Proceedings of the 4th International Conference on Computing and Informatics, ICOCI 2013
28-30 August, 2013 Sarawak, Malaysia. Universiti Utara Malaysia (http://www.uum.edu.my)

Paper No. 091

ROULETTE-TOURNAMENT SELECTION FOR SHRIMP DIET FORMULATION PROBLEM

Rosshairy Abd. Rahman[1] and Razamin Ramli[2]

[1,2]Universiti Utara Malaysia, Malaysia, [1]shairy@uum.edu.my, [2]razamin@uum.edu.my

ABSTRACT. This paper aims to propose a new selection procedure for real value encoding problem, specifically for shrimp diet problem. This new selection is a hybrid between two well-known selection procedure; roulette wheel selection and binary tournament selection. Shrimp diet problem is investigated to understand the hard constraints and the soft constraints involved. The comparison between other existing selections is also described for evaluation purposes. The result shows that roulette-tournament selection is better in terms of number of feasible solutions achieved and thus suitable for real value encoding problem. However, the combination with other crossover or mutation might be investigated to find the most suited combination that can obtain better best so far solution.

Keywords: evolutionary algorithm, diet problem, selection, shrimp diet

INTRODUCTION

Selection or reproduction is one of the important operators used in genetic algorithms (GA). Basically, the purpose of selection process is to choose better solutions to be parents for the next step and delete the remaining worse solution (Deb, 2000). Sivaraj and Ravichandran (2011) review several selection procedures in GA. These procedure are roulette wheel selection, deterministic sampling, linear ranking selection, binary tournament selection, range selection and many more. Different selection mechanisms work well under different problem (Sivaraj & Ravichandran, 2011). Thus, the most suitable procedure has to be chosen for the specific problem to increase the optimality of the solution.

Roulette wheel selection (RWS) has been proposed by Holland in 1975 and has been used widely in the application of GA. It becomes one of the most popular selection procedures that are based on the concepts of proportionate. Conceptually, the fitness value of each individual in the population is corresponds to the area on the roulette wheel proportion. Then, the roulette wheel is spin; a solution marked by the roulette wheel pointer is selected. Higher fitness with bigger area is likely to have more chances to be chosen. The segment size and selection probability remain the same throughout the selection phase (Chipperfield, 1997). The advantage of RWS technique is it gives no bias with unlimited spread (Chipperfield, 1997). However, one of the disadvantages of RWS is it cannot handle negative fitness values due to the proportionate concept (Deb, 2000). Also, RWS cannot handle minimization problem directly. However, this limitation can be overcome by transforming it into equivalent maximization problem.

Meanwhile, competition among a group of parents is the basis of tournament selection procedure. Measurement of fitness of solution is made among all parents and the parent having the best fitness is selected. The term ‘binary tournament’ refers to two tournament size which is the simplest form of tournament selection (Deb, 2000). Binary tournament selection starts by selecting two individuals in random. Then, fitness values of these individuals are evaluated. The one having the better fitness is chosen. One advantage of tournament selection is its ability to handle either minimization or maximization problems without any structural changes. In addition, negative value is allowed without any restriction.

The main concern of this paper is to develop a new selection procedure that combines two established selection schemes; RWS and binary tournament selection. GA with this new proposed roulette-tournament selection is applied for diet formulation model for juvenile Whiteleg shrimp which satisfy all the constraints with minimum cost. Whiteleg shrimp is chosen because this species is the most popular cultured shrimp in Asia and Malaysia as well. Whiteleg shrimp contributes to nearly 80% of total shrimp production in Malaysia (FAO, 2012).

Several constraints on shrimp diet were considered including ration weight, nutritional range, and ingredient range. These were define through experts opinion and literature review. Nutritional range is classified into three; single nutrient, combination of nutrients, and ratio between two nutrients. A system prototype is then developed to allow user to put preferred ration weight and choose the preferred ingredients and search for the most economic diet. The final result is the list of selected ingredients in specific quantities that will satisfy all the stated constraints.

