3D Neighborhood Relationship of Cellular Genetic Algorithm for Tour Guide Assignment Problem

Lina Setiyani[1] and Takeo Okazaki[2]
[1] Graduate School of Engineering and Science, University of the Ryukyus
1 Senbaru, Nishihara, 903-0213, Japan
[2] Department of Information Engineering, University of the Ryukyus
1 Senbaru, Nishihara, Okinawa 903-0213, Japan
Email: [1]tya.sachi@ms.ie.u-ryukyu.ac.jp, [2]okazaki@ie.u-ryukyu.ac.jp

Abstract: Management optimization is very important in tourism, one of the problems in management is tour guide assignment. The objective of this research is to give a clear understanding into the advantages of increasing cellular dimensionality on Tour Guide Assignment Problem by using Cellular Genetic Algorithm.
Keywords— **3D Neighborhood, Tour Guide Assignment Problem, Cellular Genetic Algorithm.**

1. Introduction

Good service quality with high productivity is the key factor in tourism service center that can attract more tourist and increase profit. For this reason, management optimization becomes a very important thing in tourism. By maintaining good management optimization, a Service Centers (SCs) can provide good quality service while keeping the operational cost as minimize as possible. One of the problem in management is tour guide assignment. In general, assignment problem is a special type of linear programming problem which deals with the allocation of the various objects to the various activities on one to one basis in the most effective way.

Genetic algorithm (GA) and cellular genetic algorithm (cGA) are two methods that have been chosen and proven by researchers for solving many kinds of optimization problems from both classical and real world settings [1]. However, most of these studies focused on implementing the method on two-dimensional (2D) environment. There are not so many studies implementing 3D neighborhood.

Higher cellular dimensions have shown promising benefits. Previous investigation on a 3D architecture has showed improvements in the performance of the algorithm when compared with smaller grid dimensions [2,3]. Increasing the dimension of the neighborhood topology also give more efficient solution in terms of convergence [4].

The purpose of this research is to introduce the 3D neighborhood relationship to cGA in order to reduce computational time while keeping the quality of the solution on Tour Guide Assignment Problem (TGAP).

2. Tour Guide Assignment Problem

The assignment problem is usually recognized by its two components of the assignments, which represent the unrevealed combinatorial structure. Each assignment problem has a table or matrix which associated with it. In general, the row represent the objects or people that we wish to assign, while the column consist of the jobs or tasks that we assign to them. Considering a problem of assignment of n resources to m activities so as to minimize the overall cost or time in a condition where each resource will be assigned to handle one task only for each cycle. The cost matrix (C_{ij}) is represent as below:

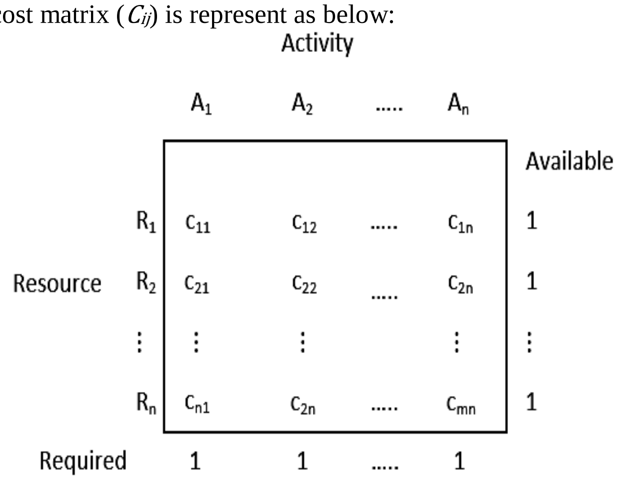

Figure 1. Matrix of Assignment Model

In managing the assignment problem of tour guides, one of the most important issues is about the service time of tour guides. An assignment of tour guides that provides lower total service time is preferred. To reduce the total service time, the waiting time should be lowered as much as possible.

To ensure good quality, a guide should not assign direct too many visitors at a time. However, directing fewer visitors will cost more because of needing more guides.

The service center provides services where $T = \{1, 2, \dots, x\}$ numbers of tour guides will guide $S = \{1, 2, \dots, y\}$ different visitor groups. Each guide has their own fixed guiding time to serve a visitor group. When a visitor group is too large to handle, it will be divided into some smaller subgroups before being served then each subgroup will be handled by only one tour guide (Fig. 2)

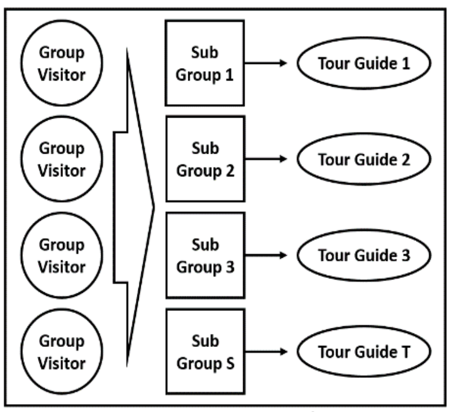

Figure 2. The assigning of tour guide

The fitness function F measures the fitness of the minimum total guiding time. An assignment which has the minimum total guiding time will be the best schedule.

