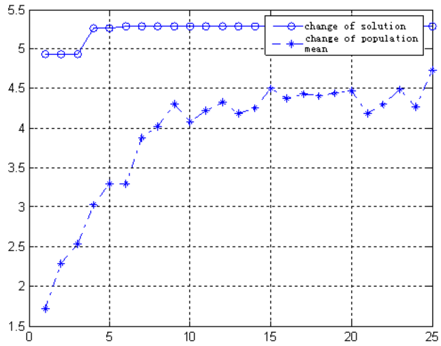

Figure 3. Change of case solution

In this research, the main parameters set as follows:

Length of a string 26

Population size 100

Maximum cycle number 200

Select probability 10%

Crossover probability 65%

Mutation probability 15%

Based on calculation of every chromosome fitness, optimal individual is derived, then analyzing and decoding the chromosome, spatial distribution is determined as show in Figure 4， I, R, N, P represent respectively industrial land, residential land ,non-constructive land and public service land.

Figure 4. Spatial distribution of some area

V. CONCLUSIONS

In this paper, we have presented novel genetic algorithm for spatial distribution in urban planning in which premature convergence hampering conventional GA is avoided. Use simple genetic algorithm to optimize binary sequence and to improve it, we reserve the best individuals and search the individual locally, which in general are far away from one another in configuration space, is a challenge to conventional GA since there is no selection in the direction of a single global minimum. The conventional genetic algorithm has a synchronization problem and the crossover fails to work properly when the diversity of the population is lost, leaving the problem of finding a global minimum to the mutation operator alone. We solve the synchronization problem using a symmetry-adapted crossover in combination with representations which reflect the spatial arrangements of urban, by replacing the offspring with a symmetrically equivalent configuration. The use of symmetry-adapted mating operations allows one to carry out large swaps in configuration space, enabling the land area of good solution to combine to form higher order land area of even better solutions.

REFERENCES

[1] Y. Sadahiro. “Exploratory Method for Analyzing changes in Polygon Distribution”. Environment and Planning B: Planning and Design, Vol. 28, pp. 595-609, 2001.

[2] Z. Yu, Z. Wu, H. Xu. “Urban Spatial Growth Model Based on Genetic Algorithms”.City Planning Review, Vol. 32, No. 5, pp. 83-87, 2008.

[3] D.E. Goldberg. Genetic algorithms in search optimization and machine learning, Addison-Wesley Publishing Company, Inc., Reading, Mass, 1989.

[4] S. Hartmann. “A competitive genetic algorithm for resource-constrained project scheduling”. Naval Research Logistics Vol.45, pp.279–302, 1998.

[5] S. Feng, L.. Xu. “Decision support for fuzzy comprehensive evaluation of urban development”. Fuzzy Sets and Systems Vol.105, pp. 1-12, 1999.

[6] J.M.C. Pereira, L. Duckstein. “A multiple criteria decision-making approach to GIS-based land suitability evaluation”. Internat. J.Geogr. Inform. Systems No.7, pp. 407-424, 1993.

[7] S. Hartmann. “A self-adapting genetic algorithm for project scheduling under resource constraints”. Naval Research Logistics Vol. 49, pp. 433–448, 2002.

[8] U. Kohlmorgen, H. Schmeck and K. Haase, “Experiences with fine-grained parallel genetic algorithms”. Annals of Operations Research Vol. 90, pp. 203–219, 1999.

[9] K. Deb, Multi-objective optimization using evolutionary algorithm, John Wiley & Sons, Ltd., New York, NY, 2001.