Comparing Selection Method in Course Scheduling Using Genetic Algorithm

Ni Gusti Ayu Putu Harry Saptarini[1], Putu Indah Ciptayani[2], Ni Wayan Wisswani[3], I Wayan Suasnawa[4], Nyoman Eddy Indrayana[5]

Electrical Engineering Department
Politeknik Negeri Bali
Bali, Indonesia
[1]ayu.harry@pnb.ac.id, [2]putuindah@pnb.ac.id

Abstract— **Selection technique is an important operator in genetic algorithm (GA). Defining the best selection technique is critical in order to get the optimum solution for certain problem. The purpose of this study was to compare 3 selection techniques in high school scheduling problem using distributed GA (DGA). The selection techniques implemented in this study were roulette wheel, tournament and truncation selection. The migration probabilities used in DGA was 0.5. The results showed that the best selection method in avoiding early convergence is the tournament method**

Keywords— *DGA; roulette wheel; tournament; truncation; genetic algorithm, selection technique*

I. INTRODUCTION

The process of build a schedule between rooms, subjects, students and teachers to avoid the constraints violations is called timetable problem (TTP) [1]. The constraints can be hard or soft. To build a schedule for senior high school, some resources need to be arranged, so that the hard constraints violation can be avoided. The difficulty in preparing a schedule is the limited resources and the resulting schedule cannot violate existing constraints.

There are many algorithms in solving TPP such as local search [2,3], simulated annealing[4,5], cultural algorithm [6], and genetic algorithm(GA) [1,7-16]. An improved version of GA was using distributed population in GA that called distributed GA (DGA). Study [17-19] used DGA to solve combinatory problems. In DGA, there is an important operator, the migration rate. This operator control how individual exchange among groups. This rate can be constant or adaptive as found in study [20]. The adaptive parameter generally based on the fitness on certain generation, so the user will not be difficult to determine the parameter. Earlier study [21] using DGA to solve senior high school scheduling. The study showed that probability of migration (pmig≥0.3) can reach the maximal generation of GA.

The DGA operators include selection technique, crossover rate, mutation rate and migration rate. Selection plays an important role in DGA. The appropriate selection technique will lead to the optimum solution. There were many studies that comparing selection technique. Zhong et al. [22]

compared tournament and roulette wheel in simple GA (SGA). The numerical experiments showed that roulette wheel selection converges much slower than tournament selection in SGA. Other study done by Razali et al. [23] and Pandey et al. [24] compared tournament, roulette wheel and rank selection in travelling salesman problem (TSP). The first study [23] found that rank-based and roulette wheel are suitable for large size problem while the tournament selection can be used for smaller size problem. The second study [24] showed that tournament selection was the worst technique in terms of distance resulted. The better technique was roulette wheel, while rank-based selection was the best technique. Comparing selection technique also done by [25]. In that study three selection techniques: stochastic universal sampling, tournament selection and roulette wheel were compared. The study found that the roulette wheel technique produces more stable fitness value than the stochastic universal sampling and tournament selection. Chudasama and Panchal [26] compared roulette wheel, tournament selection and elitism selection method. The study conclude that elitism method was the best method. While Gangwar et al. [27] compared truncation selection, proportionate selection, roulette wheel selection, rank selection, steady state selection and tournament selection. The study showed that tournament selection was the best technique that gives the best chromosomes to fight in the tournament. Sohal et al [28] compared Elitism selection, rank selection, roulette wheel selection, tournament selection, Boltzmann selection and steady state selection. The result showed that the elitism in GA give the better speed to converge and result the optimal chromosome (solution) than the GA without elitism. Shukla et al [29] did a comparative review of ranking selection, tournament selection and roulette wheel selection. The study conclude that tournament selection give the better result in time complexity and convergence rate.

This study aims to explore the best selection technique in DGA for senior high school scheduling in order to reach the most optimal fitness and generation. Here three kinds of selection techniques: the tournament, roulette wheel and truncation selection were compared.

II. METHODOLOGY

A. Distributed Genetic Algorithm

One of heuristic technique for searching to solve the combinatory problem is genetic algorithm (GA)[30]. Genetic algorithms is widely use in several studies. The first step in GA is generate a set of random solution called population. A solution will represent as an individual called chromosome. The initial population will evolve during the iteration. Each iteration will produce one new generation. Each new generation is produced by selection, crossover and mutation. Each individual will be evaluated by a fitness function that indicates the resilience degree of the individual in adapting to the problem. After several iterations, the GA will converge on the best individual, which is expected to be the optimal solution for certain problem.

