INTERNATIONAL JASIN MULTIMEDIA & COMPUTER SCIENCE INVENTION & INNOVATION EXHIBITION
i-JaMCSIIX 2020

i-Saturate: The New Discovery of Stopping Criterion in Genetic Algorithm

Foo Fong Yeng[1], Azrina Suhaimi [2], PM Dr Soo Kum Yoke[3]
UiTM Kampus Pasir Gudang[1]
foofo931@uitm.edu.my
UiTM Kampus Pasir Gudang[2]
azrin253@uitm.edu.my
UiTM Kampus Seremban[3]
sooku607@uitm.edu.my

ABSTRACT

A stopping criterion for evolutionary algorithms like Genetic Algorithm (GA) is crucial in determining the optimum solution. It is common for a stopping criterion like *maximum generations* or *fittest chromosome repetition* used in GA to solve hard optimization problems. However, these stopping criteria require human intervention to make certain changes. In this study, a new stopping criterion called *i-Saturate* that measures saturation of population fitness of every generation chromosome (in GA searching process) is reported. The searching process would stop when the fitness deviation of the population was small. A model using *fittest chromosome repetition* was developed to compare the efficiency with *i-Saturate*. It was found that the performance of the developed model was good at the low mutation rate (0.01,0.02) but the *i-Saturate* model was better when mutation rate was greater than 0.03. The probabilities of the *i-Saturate* model finding global optimum solution were very close to 1 when mutation rate was above 0.07. It was concluded that the *i-Saturate* model has demonstrated better searching ability than the comparative model and it intelligently stops searching without human intervention.

KEYWORDS: Stopping Criterion, Genetic Algorithm, Optimization, Machine Learning

1 INTRODUCTION

Historically, the Darwin Evolution Theorem has inspired John Holland [1] to introduce GA in 1975 that imitates the process of genetic inheritance evolution [2]. GA begins the search by a population of randomly generated feasible solutions where they are encoded into chromosomes [3] and each of them is assigned with a fitness function. They are placed into an environment analogue of natural evolution where they need to survive, adapt, and propagate their genetics to the future generations [4]. The evolution takes many generations to converge to a perfectly adapted chromosome (global optimum solution) [4]. The timing for ending the evolution (searching process) is crucial. If the search ended too early (premature convergence), the identified solution might be only the best in that generation (local optimum) but not the best of all possible solutions (global optimum solution) [5]. The choice of the stopping criterion would determine if the global optimum solution could be found before the searching is terminated.

Hence, the stopping criterion of GA was investigated in this study since the right choice of stopping criterion could greatly help in determining the optimum solution.

2 OBJECTIVES

To accomplish this study, we proposed the density or saturation of population fitness as the new stopping criterion which served as a measurement key to end the searching process. We named this proposed stopping criterion as *i-Saturate* (Fig 2). *i-Saturate* measures density or saturation of population fitness (F) of M chromosomes of every generation. It stops searching when the population becomes saturated with fittest chromosome where the fitness deviation of the population was small ($[(1/M)\sum(Fij - F)^2] < \delta, \delta \rightarrow 0$).

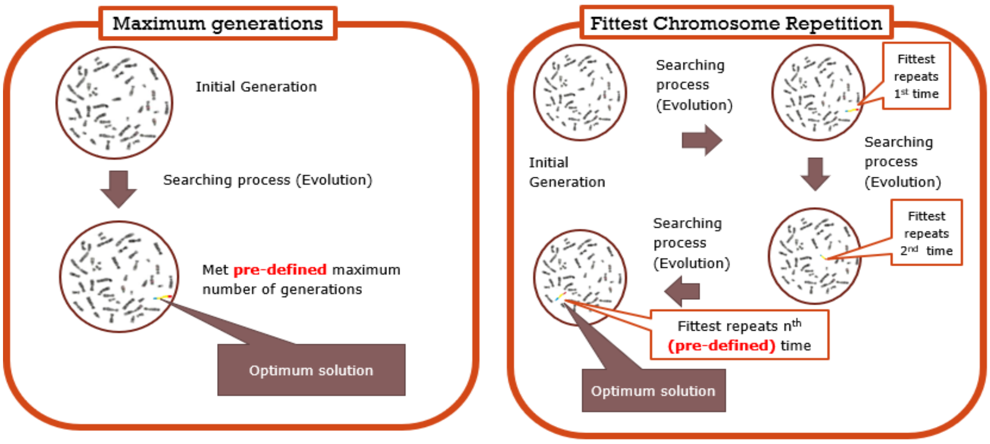

Fig. 1. Common stopping criteria

i-Saturate would be compared to a conventional stopping criterion (*Fittest chromosome repetition*) (Fig 1) which the searching stops when there is no improvement of fittest chromosome for some successive generations. Therefore, besides proposing a new stopping criterion, the research objective of this study is to compare the efficiency of *i-Saturate* with the conventional stopping criterion in the hope that the right choice of stopping criterion could be discovered.