For $i \in T, j \in S$

$$\text{Minimize } F^* = \sum_{i=1}^{x} \sum_{j=1}^{y} \{GijVij + piqi\} \quad (1)$$

$$Vij \in \{0,1\} \; \forall i \in T, j \in S$$

$$qi \in \{0,1\} \; \forall i \in T$$

Where G_{ij} represents actual guiding time of guide and V_{ij} indicate variable that takes value 1 of guide i when assigned to group j and otherwise 0. While p_i represents preparation time of the i^{th} tour and q_i will indicate the value of 1 when the service sequence number of the former subgroup is higher, otherwise 0.

An assignment which has the minimum total guiding time stands for the best schedule assignment of subgroups to tour guides. The lower the total guiding time, the better is the assignment of subgroups to tour guides. The fitness function is composed of three parts:

$$F^* = \sum_{i=1}^{x} \sum_{j=1}^{y} \{G_{ij}V_{ij} + p_i q_i + R_j o_j\} \quad (2)$$

The first part is the actual guiding time. In the second part, an additional preparation time is considered. The final term in *Eq.2* is the product of *{0,1}* variable o_j and the penalty function R_j, which is related to the relative importance of the visitor subgroups. The penalty function value for a more important subgroup is higher than other subgroups. The minimum total service time will lead to an effective and less operational cost.

TGAP itself is a combinatorial optimization problem that belongs to the NP-Hard problem. For this reason, heuristic approach then applied to solve this problem. Since the best solutions spread smoothly through the whole population, genetic diversity preservation then becomes longer. This soft dispersion of the best solutions through the population can be tuned by tuning the selective pressure on the population, and try to find an optimal (or near-optimal) ratio between exploitation and exploration. Basic GA and cGA was previously shown and discussed to solve the tour guide scheduling problem within a short time [5, 6]. In general, the cellular algorithm proposed to follow these steps:

Step 1 Population Initialization
Step 2 Fitness Calculation and Evaluation
Step 3 Parent Selection
Step 4 Crossover
Step 5 Mutation
Step 6 Fitness Calculation and Evaluation (replace if better)

3. 3D Neighborhood

Compare to other genetic algorithm methods, cGA has several advantages, such as a high diversity level which can be maintained for much longer in comparison with centralized ones. Although in crossover process, 3D neighborhood will take longer time than 2D, the mutation process will go faster since mutation only works over single individuals, therefore no communication needed among individuals in this process.

In the crossover process, the offspring were being produced through two-point crossover method. After two points were randomly chosen, there are three fragments available in each chromosome. The offspring is then created by copying the first and third fragment from chromosome 1, while the second fragment from chromosome 2. As illustrated in the figure below, 6, 8 and 3 are copied from chromosome 1. The remained numbers are 9, 2 and 5. They are then copied to create the offspring. The procedure is quite straightforward and it is not necessary for the number of each gene to be unique.

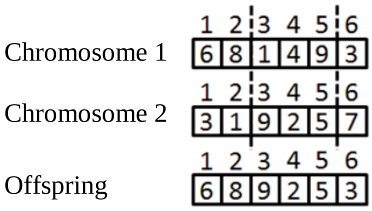

Figure 3. Crossover is done by two-point crossover method.

During the process, the overlap of the neighborhoods provides an implicit mechanism of migration to the cGA. This leads to fast spread of good individuals among population and genetic diversity preservation then becomes longer. As the result, cGA with 3D topology has better convergence time than 2D-cGA. The convergence rate of the algorithm then depended on the shape and size of the grid and neighborhood [7].

The comparison will be learned based on cellular genetic algorithm results when a 2D or 3D grid structure is applied to solve the tour guide assignment problems. The minimum total service time will lead to an effective and less operational cost. In order to additional improvement also made to the waiting time. The waiting time is also effective for the total service time, so it should be lowered.

In the previous work on TGAP, the algorithm was applied to only 2D neighborhood world. Increasing the dimension of the neighborhood dimension is expected to give more efficient solution within shorter time

Based on their algorithm, the difference between 2D and 3D neighborhood cGA is the addition of a third dimension, which refers to the layers of the grid, as the position of individuals located also on z axis. In this research, a 3D cubic topology is used. Previous study showed that a cubic topology allows good solutions to spread quickly due to its shorter diameter as well as diverse degrees of exploration and exploitation.

The topology of both 2D and 3D are considered to be linear with one step distance from the central of the cell but differ in the size of the neighborhood based on their dimension. 2D neighborhood has 4 neighbors (North, East, West and South) with a radius of 0.89, while 3D neighborhood has 6 neighbors (horizontal north and south, vertical north and south, east and west) with a radius of 0.925.

As consequence, 3D neighborhood has a denser neighborhood than 2D, leading to additional time for communication or computation, especially in tournament selection system, since this system needs access to all individuals in the neighborhood.