METHODOLOGY

The roulette-tournament selection procedure introduced in this study is a combination of RWS procedure and binary tournament selection. The procedure starts with the same steps as RWS. Then, the binary tournament procedure take place by choosing two individuals as parents. As in binary tournament, two individuals are randomly picked from all solutions, and the fitter parents will be chosen as parent one. The same step is repeated to find parent two. The hybridization of this procedure will merge the advantages from both RWS and binary tournament.

The evolutionary model consists of initialization, roulette-tournament selection, one-point crossover, power mutation and steady state reproduction as shows in Figure 1. In addition, elitism procedure is also inserted because it can increase GA performance as it prevents the loss of best found solution (López-Pujalte, Bote & Anegón, 2002 and Sharief, Eldho & Rastogi, 2008). However, this paper specifically focused on roulette-tournament selection procedure. In order to develop the model, objective function and the constraints involved in shrimp diet problem are illustrated in mathematical formulation in the next subsection. Meanwhile, for comparison purposes, two existing selection schemes; RWS and Queen Bee selection, are also developed. The results are then evaluated by comparing the results from these selection schemes.

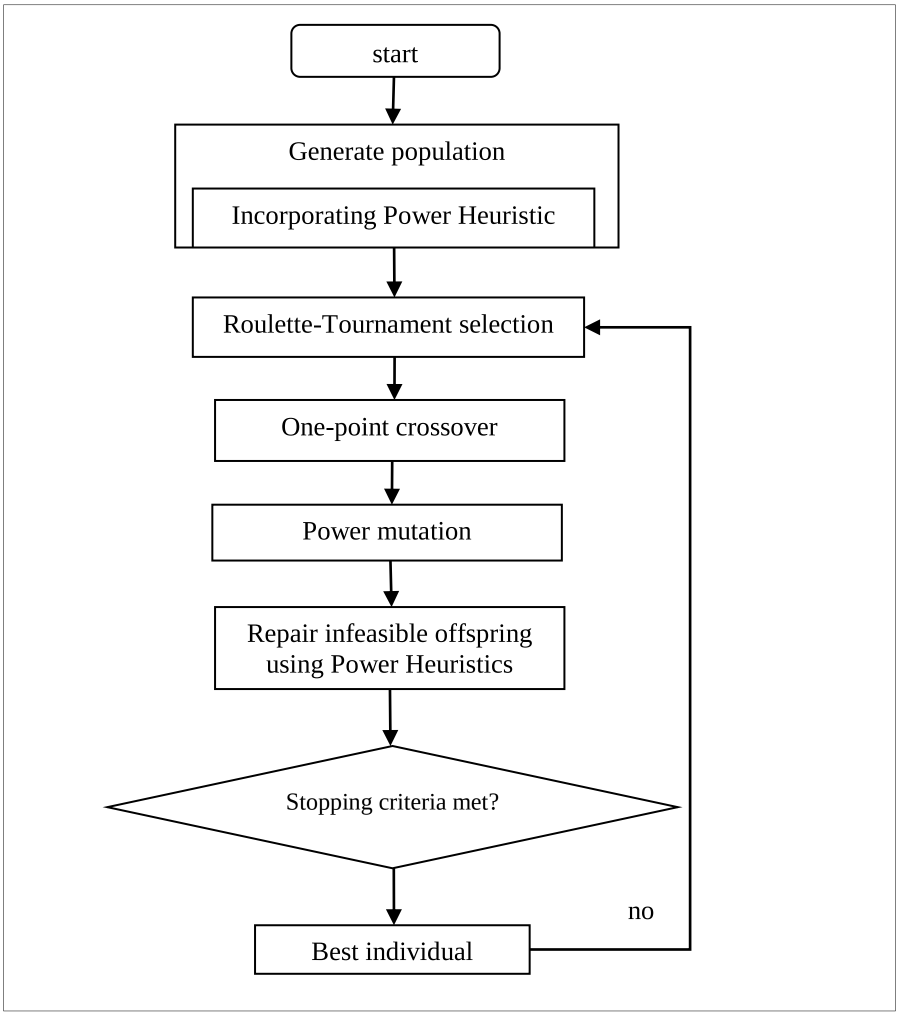

Figure 3. Evolutionary Model.

