Comparative Review of Selection Techniques in Genetic Algorithm

Conference Paper · February 2015

DOI: 10.1109/ABLAZE.2015.7154916

3 authors:

Anupriya Shukla
Amity University

Hari Mohan Pandey
Edge Hill University

Deepti Mehrotra
Amity University

Comparative Review of Selection Techniques in Genetic Algorithm

Anupriya Shukla
M.Tech CSE
ASET, Amity University, Noida, India
anupriyashukla2603@gmail.com

Hari Mohan Pandey
Assistant Professor, CSE Department
ASET, Amity University, Noida, India
hmpandey@amity.edu

Deepti Mehrotra
Professor, ASET,
Amity University, Noida, India
dmehrotra@amity.edu

***Abstract*— This paper compares various selection techniques used in Genetic Algorithm. Genetic algorithms are optimization search algorithms that maximize or minimizes given functions. Indentifying the appropriate selection technique is a critical step in genetic algorithm. The process of selection plays an important role in resolving premature convergence because it occurs due to lack of diversity in the population. Therefore selection of population in each generation is very important. In this study, we have reported the significant work conducted on various selection techniques and the comparison of selection techniques.**

***Index Terms*— Genetic Algorithm, tournament selection, Ranking Selection, proportionate roulette wheel selection.**

I. Introduction

The focus of this paper is towards comparative analysis of different selection techniques used in genetic algorithms (GAs). GAs are "adaptive heuristic search algorithms" that are designed to mimic Darwin's principle of "survival of the fittest" [22]. Which includes the following approach: individuals within a population compete for resources and the individuals that are most successful in the competition will produce offspring instead of those who didn't get success. Genes are transmitted over generations where pair of individuals get selected to produce better results/offspring in next generations [17]. Therefore, the individual that emerges out at the end is the fittest and more dominating than the weaker ones.

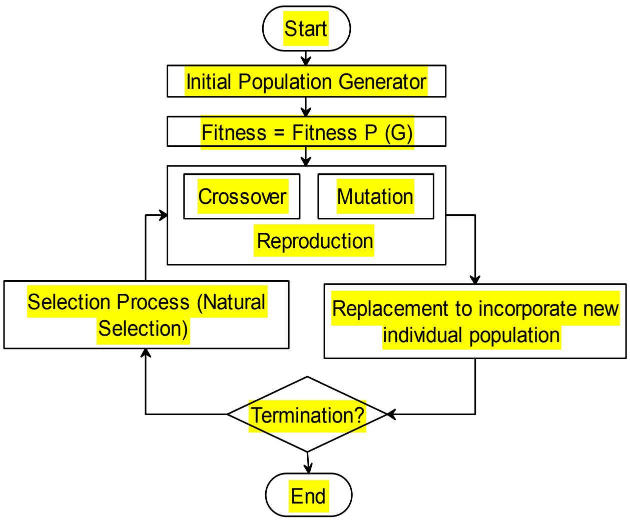

Fig. 1. Steps involved in Simple Genetic Algorithm [16]

Figure 1 depicts the working of simple genetic algorithm. The selection of the pair of individual starts immediately, as soon as reproduction operation ends. There are various approaches suggested to select the parent's string which may vary depending on the difficulty level of the problems. But, all in all making decision about appropriate selection method to be applied is considered as a difficult step and the same step is responsible for the speed of evolution and sometimes called as selection pressure. The different selection techniques applied in GA search process influence the overall performance.

GA comprises of a population of strings of character in each generation that are similar to the chromosomes of a DNA [1]. In which every chromosome of the population represents a possible solution in the search pool [18]. GA finds applications in various domains such as image processing, information retrieval, grammar induction [16] [19] [20], data mining, natural language processing [19] and others.