To increase the solution in shorter computational time while keeping the quality of solution with 3D structure, the pseudo code used in this research are being proposed as shown in Algorithm 1 below:

Algorithm1. Pseudo code of 3D-cGA.

```
1. procedure cGA
2. Generate_initial_population(P(0));
3.Evaluation(P(0));
4. while ! stop_condition() do
5.     for i ← 1 to ROWS do
6.       for j ←1 to COLUMNS do
7.         for k ←1 to LAYERS do
8. Neighbors ← compute_neighbors (position(i,j,k));
9. Parent1← position(i,j,k);
10. Parent2 ←Local_selection (neighbors);
11. offspring ← Recombine (Pc, parent1, parent2);
12. offspring ← Mutate(Pm);
13. Evaluation ← Fitness(offspring);
14. Replace_if_better (position(i,j,k), offspring,P_aux(t));
15. end for;
16. end for;
17. P(t+1) ← P_aux(t);
18. t ← t+1;
19. end while;
20. end procedure cGA
```

It starts with creating a random population (step 2), which is then the fitness function of each individual will be evaluated (step 3) and keep updating individuals by using genetic operators, until a termination condition is met (step 4). The current individual then calculates which can only interact with its neighbors (step 6). First parent will be chosen from the current one (step 9). Binary tournament then uses as the local selection method in the neighborhood to choose the second parent (step 10). Two child with the best fitness will be delivered after a one-point crossover is applied with a 90% probability Pc (step 11), followed by a bit-flip mutation with probability Pm of 0.02 (step 12).

Algorithm then calculates the fitness value (step 13) and replaces the current individual with the new one if it has better fitness value (step 14). Each population then replaces the previous one on a synchronous update, (step 17). The loop continues until termination condition is met (step 4).

In the binary tournament selection, the population fitness in generation t will be distributed normally $N(\mu_{F,t},\sigma^2_{F,t})$. Individual with fitness F from the binary tournament (s = 2) will have the expected value as the tournament winner, which can be calculated as:

$$\mu_{F,t+1} = \mu_{F,t} + \sigma_{F,t}\frac{1}{\sqrt{\pi}} \quad (3)$$

4. Verification Experiment

To confirm the efficiency of 3D neighborhood relationship, we made comparison data result obtained from 2D and 3D neighborhood of cGA.

Some data were input before the whole process began. These data can be divided into two parts:

- First, the visitor related data, such as visitor group index number, name, and stay time.
- Secondly, the tour guide related data, for example service capacity and preparation time.

Table1. Visitor Related Data

	Description	Value
No. of Visitors	Number of visitors	Positive Integer
Stay Time	Time that visitors will choose for stay	4, 8 hours

The service capacity varies from the experienced to the beginner tour guide. The preparation time also varies from an experienced to a beginner tour guide. It is always the experienced tour guide requires less preparation time.

Table2. Tour Guide Related Data

	Description	Value
Service Capacity	The tour guide has different capacity; it depends on the experience.	15, 20, 25
Preparation Time	This preparation time depends on service capacity of the tour guide itself.	1-2 hours

The tour guide and subgroup then are represented by a gene, which build the chromosome. Two-point crossover and two-point swapping mutation were being used for the crossover and mutation methods, respectively.

We set the quantity of visitors to be 1.500 and consider it as the base case. We determine the generation number varied from 50 to 500 and set the crossover rate to 0.9 while the mutation rate is equal to 0.02. Meanwhile, the population size is varied from 50 to 300.

From the number of tour guide and the number of visitor data, we calculated the best fitness value $\boldsymbol{FBest}$ from $\boldsymbol{Eq. 2}$ over generations. The average and the minimum best value then also determined.

The results of the experiments expected to be well solved within short time while keeping the quality of the solution by increasing the dimensionality of neighborhood.

References

[1] Alba, E., Dorronsoro, B., "Cellular Genetic Algorithm", ser. Operations Research/Computer Science Interfaces, Springer-Verlag Heidelberg, 2008.

[2] Breukelaar, R. and Back, "Using a Genetic Algorithm to Evolve Behaviour in Multidimensional Cellular Automata", *In Proceedings of GECCO'05*, 107–114, 2005.

[3] Morales-Reyes, A., Al-Naqi, A., Erdogan, A. T., and Arslan, T. , Towards 3D Architectures: "A Comparative Study on Cellular GAs Dimensionality". *In Proceedings of the NASA/ESA Conference on Adaptive Hardware and Systems, San Francisco, California, USA,* 223–229, 2009.

[4] Reyes, A.M., Arslan, T. and Erdogan, A., "Towards 3D Architectures: A Comparative Study on Cellular GAs Dimensionality", *NASA/ESA Conference on Adaptive Hardware and Systems*, 2009.

[5] Chen, R. and Wang, H., "Application of Genetic Algorithm to Scheduling of Tour Guides for Tourism and Leisure Industry", *INFOS*, Cairo-Egypt, 2008.

[6] El Zoghby, M., Badr, A. and Hegazy, A. E., "Schedulling of Tour Guides for Tourism and Leisure Industry using Cellular Genetic Algorithm", *IJCSNS*, 2011.

[7] Sarma, J. and De Jong, K., "An Analysis of the Effects of Neighborhood Size and Shape on Local Selection Algorithms", *Proceedings of the 4th International Conference on Parallel Problem Solving from Nature*, 1996.