Mathematical Formulation

The performance of three GA models with different selection schemes are tested using real data for animal diet formulation problem. In this problem, the aim is to satisfy all the nutritional needs of farmed shrimps at a minimum cost. The minimization problem takes into account 14 ingredients and 18 nutrients. The following are the objective function and constraints involved in this problem. Objective function of the feed cost is defined as:

$$f(s) = \min \sum_{i=1}^{n} (X_i \times C_i), \tag{1}$$

where *Ci* is the cost of ingredient *i*,

Xi equals the weight of the *i*th ingredient, and

s is cumulative cost in a string of chromosome.

However, the aim of this study is to firstly reduce the penalty function value based on all identified constraints. The constraints consist of ingredients' range, ingredient (ration) weight, number of ingredients, single nutrient's range, combination nutrients' range, and ratio of nutrients.

- Ingredients' range:

$$X_i = 0 \text{ or } L_{Xi} \le X_i \le U_{Xi} \quad \text{for all } X_i, \tag{2}$$

where L_{Xi} = lower bound of ingredient *i,*

U_{Xi} = upper bound of ingredient *i,*

X_i equals the weight of the *i*th ingredient.

- Ingredient weight:

$$\sum_{i=1}^{n} X_i = Y, \tag{3}$$

where Y is a weight predefine by user in user interface.

- Number of ingredient:

$$n \le 14. \tag{4}$$

- Single nutrients' range:

$$L_{Nk} \le \sum_{i=1}^{n} N_{ki} X_i \le U_{Nk}, \tag{5}$$

where L_{Nk} = lower bound of nutrient *k,*

U_{Nk} = upper bound of nutrient *k,*

N = total value of nutrient *k.*

- Combination nutrients' range:

$$L_{Nk(i+j)} \le \sum_{i=1}^{n} N_{k(i+j)} X_i \le U_{Nk(i+j)}, \tag{6}$$

where $L_{Nk\,(i+j)}$ = lower bound of combination nutrient $i+j$,

$U_{Nk\,(i+j)}$= upper bound of combination nutrient $i+j$.

- Ratio nutrients' range:

$$L_{ratio} \le \frac{\sum_{i=1}^{n} N_{ki}}{\sum_{i=1}^{n} N_{kj}} \le U_{ratio}, \tag{7}$$

where L_{ratio} = lower bound of ratio between nutrient *i* and *j*,

U_{ratio} = upper bound of ratio between nutrient *i* and *j*.

Fitness calculation for the GA is basically based on penalty value for each constraint. There are two types of constraint; hard and soft constraints. In this study, hard constraints are ingredient (ration) weight, number of ingredient, and protein range constraint. Else, for soft constraints, different penalty values are given for different constraints based on in depth discussion with experts. Penalty value of 20 is given for violating each ingredient constraint, except for certain important ingredients; 30 is given for single nutrient, 20 for combination of nutrients, and 20 for ratio of nutrient.

RESULT AND DISCUSSION

In our experiments, GA parameters were set as follow: size of a population is 30, number of generation is 200, crossover rate is 0.60, and power value for power mutation is 0.25. Table 1 illustrates the simulated results of all GA models. From the Table, we summarize the best so far solution, average fitness, standard deviation and processing time (in second) taken to produce the best-so-far solution. These values are used as an indicator to evaluate the performance of these GA models.

Table 1. The Results of GA Models with Different Selection Schemes

Model	Best so far solution	Average Fitness	Standard Deviation	Run Time
Roulette-Tournament	460	699.6000	129.4694	1590.1455
Roulette Wheel	340	554.5833	127.1418	5144.0216
Queen Bee	410	687.1429	145.5192	1583.2123

Normality test is done with the intention to check either the data is normally distributed or not. The result from Shapiro-Wilk test confirmed that all the distribution from each model is normal. From Table 1, Roulette-Tournament One-point gives the worst solution with 460 fitness value. However, from 30 runs, Roulette-Tournament One-point obtained only 5 infeasible solutions, compared to Roulette Wheel-One-point with 6 infeasible solutions, and Queen Bee One-point with 16 infeasible solutions.

