International Journal of Engineering, Science and Mathematics
Vol. 6 Issue 3, July 2017,
ISSN: 2320-0294 Impact Factor: 6.765
Journal Homepage: http://www.ijesm.co.in, Email: ijesmj@gmail.com
Double-Blind Peer Reviewed Refereed Open Access International Journal - Included in the International Serial Directories Indexed & Listed at: Ulrich's Periodicals Directory ©, U.S.A., Open J-Gage as well as in Cabell's Directories of Publishing Opportunities, U.S.A

# Study of the various selection techniques in Genetic Algorithms

[1]Saneh Lata Yadav, Assistant Professor, Department of Computer Science and Engineering,, K.R. Mangalam University, Gurgaon
[2]Asha Sohal, Associate Professor, Department of Computer Science and Engineering, K.R. Mangalam University, Gurgaon

**Information Technology**

***Abstract-*** *This paper considers a number of selection schemes commonly used in modern genetic algorithms. The categorization of the various selection techniques and their advantages and disadvantages are considered in this paper.*

***Keywords*** *– Convergence, Diversity, Fitness proportionate, Genetic algorithm, Roulette wheel selection.*

## 1. INTRODUCTION

Genetic algorithms are the methods of optimization which are analogous to the natural evolution and natural genetics. The theory of genetic algorithms was originally developed by John Holland in 1960 and was fully developed in his book "Adaptation in Natural and Artificial Systems ", published in 1975 [1]. The goal of their research has been to design artificial systems software that retains the important mechanisms of natural systems. Today there are many applications of genetic algorithms in science, economy and research development.

In genetic algorithms, initially a population of individuals is randomly generated. This population represents the candidate solutions to the problem that is to be optimized. The individuals in this population are evaluated through adaptation function or fitness function. A selection mechanism is then used to select individuals to be used as parents to those of the next generation. After that crossover and mutation operators are applied. The next generation is finally formed. This process is repeated until the convergence criterion is reached.

The genetic algorithms start working from Initialization phase, where a population of candidate solutions to the problem is generated in a random manner. Fitness of each candidate solution is evaluated using fitness function in order to determine their utility and a fitness value is associated with each candidate solution. Then only those solutions which possess high probability i.e. the fit individuals are filtered out using selection operator. Depending on the nature of the optimization problem a selection technique is used to select the intended parents. The selected candidates are the ones which are likely to reproduce. This list of selected candidates forms the mating pool. Then in reproduction phase, the offspring (new candidate solutions) are produced by using the crossover operator. Here, the set of new solutions is the new population. Now, if some termination criteria is met the evolution process stops, otherwise continues to evaluate.

The three most important aspects of a genetic algorithm are: objective function, the genetic representation and the genetic operators. Once these three have been defined, the genetic

International Journal of Engineering, Science and Mathematics
Vol. 6 Issue 3, July 2017,
ISSN: 2320-0294 Impact Factor: 6.765
Journal Homepage: http://www.ijesm.co.in, Email: ijesmj@gmail.com
Double-Blind Peer Reviewed Refereed Open Access International Journal - Included in the International Serial Directories Indexed & Listed at: Ulrich's Periodicals Directory ©, U.S.A., Open J-Gage as well as in Cabell's Directories of Publishing Opportunities, U.S.A

algorithm works fairly well. [2] The objective functions are subject to optimization. The various genetic operators are selection, crossover and mutation.

## 2. RELATED WORK

Genetic algorithms involve the following three operators

1. **Selection**: - This operator selects chromosomes in the population for reproduction. The fitter the chromosome, the more times it is likely to be selected to reproduce. [**3**]
2. **Crossover: -** This operator randomly chooses a point and exchanges the subsequence before and after that point between two chromosomes to create two offspring. These strings are randomly selected from the mating pool and the crossover point is also selected randomly.
3. **Mutation: -** This operator is used for occasional introduction of new features in the solution strings of the population. The objective of this operator is to maintain the diversity.

**Disadvantage**

- It can result in premature convergence as less fit candidates are truncated and they are not given the opportunity to reproduce.
- It doesn't preserve the diversity because only the best individuals always contribute.
- It cannot be used in the situation where the each individual is not better than the other; it means the set is optimal.
- It requires sorting technique, if the population size is very large. It may add to the cost.

## 3. DISCUSSION

Selection Techniques are used to select the individuals for reproduction. The motive of the selection operator is to propagate the good individuals to the next generation while maintain the population size. It is expected that each time a fit individual is selected to be a part of the mating pool. The various selections techniques used in genetic algorithm are as following:

### 3.1 Truncation Selection

Truncation selection is the simplest selection technique. As the name implies it uses truncation for selection. In this technique individuals are first sorted in accordance to their fitness. And the individuals above threshold value are selected for being a part of the mating pool and the rest of the individuals are truncated. But it may be the case that the number of individuals selected is not equal to the population size. So as to maintain the population size, copies of the selected individuals are created. For example, if the population size is 80 we may select 25% fittest individuals from this population, i.e. 20 individuals are selected but the population size is 80. Now in order to maintain the population size we have to four times copy the selected fittest individuals.