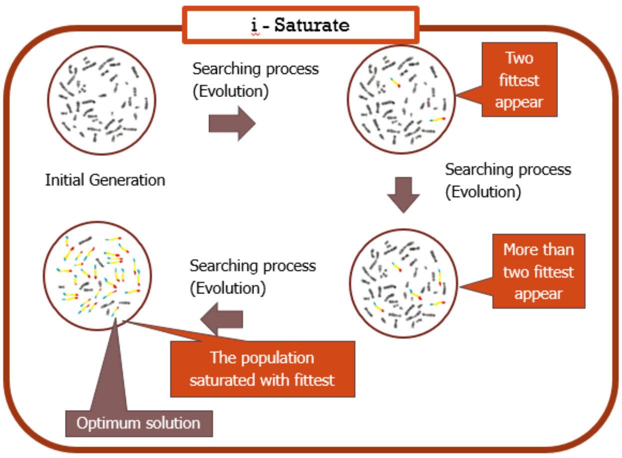

Fig. 2. *i-Saturate* (proposed stopping criterion)

3 SIGNIFICANCE (S)

GA has been widely applied in research areas such as pure sciences, engineering and social sciences [6]-[10]. Though GA can provide good solutions at reasonable computational cost

[11], it may not be able to guarantee the optimality of solution due to their stochastic nature [5]. The timing to end the evolution in GA processes is crucial to achieve optimal solutions. Therefore, the decision of having an appropriate stopping criterion would greatly affect the capability of the algorithm [12].

Generally, there are two stopping criteria (Fig. 1) that have been widely used: (1) *Maximum generations* [12] and (2) *Fittest chromosome repetition* [9]. For the first criterion, the user needs to define a maximum number of generations and the size of the population would sometimes influence the duration required for convergence [13]. For the second criterion, again, the user needs to determine the appropriate number of successive repetitions for dismissing the search. The successive repetition of the fittest chromosome is greatly influenced by the size and complexity of the research problem [12]. Since these two criteria need human intervention in pre-setting/predefined, a research that reduces the possibility of human intervention and dependency on human decision could give contributions in the field of enhancing GA works.

4 METHODOLOGY/TECHNIQUE

To test the proposed stopping criterion, two models with different stopping criteria were developed. The first model, named as Normal stopping criterion (Nsc) model, was designed with an algorithm that ends the searching process when *fittest chromosome repetition* have reached the plateau, prefix upper bound. The second model, named as the Saturation stopping criterion (Ssc) model, was created with the proposed stopping criterion, *i-Saturate*. The searching was stopped when the generation saturated with the fittest chromosomes.

5 RESULT

A good model of algorithm should be able to produce a simulated result that is very close to actual value under various circumstances of different crossover rate and mutation rate. Therefore, Nsc and Ssc models were tested for their capabilities in forecasting simulation by using a set of time series data range [1300, 1600]. In this research, both models carried out a forecasting simulation experiment with 1000 trials to accumulate the statistical records of stopping criteria efficiency. The genetic drift / parameters of both experiment models are standardized as in Table 1.

Table 1. Genetic drift / parameters for experiment models

Population size, M = 20 chromosomes
Crossover rate, CR = [0.5, 0.9] with rate interval 0.1
Mutation rate, MR = [0.01,0.1] with rate interval 0.01

Fig. 3 shows the comparison of Ssc and Nsc models performances under (a) CR=[0.5, 0.9] with interval rate 0.1 and (b) under MR=[0.01, 0.1] with interval rate 0.01. From both the graphical representations, the Ssc model has demonstrated tracking ability better than the Nsc model. Fig.3 (a) illustrates the probabilities of the Ssc model in generating good simulation model were close to 1 especially when CRs were high, meanwhile, in Fig. 3 (b), the Nsc model was better at the low mutation rate (MR=[0.01, 0.02]) but the prevailing characteristic of the Ssc model emerged when MR was greater than 0.03. The possibilities of the Ssc model having a good forecast were very close to 1 when MR was above 0.07. From both the graphical representations, the Ssc model has demonstrated better searching ability than the Nsc model.