In evolutionary search, the diversity decreases as the population converges to a local optimum. It is the situation when an extraordinary individual takes over a significant proportion of finite population and leads towards an undesirable convergence. In other word, when a non-optimal genotype taking over a population, results in every individual being either identical or, the consequence of which is population that does not contains sufficient genetic diversity to evolve further. Hence, to reach to the global solution and explore search space adequately, maintaining population diversity is very important. Also, it had been explained clearly that degree of population diversity is one of the main cause of premature convergence. It is necessary to select the best solution of the current generation to direct the GA to reach to the global optimum. The tendency to select the best member of the current generation is known as selective pressure. It is also a key factor that plays an important role in maintaining genetic diversity. A proper balancing is required between genetic diversity and selection pressure to direct GA search to converge in a time effective manner and to achieve global optima. High selective pressure reduces the genetic diversity; hence in this situation premature convergence can occur while the little selective pressure prohibits GA to converge to an optimum in reasonable time. Pandey et al. [28] presented a detailed and comprehensive comparison of various premature convergence handling approaches on the basis of their merits, demerits and other important factors

978-1-4799-8433-6/15/$/31.00©2015IEEE

The primary interest of this study is to present the working of various selection techniques, their merits and demerits with their comparisons. The underlying motivation is to identify new approaches to address the issue of premature convergence. The Section II is given to understand the previous work done by researchers to utilize various selection techniques. The purpose of Section III is to discuss the working of different selection techniques studied for this paper. The following selection techniques like Proportionate Roulette wheel selection, Linear Ranking Selection, Exponential Ranking Selection and Tournament Selection are discussed. Section IV shows the comparison between these selection techniques on the basis of the results found by the researchers in past. The conclusion is drawn in Section V.

II. Previous work on selection techniques

As discussed in the previous section, GA is heuristic search technique and found variety of applications in various domains. The selection of parent population is one of the main steps in GA performance. Many researchers had studied GA considering different selection techniques. The performance of GA largely depends on the convergence rate and total number of generations to reach to the global optima.

Zhong et al[22] presented a comparative study of the selection techniques. For the study purpose, roulette wheel selection and tournament selection were considered. To conduct the experiment, seven different test functions were chosen and obtained results were collected for the analysis purpose. It was found that SGA outperformed with tournament selection whereas SGA with roulette wheel selection showed average performance.

Julstrom [23] considered the computational time as a parameter to examine the efficiency of the two types of rank-based selection probabilities namely linear ranking and exponential ranking probabilities which was compared with two types of tournament selection i.e. 2-tournament selection without replacement and k-tournament selection with replacement. Based on the analysis it -was found that tournament selection is preferred over rank-based selection. The reason is repeated tournament selection is much faster than arranging the population in order to assign rank-based probabilities.

Mashohor et al. [24] compared three selection approaches such as deterministic, tournament and roulette wheel to evaluate the performance of PCB inspection system. It was found that deterministic selection showed the ability to come out with the maximum fitness, computational time and accuracy in lowest number of generations. This is followed by roulette wheel and tournament selection.

III. Selection Strategies

The selection phase determines which individuals are chosen for mating (reproduction) and how many offspring each selected individual produces. The main principle of selection strategy is "the better is an individual; the higher is its chance of being a parent [3]." It is the process that determines which solutions are to be preserved and allowed to reproduce and which ones deserve to die out. The primary objective of the selection operator is to emphasize upon the good solutions and eliminate the bad solutions in a population, while keeping the population size constant.

Selection introduces the influence of the fitness function to the genetic-algorithm optimization process. Selection must utilize the fitness of a given individual, since fitness is the measure of the "goodness" of an individual. However, selection cannot be based solely on choosing the best individual, because the best individual may not be very close to the optimal solution. Instead, some chance that relatively unfit individuals are selected must be preserved, to ensure that genes carried by these unfit individuals are not "lost" prematurely from the population. In general, selection involves a mechanism relating an individual's fitness to the average fitness of the population. A number of selection strategies have been developed and utilized for genetic algorithm optimization.

A. *Proportionate Roulette Wheel Selection*

Schema Theorem, which was given by Holland [27], assumes that the chromosomes are selected based on their probabilities that are proportional to their fitness value. Its selecting principle is similar to that of a roulette wheel.