International Journal of Engineering, Science and Mathematics
Vol. 6 Issue 3, July 2017,
ISSN: 2320-0294 Impact Factor: 6.765
Journal Homepage: http://www.ijesm.co.in, Email: ijesmj@gmail.com
Double-Blind Peer Reviewed Refereed Open Access International Journal - Included in the International Serial Directories Indexed & Listed at: Ulrich's Periodicals Directory ©, U.S.A., Open J-Gage as well as in Cabell's Directories of Publishing Opportunities, U.S.A

**Advantages**

- This is an easy selection strategy and it is simpler to implement.

Threshold takes values ranging from 50%-10%. Individuals below the truncation threshold do not produce offspring. [**4**]

### 3.2 Roulette wheel selection

It is the most common and simplest fitness-proportionate selection technique. Each individual of the population is allocated a section of an imaginary roulette wheel, which is proportionate to its fitness. The fittest candidate has the biggest section of the wheel and the weakest candidate has the smallest. The wheel is then spun *n* number of times, where *n* is the population size and every time the individual associated with the winning section is selected. The characteristic of this selection method is the fact that it gives to each individual *x* of the current population a probability *p(x)* of being selected, proportional to its fitness *f* (x) Where *n* denotes the population size.

| No. | Chromosome | Fitness f(x) | %age |
|---|---|---|---|
| 1 | -1.0159 3.7160 | 14.8407 | 22.45 |
| 2 | -1.9395 -2.3770 | 9.4118 | 14.24 |
| 3 | 1.5749 0.3961 | 2.6372 | 3.990 |
| 4 | -1.1077 0.5270 | 1.5047 | 2.276 |
| 5 | -3.8980 4.7438 | 37.6980 | 57.03 |
| | | **66.0925** | **100.0 0** |

**Table**: **Percentage Fitness**

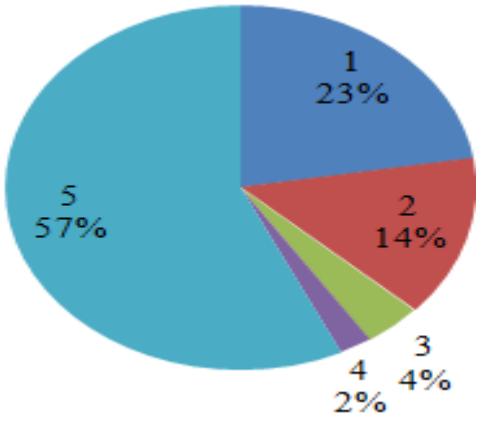


**Figure 2 Roulette Wheel**

International Journal of Engineering, Science and Mathematics
Vol. 6 Issue 3, July 2017,
ISSN: 2320-0294 Impact Factor: 6.765
Journal Homepage: http://www.ijesm.co.in, Email: ijesmj@gmail.com
Double-Blind Peer Reviewed Refereed Open Access International Journal - Included in the International Serial Directories Indexed & Listed at: Ulrich's Periodicals Directory ©, U.S.A., Open J-Gage as well as in Cabell's Directories of Publishing Opportunities, U.S.A

In Figure 2 each individual is allocated area proportional to its fitness. Where, the individuals in this population are values within the -5.12 to 5.12 and are being used to optimize a DeJong Function, *f(x)*. The fitness values are the function of *x*.

**Modified Percentage Fitness** 0%

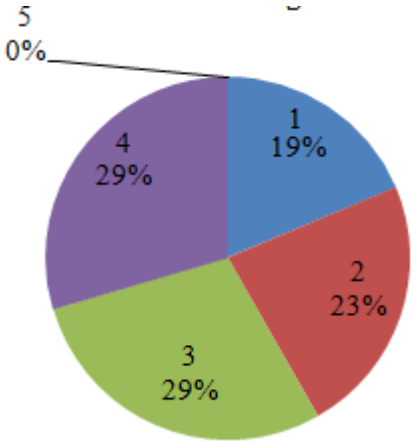


**Figure 3 Modified Roulette Wheel**

As the optimization function is a minimization problem, individual No. 4 is the fittest and No. 5 is the weakest. To increase the number of chances a fit individual being selected, some mechanism is applied. And as a result of it, the strongest individual a value of 2.27% and the weakest 57.03% now has percentage fitness of 29% and 0% respectively.

**Advantages**

- It is possible that one or more individuals are selected multiple times.
- It is simpler and easy to implement technique.
- This selection technique is free from bias, as a fair chance of selection is given to each individual thereby preserving the diversity

**Disadvantages**

- The risk of premature convergence to a local optimum, due to the possible presence of an individual which is dominant and is selected as a parent every time.
- The fittest individual occupies the largest circumference on the wheel; hence the chances of its selection increases and the other individuals have few chances.