Standard deviation for Roulette Wheel One-Point is the lowest value. It is then followed by Roulette-Tournament One-point model and Queen Bee One-point. Standard deviation shows the deviation or dispersion of the data from mean. The lower standard deviation indicates that the model give stable solution, which mean it is always approaching mean. Run times shows that Roulette-Tournament One-point and Queen Bee-One-point each give approximately equal time.

CONCLUSION

The performance of basic GA model with different selection scheme is described. In this paper, we extend the basic GA model by introducing the roulette-tournament selection. The results show that our proposed selection procedure can be used in problems with real value encoding. In future research, this new selection scheme might be used with other crossover or mutation scheme to get the most appropriate combination for real value encoding problem.

REFERENCES

Chipperfield, A. (1997). Introduction to genetic algorithms. In A. M. S. Zalzala & P. J. Fleming, *Genetic Algorithms in engineering systems*. (pp. 1-41). London: The Institution of Electrical Engineers. Food and Agriculture Organization of the United Nations (FAO). Retrieved September 26, 2012, from http://www.fao.org/fishery/statistics/software/fishstat/en

Deb, K. (2000). Introduction to selection. In T. Bäck, D. B. Fogel, & T. Michalewicz, *Evolutionary Computation 1: Basic algorithms and operators* (pp. 166-171). Bristol: Institute of Physics Publishing.

Food and Agriculture Organization of the United Nations (FAO). Retrieved September 26, 2012, from http://www.fao.org/fishery/statistics/software/fishstat/en

López-Pujalte, C., Bote, V. P. G. and Anegón, F, d., M. (2002). A test of genetic algorithms in relevance feedback. *Information Processing and Management, 38*, 793–805.

Sharief, S., Eldho, T., & Rastogi, A. (2008). Elitist GA Based Evolutionary Algorithm for Groundwater. *The 12th International Conference of International Association for Computer Methods and Advances in Geomechanics (IACMAG).* Goa, India.

Sivaraj, R., & Ravichandran, T. (2011). A Review of selection methods in genetic algorithm. *International Journal of Engineering Science and Technology*, 3792-3797.

Charkha, P. R. (2008). Stock Price Prediction and Trend Prediction Using Neural Networks. *First International Conferenceon Emerging Trends in Engineering and Technology* (pp.592–594). IEEE. doi:10.1109/ICETET.2008.223

Dzikevičius, A., & Šaranda,S. (2010). EMA versus SMA Usage to Forecast Stock Markets: The Case ofS#38;P500 and OMXBaltic Benchmark. *Business: Theory and Practice*, *11*(3), 248–255. doi:10.3846/btp.2010.27

Gupta,A., & Dhingra, B. (2012). Stock Market Prediction Using Hidden Markov Models. *IEEE*. Hassan,M. R., Nath,B., & Kirley, M.(2007). Afusion model of HMM, ANN and GAfor stock market forecasting. *Expert Systems with Applications*, *33*(1), 171–180. doi:10.1016/j.eswa.2006.04.007

Kamijo, K., & Tanigawa, T. (1990). Stock price pattern recognition- a recurrent neural network approach. *1990I JCNN International Joint Conference on Neural Networks,1*(1),215–221. doi:10.1109/IJCNN.1990.137572

Lee, K.H., & Jo, G.S.(1999). Expert system for predicting stock market timing using a candle stick chart. *Expert Systems with Applications*, *16*(4),357–364. doi:10.1016/S0957-4174(99)00011-1

Maknickienė, N., & Maknickas, A. (2012). Application of Neural Networkf or Forecasting of Exchange Rates and Forex Trading. *The 7th International Scientific Conference "Business and Management 2012". Selected papers* (pp.122–127). Vilnius, Lithuania: VilniusGediminas Technical University Publishing House Technika. doi:10.3846/bm.2012.017

Niels,R. (2004). Dynamic Time Warping: Anintuitive way of handwriting recognition ? Radboud University Nijmegen.

Olaniyi, S. A. S., Adewole, K. S., & Jimoh, R. G. (2011). Stock Trend Prediction Using Regression.