The probability of selection of a sector in a roulette wheel is proportional to the magnitude of the central angle of the sector. Similarly in Genetic Algorithm, whole population in partitioned on the wheel and each sector represents an individual. The proportion of individual's fitness to the total fitness values of whole population decides the probability of selection of that individual in the next generation. This consequently decides the area occupied by that individual on the wheel [22].

Following are the steps for Roulette Wheel selection:

- Calculate the sum of the fitness values of every individual in the population.
- Calculate the fitness value of each individual and their probability of selection by dividing individual chromosome's fitness by the sum of fitness values of whole population.
- Partition the roulette wheel into sectors according to the probabilities calculated in the second step.
- Spin the wheel 'n' number of times. When the roulette stops, the sector on which the pointer points corresponds to the individual being selected.

The probability of selection of an individual aj:

$$ps(a_i) = \frac{f(a_i)}{\sum_{i=1}^{n} f(a_j)}; j = 1, 2, ..., n \tag{1}$$

Where 'n' is the size of population, f(ai) is the fitness value of the individual ai

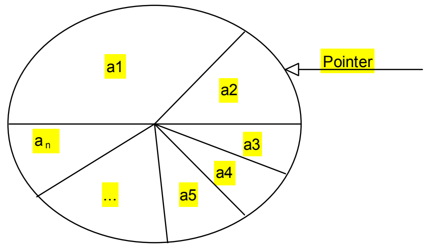

Fig. 2. Roulette wheel selection [22]

B. Linear Ranking Selection

Ranking Selection in Genetic Algorithm was introduced by Baker to eliminate the disadvantages of Proportionate selection [25]. In Linear Ranking selection method, individuals are first sorted according to their fitness value and then the ranks are assigned to them. Best individual gets rank 'N' and the worst one gets rank '1'. The selection probability is then assigned linearly to the individuals according to their ranks.

$$P_i = \frac{1}{N}(n^- + (n^+ - n^-)\frac{i-1}{N-1}); i \in \{1,\ldots,N\} \quad (2)$$

Where P_i is the Selection Probability of ith individual, $\frac{n^-}{N}$ is probability of selection of the worst individual and $\frac{n^+}{N}$ is probability of selection of the best individual.

Each individual gets a different rank even if their probabilities are same.

Ranking process is of two steps. In first step we sort the population and in second, we assign the ranks in order corresponding to proportionate Selection. Best sorting algorithm takes O (n log n) time complexity. Therefore, the time complexity of Ranking selection would be O(n log n) + time of selection (i.e. somewhere between O(n) and O(n2)) [5].

C. Exponential Ranking Selection

This technique is different from linear ranking selection technique in a way that the probabilities here are exponentially weighted. The base of the exponent is c, where 0 < c < 1.

$$p_i = \frac{c^{N-i}}{\sum_{j=1}^{N} c^{N-j}}; i \in \{1,\ldots,N\} \quad (3)$$

The sum $\sum_{j=1}^{N} c^{N-j}$ normalizes the probabilities to ensure that $\sum_{i=1}^{N} p_i = 1$

The algorithm of both Linear Ranking and Exponential Ranking is the same. The only difference is in the calculation of selection probabilities [3]. In this also the best individual is assigned rank 'N', and the worst one is assigned '1'.

D. Tournament Selection

Due to the efficiency and ease of implementation, Tournament selection is the most popular selection technique of Genetic Algorithm. In Tournament Selection, 'n' individuals are chosen at random from the entire population. These individuals compete against each other. The individual with the highest fitness value wins and gets selected for further processing of Genetic Algorithm [5].

The number of individual taking part in every tournament is referred as tournament size. Most commonly used tournament size is 2 i.e. in Binary Tournament Selection.

There are several advantages of Tournament selection strategy that makes it more efficient than other techniques. These include less time complexity i.e. O(n)[5] , easy parallel implementation, low vulnerability to takeover by dominant individuals, and no requirement for fitness scaling or sorting [8][12].