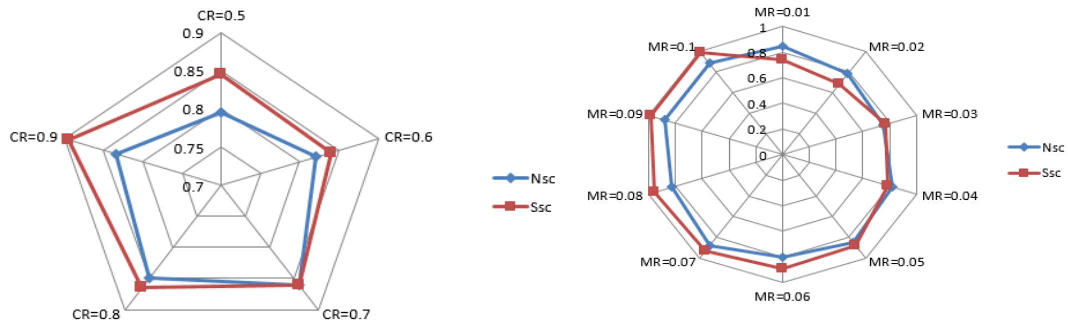

Fig.3: under CR=[0.5, 0.9] with interval rate 0.1 and under MR=[0.01, 0.1] with interval rate 0.01

6 CONCLUSIONS

The right choice of stopping criterion could enhance the capability of Genetic Algorithm. As the research has demonstrated, the Ssc model that adopted *i-Saturate* was better in finding a global optimum solution though MR and CR were not high. The capability of the Ssc model boosted when CR and MR had been set high. Thus, our proposed stopping criterion, *i-Saturate* has shown a great improvement in enhancing the algorithm ability in solving the optimization problem and reducing the risk of premature convergence. *i-Saturate* is also developed with no pre-setting by humans makes this stopping criterion diverse from other common stopping criteria. For future study, *i-Saturate* may also be tested in other fields of optimization such as engineering and one may research on the influence of different genetic operators in evolution.

REFERENCES

[1] H. M. Pandey, A. Chaudhary, and D. Mehrotra, "A comparative review of approaches to prevent premature convergence in GA," Applied Soft Computing, vol. 24, pp. 1047–1077, Nov. 2014.

[2] Benlahbib, F. Bouchafaa, S. Mekhilef, and N. Bouarroudj, "Wind Farm Management using Artificial Intelligent Techniques," International Journal of Electrical and Computer Engineering (IJECE), vol. 7, no. 3, p. 1133, 2017.

[3] S. Kalantari and M. S. Abadeh, "An Effective MultiPopulation Based Hybrid Genetic Algorithm for Job Shop Scheduling Problem," Bulletin of Electrical Engineering and Informatics, vol. 2, no. 1, pp. 59–64, 2013.

[4] X. Yan, W. Luo, W. Li, W. Chen, C. Zhang, and H. Liu, "An Improved Genetic Algorithm and Its Application in Classification," TELKOMNIKA Indonesian Journal of Electrical Engineering, vol. 10, no. 1, pp. 337–346, 2013.

[5] V. N. Wijayaningrum and W. F. Mahmudy, "Fodder Composition Optimization using Modified Genetic Algorithm," International Journal of Intelligent Engineering and Systems, vol. 9, no. 3, pp. 1–8, 2016.

[6] A. P. Unnikrishnan, R. Romeo, and F. A. Rawther, "License Plate Localization Using Genetic Algorithm including Color Feature Extraction," Procedia Technology, vol. 24, pp. 1445–1451, 2016.

[7] F. Jolai, M. Rabbani, S. Amalnick, A. Dabaghi, M. Dehghan, and M. Yazadn Parast, "Genetic algorithm for bi-criteria single machine scheduling problem of minimizing maximum earliness and number of tardy jobs," Applied Mathematics And Computation, vol. 194, pp. 552–560, 2007.

[8] R. Venkatesan and V. Kumar, "A genetic algorithms approach to growth phase forecasting of wireless subscribers," International Journal of Forecasting, vol. 18, no. 4, pp. 625–646, 2002.

[9] R. Venkatesan and V. Kumar, "A genetic algorithms approach to growth phase forecasting of wireless subscribers," International Journal of Forecasting, vol. 18, no. 4, pp. 625–646, 2002.

[10] B. C. Ervural, O. F. Beyca, and S. Zaim, "Model Estimation of ARMA Using Genetic Algorithms: A Case Study of Forecasting Natural Gas Consumption," Procedia - Social and Behavioral Sciences, vol. 235, no. October, pp. 537–545, 2016.

[11] N. Hassan et al., "Number of iteration analysis for complex fss shape using GA for efficient ESG," Bulletin of Electrical Engineering and Informatics, vol. 7, no. 4, pp. 505–513, 2018.

[12] C. Oĝuz and M. F. Ercan, "A Genetic Algorithm for Hybrid Flow-shop Scheduling with Multiprocessor Tasks," Journal of Scheduling, vol. 8, no. 4, pp. 323–351, Jul. 2005.

[13] A. R. Melo, M. M. S. Loureiro, and F. Loureiro, "Blood Perfusion Parameter Estimation in Tumors by means of a Genetic Algorithm," Procedia Computer Science, vol. 108, pp. 1384–1393, 2017.