The DGA is a modification of a simple GA (SGA) to avoid the premature convergence. The different between SGA and DGA is DGA generates several population groups at once. For each iteration, the GA process will be conducted in each group. The migration of individual/chromosome between groups is conducted in the end of iteration process based on the migration rate. This migration is expected to keep the variation of the member in population, so that the local optimum can be avoided. The DGA flowchart in this study is shown in Figure 1.

B. The Representation of Chromosome

The solution or chromosome was represent using data structure as below :

```
TClass
    nameOfClass : String,
    shift : byte
end of TClass;
Class = array[1..j] of TClass;

TSection
    dayName  : String,
    timeslot : byte
end of TSection;
Section = array[1..k] of TSection;


TChromosome
    course
    teacher
end of TChromosome
Individual = array[1..j,1..k] of TChromosome
```

A chromosome will be composed of course and teachers. Individuals are represented in a two-dimensional array, where the first index represents the class index, while the second index represents the session index. For example, when index 1,1 shows the geography lesson and instructor Andi, then using class 1 will get a Geography taught by Andi in a session with index 1 (in the session there are course and hours).
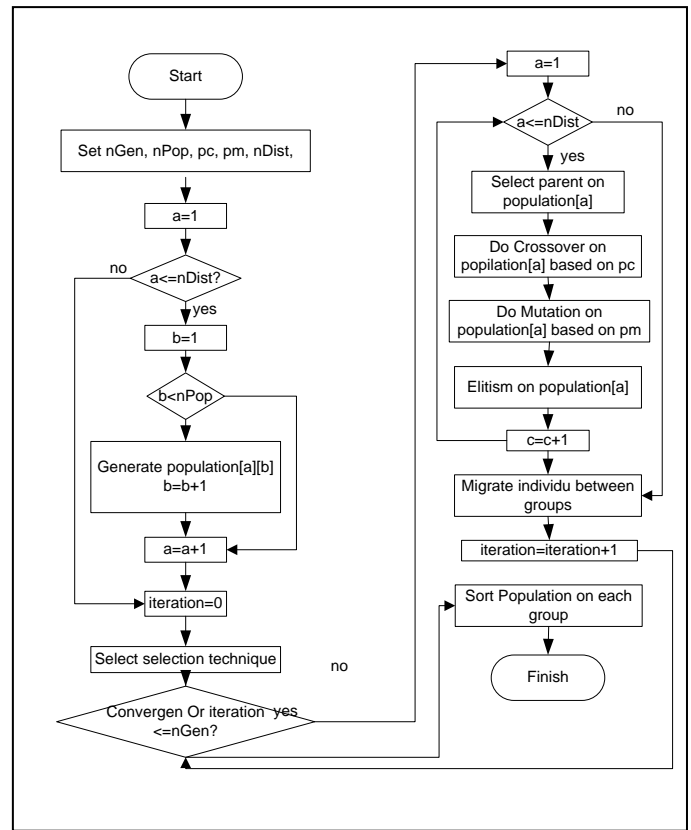

Fig. 1.   The modified DGA flowchart[21]

C. Fitness Function

There are two kind of constraint in this study, the soft and the hard contraints. The hard constraints are cannot be violated in the solution provided by GA, while the soft constraints can be violated, but the violation will decrease the fitness value of the solution. Any violation of the soft constraint will be subject to penalty. The more violations will result the bigger penalty for the solution. The objective of this study was to minimize the total penalty. The fitness function of this case is as follow:

$$\text{fitness} = 1/(1+\text{total\_penalty}) \qquad (1)$$

where total_penalty is the sum of all penalty caused by the soft constraint violation.