The process of Tournament Selection is shown below (Fig 3)

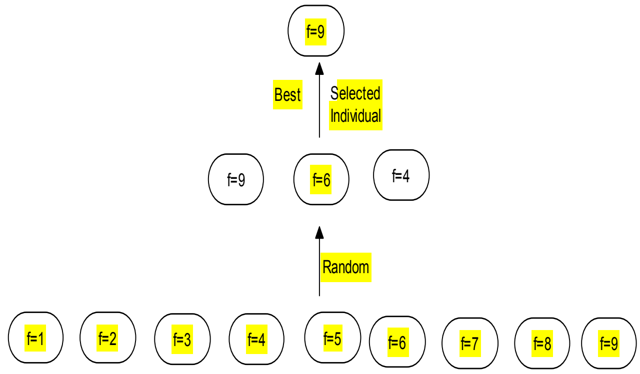

Fig. 3. Tournament Selection Mechanism [15]

In the above figure, Tournament size is three, which means three individuals compete against each other in one tournament.

The larger the tournament size, the greater is the probability of loss of diversity [12][14]. There are two reasons for loss of diversity. Either the individual did not get the opportunity to be selected (because of random selection), or the individual didn't get selected in the intermediate population because they lost some tournament.

IV. COMPARISON BETWEEN DIFFERENT SELECTION STRATEGIES

TABLE I. COMPARISON BETWEEN DIFFERENT SELECTION STRATEGIES

Citation	Publication Year	Author and year	Factor Considered for study	Selection Technique	Observation
[5]	1991	Goldberg, David E., and Kalyanmoy Deb	Expected fitness ratio and time to converge.	Compared proportional, ranking, tournament and Genitor (steady state) selection.	Ranking and tournament selection is better than proportional selection in terms of maintaining

					steady pressure toward convergence. Binary tournament selection is recommended over linear ranking selection because of its more efficient time complexity.
[23]	1999	Julstrom, Bryant A	Probability of ith chromosome	Studied the computing time efficiency of 2 types of Rank based selection (linear and exponential),and compared them with Tournament selection method.	Tournament selection is better than rank based selection, since repeated tournament selection is faster than sorting the population and then assigning rank based probabilities to them.
[22]	2005	Zhong, J., Hu, X., Gu, M., & Zhang	Using 7 general test functions.	Compared Proportionate roulette wheel selection method and Tournament selection method. (tournament size=6)	Tournament selection is more efficient than Proportionate roulette wheel selection in terms of Convergence rate.
[24]	2005	Mashohor, Syamsiah, Jonathan R. Evans, and Tughrul Arslan	Maximum fitness Rate of accuracy Computation time	Compared Deterministic, Tournament and Roulette wheel selection methods.	The ability to reach the maximum fitness with lowest no. of generations is highest in deterministic. This is followed by roulette wheel and then tournament selection.

V. CONCLUSION

This paper presents a comprehensive study of various selection techniques within GA. The comparison and observations are presented in tabulated form for a quick review. During the study it was observed that tournament selection perform better than other techniques in terms of convergence rate and time complexity.

REFERENCES

[1] Sivaraj, R., and T. Ravichandran. "A REVIEW OF SELECTION METHODS IN GENETIC ALGORITHM." International Journal of Engineering Science & Technology 3.5 (2011).

[2] Chen, Yang, et al. "Multiple sequence alignment based on genetic algorithms with reserve selection." Networking, Sensing and Control, 2008. ICNSC 2008. IEEE International Conference on. IEEE, 2008.

[3] Blickle, Tobias, and Lothar Thiele. "A comparison of selection schemes used in genetic algorithms." (1995).

[4] Johnson, J. Michael, and V. Rahmat-Samii. "Genetic algorithms in engineering electromagnetics." Antennas and Propagation Magazine, IEEE 39.4 (1997): 7-21.

[5] Goldberg, David E., and Kalyanmoy Deb. "A comparative analysis of selection schemes used in genetic algorithms." Urbana 51 (1991): 61801-2996.

[6] MAGALHÃES-MENDES, J. O. R. G. E. "An Experimental Study of Genetic Crossover Operators for the Job Shop Scheduling Problem." Proceedings of 3rd International Conference on Mathematical Models for Engineering Science. 2012.