International Journal of Engineering, Science and Mathematics
Vol. 6 Issue 3, July 2017,
ISSN: 2320-0294 Impact Factor: 6.765
Journal Homepage: http://www.ijesm.co.in, Email: ijesmj@gmail.com
Double-Blind Peer Reviewed Refereed Open Access International Journal - Included in the International Serial Directories Indexed & Listed at: Ulrich's Periodicals Directory ©, U.S.A., Open J-Gage as well as in Cabell's Directories of Publishing Opportunities, U.S.A

### 3.3 Stochastic Universal Sampling

Stochastic Universal Sampling is similar to Roulette Wheel Selection, but unlike Roulette Wheel Selection in this technique there are *n* numbers of pointers, where *n* represents the number of individuals to be selected. These n pointers are equally distant.

The distance between two pointers is *1/n.* The selection is done by generating random number only once. The first pointer of the comb like assembly of the equidistant pointers moves to the random number generated and rest of the pointers moves accordingly. In this way *n,* required number of individuals are selected in one go.

**Advantages**

- It is a fast method and free from bias.
- Choosing individuals from equally spaced intervals gives fair chance of reproduction to the weak individuals also.

### 3.4 Rank selection

Rank Selection is similar to fitness-proportionate Roulette Wheel Selection except the fact that the selection probability in this technique is proportional to relative fitness rather than absolute fitness. In this technique the probability of selection of an individual is a function of rank rather than fitness. It doesn't make any difference whether the fittest candidate is ten times fitter than the next fittest or 0.1% fitter. In every case the selection probabilities remains the same; all that matters is the ranking relative to other individuals [**5**]. Rank selection is a two-step process. First the list of individuals must be sorted, and next the assignment of ranks to the individuals.

**Advantages**

- It preserves the diversity.
- Free from bias.

**Disadvantages**

- This method can lead to slower convergence, because there is not much difference between the best individuals.
- Computationally expensive as a sorting is also required

### 3.5 Tournament selection

In tournament selection, every individual in the population is paired at random with another. The fitness values of each pair are compared. The fitter individual of the pair moves on to the next round, while the other is disqualified. This continues until there are a number of winners equal to the desired number of parents. Then this last group of winners is paired as the parents for new individuals [**6**]. In this selection technique the selection pressure is

controlled by the size of the tournament. If each time the size of the tournament is equal to the size of the population, winner will always be the same.

**Advantages**

- It can be implemented very efficiently as no sorting of the population is required.
- It preserves the diversity.
- It can be implemented in parallel.

**Disadvantages**

- If the tournament size is large, each time the same individual turns out to be the winner; in this case diversity may be lost.

### 3.6 Boltzmann Selection

Boltzmann selection is based upon the concept of annealing; heating a metal and then gradually letting it cool down. Analogous to the heat or temperature in the annealing of metals is the selection pressure in the Boltzmann Selection. In this selection technique initially the selection pressure is kept low so that more and more individuals get selected, and then regularly increasing the selection pressure.

For instance, let $f_m$ represents the fitness of the best individual and $f_x$ is the fitness of the next individual(x) then fitness are compared, if it found that the fitness of next individual is better than the currently available best individual, then the best is replaced, but if it is not the case then the x is selected with a probability given by the Boltzmann probability. Here $T=T_0 (1-\alpha)^k$ , $\alpha$ is a constant in the range [0,1] and $T_0$ ranges between [5,100] and k=(1+ 100*g/G), g and G represents the generation number and total number of generations respectively. The convergence criterion is approaching 0 for the value of T which symbolizes that the global optimum is reached. [7]

**Advantages**

- Global optima is reached, it doesn't struck to local optima.

### 3.7 Elitism

The idea of elitism is to preserve the best individual of a population. Due to crossover and mutation operator the best individuals are lost. The aim of elitism is to not only preserve the best individual but at the same time allowing it to participate in reproduction so as to propagate the best characteristics to the next generations.

## 4. CONCLUSION

In this paper, some of the well-known selection methods in genetic algorithms are studied. Each selection technique has some advantages and disadvantages over the other; no single technique is the best. Depending upon the nature of the problem at hand, a selection technique should be used. If there is not a single selection technique that solves the purpose then two or more techniques can be blended to get the problem solved.

## REFERENCES

[1] John H. Holland, Adaptation in Natural and Artificial Systems: The University of Michigan Press, 1975.

[2] Sailing Matthew Wall. Introduction to Genetic algorithms.[Online].

[3] Melanie Mitchell, An Introduction to Genetic Algorithms. London England: MIT Press, 1996.

[4] Hartmut Pohlheim. (2006) Genetic and Evolutionary Algorithms Toolbox for use with Matlab.

[5] Daniel W. Dyer. (2010) Evolutionary Computation in Java. [Online].

[6] S. M. Shah, Mahesh Panchal Chetan Chudasama, "Comparison of Parents Selection Methods of Genetic Algorithm for TSP," in International Journal of Computer Applications, 2011, pp. 85-87.

[7] S.N.Deepa S.N.Sivanandam, Introduction to Genetic Algorithms. New York.

[8] Thomas Weise, Global Optimization Algorithms- Theory and Application, 2007. Springer, 2008.