D. Selection

The selection aims to select a couple of parent to produce new offsprings. The selection method used in this study was roulette wheel, tournament and truncation method. The parent selection in roulette wheel is based on the proportion of the fitness. The probability of chromosome to be selected as a parent is affected by the fitness. The greater fitness will give the greater opportunity to be selected. Based on [31], the pseudo code of the roulette wheel method is as follow:

```
RW_pseudo_code
{
```
- Calculate the sum $S=\sum_{a=1}^{n} f(a)$;
- For each chromosome $1 \le a \le n$ do{

o Generate a random number $r \in [0,S]$;
o Sum =0; b-0;
o do{
  ▪ *Sum ←Sum+f(b);*
  ▪ *b ← b+1;*
  }
o while (*Sum< r and b<n*)
• select the chromosome *b*;}
}

The tournament selection will choose *k* number of individual and then rank all *k* individual to get the best 2 individual as a couple of parent. The pseudo code of tournament selection is as follow:

TS_pseudo-code
{
  • Select *k* random individual from population and place in an array S
  • Sort the array S based on the fitness descending
  • Pick top 2 individual
}

The truncation selection [32] is the simplest technique. The technique orders the chromosome of each population based on the fitness. Then, only a certain portion p of the fittest individuals are selected and reproduced 1 p times. It is less used in practice than other techniques, except for very large population. The pseudo-code of the technique is as follows [31]:

TRS_pseudo_code
{
  • Order the n chromosomes of P(i) based on their fitness;
  • Set the percentage pt of chromosomes to select (e.g. 10%≤pt≤50%)
  • c_number = n x p
  • Select the first c_number chromosomes;
}

E. Crossover

Crossover is conducted to produce a new offspring. The one-point crossover is used in this study. Because this technique simpler than others, so the time complexity can be reduced. The rate of crossover is defined by the probability of crossover (pc). Higher pc will produced more offprings.

F. Mutation

Mutation happened after crossover. It will change the genes in chromosome. The probability of mutation happened is based on the mutation rate/probability of mutation (pm). Generally, the value of pm is small, because based on fact the mutation probability is low. This process aims to prevent the premature convergence. In this study, the genes will be change if the mutation is happened, that means the teacher will be change by other teacher.

G. Elistism

Elitism is done in the end of iteration to eliminate the chromosomes that have the worst fitness. All of the offspring resulted by the crossover will be compared to all chromosomes in a population. If the fitness of the offspring is lower than all of the population members, so the offspring will not be a new member of population. On the other hand, if the fitness of the offspring is greater than one or more of the population member, the worst chromosome in a population will be eliminated. The elitism will keep the fittest chromosome over generation.

H. Chromosome Migration Between Groups

The DGA used in this study based on earlier study[21]. The DGA works with several number of population groups. One groups is consist of some chromosome. There will be a migration between one group to other group, so the chromosome in each group will be varied and the premature convergence is expected to be prevented. The migration will happened based on the migration rate/migration probability (pmig).

III. RESULT

This study used the scheduling case in SMA Surya Wisata, Jalan Wagimin, Kediri, Tabanan. The hard constraints are shown in Table I, while the list of soft constraint are shown in Table II.

TABLE I.   THE LIST OF HARD CONSTRAINTS

No	Hard Constraint
1	The teacher only teach the course according to their field
2	One teacher cannot be in two or more class on the same section
3	The first section on Monday cannot be assign in any course, because a national ceremony is held every Monday morning
4	One class cannot take one course in 4 hours or more in the same day

TABLE II.   THE LIST OF SOFT CONSTRAINTS

No	Soft Constraint	Penalty
1	Total teacher' section in a week exceed 50 hours or more	12
2	Total teacher' section in a week exceed 45 hours or more	11
3	Total teacher' section in a week exceed 40 hours or more	10
4	Total teacher' section in a week exceed 35 hours or more	9
5	Total teacher' section in a week exceed 30 hours or more	8
6	Distribution of uneven teaching hours for teachers in one area	8
7	Total teacher' section in a week exceed 25 hours or more	7
8	Total teacher' section in a week exceed 20 hours or more	6
9	A class get 3 hours of the same lesson on the same day	6
10	The morning lessons are completed over the 9 sections	5
11	The morning lessons are completed over the 8 sections	3
12	The afternoon section completed over 9	5
13	The afternoon section completed over 8	3
14	Lesson hours with 2 section/week load broken down on different days	2
15	There are interlude of lessons in a class	1

The hard constraint is the constraint that cannot be violated by the chromosome resulted in GA, on the other hand, the soft constraint violation will affect the fitness. The more soft

constraints are violated, the smaller fitness resulted by the solution.

The classes (group of students) were managed in this study were: 9 classes of first year students, 9 clasess of second year students and 9 classes of third year students. Some of classes were held in the morning session while the others were afternoon session. Each session will divided into time sections. Each class has their fix classroom. The DGA will managed 61 teachers for certain class, time section and course in order to generate a good schedule. The genetic parameters were shown in Table III.