[7] Mühlenbein, Heinz, and Dirk Schlierkamp-Voosen. "Predictive models for the breeder genetic algorithm i. continuous parameter optimization." Evolutionary computation 1.1 (1993): 25-49.

[8] Voráč, Jiří, Ivo Vondrák, and Karel Vlček. "SCHOOL TIMETABLE GENERATING USING GENETIC ALGORITHM." VSB–Technical University of Ostrava Czech Republic.(1995)

[9] Colorni, Alberto, Marco Dorigo, and Vittorio Maniezzo. "A genetic algorithm to solve the timetable problem." Politecnico di Milano, Milan, Italy TR (1992): 90-060.

[10] Abramson, David, and J. Abela. A parallel genetic algorithm for solving the school timetabling problem. Division of Information Technology, CSIRO, 1991.

[11] Ali, Elnima, and Elgasim Elamin. "A proposed genetic algorithm selection method." (2006).

[12] Datta, Dilip, Kalyanmoy Deb, and Carlos M. Fonseca. "Solving class timetabling problem of IIT Kanpur using multi-objective evolutionary algorithm."KanGAL, Report 2006006 (2006): 1-10.

[13] Gotlieb, C.C., The construction of class-teacher timetables. In Proceedings of IFIP Congress, North-Holland Pub. Co., Amsterdam, 1962, 73-77

[14] Datta, Dilip, et al. "Multi-objective evolutionary algorithm for land-use management problem." International Journal of Computational Intelligence Research 3.4 (2007): 1-24.

[15] Noraini, Mohd Razali, and John Geraghty. "Genetic algorithm performance with different selection strategies in solving TSP." (2011).

[16] Pandey, Hari Mohan, Anurag Dixit, and Deepti Mehrotra. "Genetic algorithms: concepts, issues and a case study of grammar induction." Proceedings of the CUBE International Information Technology Conference. ACM, 2012.

[17] Holland, John H. Adaptation in natural and artificial systems: an introductory analysis with applications to biology, control, and artificial intelligence. MIT press, 1992.

[18] Goldberg, David E. "Genetic algorithms in search, optimization, and machine learning." (1989).

[19] Pandey, H.M. 2010. "Context Free Grammar Induction Library Using Genetic Algorithms". International Conference on Computer & Communication Technology ICCCT'10.

[20] N.S. Choubey, M.U. Kharat and Hari Mohan Pandey, "Developing Genetic Algorithm Library Using Java for CFG Induction", In International Journal of Advancement in Technology. 2011.

[21] Al Jadaan, Omar, Lakishmi Rajamani, and C. R. Rao. "IMPROVED SELECTION OPERATOR FOR GA." Journal of Theoretical & Applied Information Technology 4.4 (2008).

[22] Zhong, Jinghui, et al. "Comparison of performance between different selection strategies on simple genetic algorithms." Computational Intelligence for Modelling, Control and Automation, 2005 and

International Conference on Intelligent Agents, Web Technologies and Internet Commerce, International Conference on. Vol. 2. IEEE, 2005.

[23] Julstrom, Bryant A. "It's all the same to me: Revisiting rank-based probabilities and tournaments." Evolutionary Computation, 1999. CEC 99. Proceedings of the 1999 Congress on. Vol. 2. IEEE, 1999.

[24] Mashohor, Syamsiah, Jonathan R. Evans, and Tughrul Arslan. "Elitist selection schemes for genetic algorithm based printed circuit board inspection system."Evolutionary Computation, 2005. The 2005 IEEE Congress on. Vol. 2. IEEE, 2005.

[25] Baker, James Edward. "Adaptive selection methods for genetic algorithms."Proceedings of an International Conference on Genetic Algorithms and their applications. 1985.

[26] Baker, James Edward. "An analysis of the effects of selection in genetic algorithms." (1989).

[27] John, Holland. "Holland, Adaptation in natural and artificial systems." (1992).

[28] Pandey, Hari Mohan, Ankit Choudhary, and Deepti Mehrotra. "A Comparative Review of Approaches to Prevent Premature Convergence in GA." Applied Soft Computing (2014).