TABLE III. THE PARAMETERS IN THIS STUDY

The population group number (group)	4
Chromosome number in each group	50
Maximal generation number	1000
Probablity of crossover (pc)	0.8
Probability of mutation (pm)	0.1
Probability of migration (pmig)	0.5
k for tournament selection	10
P for truncate selection	40%

The number of population generated in each group was 50 chromosomes. Then crossover is conducted based on the pc, the number of selected parent in one crossover is two chromosome, and it will result two new offspring. The mutation rate pm is set to be low. After the elitism procedure is done, the migration between groups is conducted based on the pmig. All of the process will be repeatedly conducted until the maximal generation is reach or all of the group reach their convergence state. Convergence state will be happen when the best fitness in the group is the same for 100 generation. The experiment was conducted 30 times for each selection technique. The results of experiments was shown in Table IV.

TABLE IV. THE EXPERIMENTAL RESULT

Selection method	Max gen	Average Fitness	Best Fitness
Roulette wheel	496	0.031	0.045
Tournament	687	0.043	0.051
Truncation	153	0.024	0.025

In Table IV, we can see that the comparison between each selection method. From the table we can conclude that the tournament selection give the longest run of generation, while the fastest method to converge was truncation. Because the truncation method converge so fast, this imply to the low fitness. The largest average and best fitness was resulted by tournament selection.

IV. CONCLUSION

This study used distributed genetic algorithm (DGA) and 3 selection methods to solve the time table problem in senior high school. This study managed the teacher and course into different classes and sections. Based on the experiment result, the best selection method was tournament method. This method gave the best average fitness compared to other and prevent the algorithm from premature convergence. The fastest convergence was reached by truncation selection. For further work, we suggest to investigate the various crossover technique for time table that give the better fitness.

REFERENCES

[1] A. Nanda, M. P. Pai, A. Gole, "An Algorithm to Automatically Generate Schedule for School Lectures Using a Heuristic Approach," International Journal of Machine Learning and Computing 2 (4), 2012, pp 492-495

[2] A. Schaerf, "Local Search Techniques for Large High-School Timetabling Problems," IEEE Transactions on Systems, Man, and Cybernetics - Part A: Systems and Humans 29(4), 1999, pp 368-377

[3] L. Saviniec, A. A. Constantino, "Effective local search algorithms for high school timetabling problems," Applied Soft Computing Journal vol 60, 2017, pp 363-373

[4] N. Basir, W. Ismail, N. M. Norwawi, "A Simulated Annealing for Tahmidi Course Timetabling," Procedia Technology vol 11, 2013, pp 437 – 445

[5] A. Sadegheih, "Scheduling problem using genetic algorithm, simulated annealing and the effects of parameter values on GA performance," Applied Mathematical Modelling vol 30, 2006, pp 147–154

[6] C. Sozaa, R. L. Becerrab, M. C. Riff, C. A. C. Coelloc, "Solving timetabling problems using a cultural algorithm," Applied Soft Computing vol 11, 2011, pp 337–344

[7] P. Pattanayak, P. Kumar, "A Computationally Efficient Genetic Algoritm for MIMO Broadcast Scheduling," Applied Soft Computing vol 36, 2015, pp 545-553.

[8] S. Kaiafa, A. P. Chassiakos, "A Genetic Algorithm for Optimal Resource-Driven Project Scheduling," Procedia Engineering vol 123, 2015, pp 260-267

[9] I. Sbeity, M. Dbouk, H. Kobeissi, "Combining the Analytical Hierarchy Process and The Genetic Algorithm to Solve the Timetable Problem," International Journal of Software Engineering & Applications 5 (4), 2014, pp 39-50

[10] S. Pappu, K. T. Talele, M. Mandviwala, "Application of Parallel Genetic Algorithm for Exam Timetabling Problem," International Journal of Scientific & Engineering Research 3(9), 2012, pp 1-4

[11] M. Mahto, L. Kumar, "Exam Time Table Scheduling using Genetic Algorithm," International Journal of Enhanced Research in Management & Computer Applications 4(8), 2015, pp 31-35

[12] S. Timilsina, R. Negi, Y. Khurana, "Genetically Evolved Solution to Timetable Scheduling Problem," International Journal of Computer Applications 114, 2015, pp 12-17

[13] D. Mittal, H. Doshi, M. Sunasra, R. Nagpure, "Automatic Timetable Generation using Genetic Algorithm," International Journal of Advanced Research in Computer and Communication Engineering 4(2), 2015, pp 245-248

[14] J. Gu, J. Hu, T. Zhao, G. Sun, "A New Resource Scheduling Strategy Based on Genetic Algorithm in Cloud Computing Environment," JOURNAL OF COMPUTERS 7(1), 2012, pp 42-52

[15] A. Jain, S. Jain, P.K. Chande, "Formulation of Genetic Algorithm to Generate Good Quality Course Timetable," International Journal of Innovation, Management and Technology, 1(3), 2010, pp 248-251

[16] E. A. Abdelhalim, A. A. El Khayat, "A Utilization-based Genetic Algorithm for Solving the University Timetabling Problem (UGA)," Alexandria Engineering Journal vol 55, 2016, pp 1395–1409

[17] L. Wang, A. A. Maciejewski, H. J. Siegel, V. P. Roychowdhury, "A Study of Five Parallel Approaches to a Genetic Algorithm for the Traveling Salesman Problem," Intelligent Automation and Soft Computing, 11(4), 2005, pp 217-234

[18] E. Alba, J. M. Troya, "A survey of parallel distributed genetic algorithms," Complexity 4(4), 1999, pp 31-52

[19] T. Eldos, "A New Migration Model For Distributed Genetic Algorithms," Proceedings of the International Conference on Scientific Computing (CSC' 06), Las Vegas, NV, 2015, pp 128–134

[20] H. Lee, B. Oh, J. Yang, S. Kim, "Distributed genetic algorithm using automated adaptive migration Evolutionary Computation," 2009 IEEE Congress on Evolutionary Computation, 2009, pp 1835-1840

[21] N. G. A. P. H. Saptarini, I. W. Suasnawa, P. I. Ciptayani, "Senior high school course scheduling using genetic algorithm," Journal of Physics: Conference Series, 953(1), 2018

[22] J. Zhong, X. Hu, J. Zhang and M. Gu, "Comparison of Performance between Different Selection Strategies on Simple Genetic Algorithms," International Conference on Computational Intelligence for Modelling, Control and Automation and International Conference on Intelligent Agents, Web Technologies and Internet Commerce (CIMCA-IAWTIC'06), Vienna, 2005, pp. 1115-1121

[23] N. M. Razali, J. Geraghty, "Genetic Algorithm Performance with Different Selection Strategies in Solving TSP," Proceedings of the World Congress on Engineering 2011 Vol IIWCE, 2011

[24] H. M. Pandey, "Performance Evaluation of Selection Methods of Genetic Algorithmand Network Security Concerns," Procedia Computer Science 78, 2016, pp. 13 – 18

[25] A. D. Irfianti, R. Wardoyo, S. Hartati, E. Sulistyoningsih, "Determination of Selection Method in Genetic Algorithm for Land Suitability," MATEC Web of Conference 58, 2016, pp. 1-5

[26] C. Chudasama, M. Panchal, "Comparison of Parents Selection Methods of Genetic Algorithm for TSP," International Journal of Computer Applications® (IJCA), 2011

[27] M. Gangwar, M. Din, V. K. Jha, "Comparative Study Of Selection Methods In Genetic Algorithm," International Journal of Soft Computing and Artificial Intelligence Volume-5, Issue-1, May-2017

[28] S. L. Yadav, A. Sohal, "Comparative Study of Different Selection Techniques in Genetic Algorithm," International Journal of Engineering, Science and Mathematics Vol. 6 Issue 3, July 2017, pp. 174-180

[29] A. Shukla, H. M. Pandey and D. Mehrotra, "Comparative review of selection techniques in genetic algorithm," International Conference on Futuristic Trends on Computational Analysis and Knowledge Management (ABLAZE), Noida, 2015, pp. 515-519

[30] M. Gen, and R. Cheng, Genetic algorithms and engineering optimization, New York : A Wiley-Interscience Publication, 2000

[31] K. Jebari, "Selection Methods for Genetic Algorithms," International Journal of Emerging Sciences, 3, 2013, pp 333-344

[32] P. Hancock, "A comparison of selection mechanisms," In Handbook of Evolutionary Computation, Eds . IOP Publishing and Oxford University Press., Bristol, UK, 1997