# A Novel and Efficient Selection Method in Genetic Algorithm

Smit Anand
Institute of Technical
Education and Research
Bhubaneswar, Odisha, India

Nishat Afreen
Birsa Institute of Technology
Sindri, Jharkhand, India

Shama Yazdani
Institute of Technical
Education and Research
Bhubaneswar, Odisha, India

## ABSTRACT

The performance of a Genetic Algorithm (GA) is inspired by a number of factors: the choice of the selection method the type of crossover operator, the rate of mutation, population size etc. GA allows a diverse population to evolve under a specific selection scheme to fitter population. Therefore, the choice of the selection method plays a very important role in the maximization of the fitness function of the evolved population. In this paper, a novel selection method called "Alternis" has been proposed. This study emphasizes on the comparison among the different selection methods used in GAs and the proposed method and evaluate their performance. Results of this study highlight the significant differences among the various selection schemes. The influence of the various selection methods on the performance of genetic algorithm can be estimated to assist the preference of a selection method. The aim of this paper is to propose a selection method which gives best overall performance in a widely diverse population.

## General Terms

Artificial Intelligence, Algorithms.

## Keywords

Genetic algorithm, Chromosomes, Crossover, Mutation, Fitness function.

## 1. INTRODUCTION

Genetic Algorithm inspired by Darwin's theory of evolution is a way to mimic the natural process of reproduction to solve various problems. It was proposed by John Holland in 1975. It is a machine learning model which is derived from the process of evolution in nature. GAs simulate those processes in natural populations which are essential to evolution [7]. The behavior of GA is analogous to the natural behavior of living species where they compete which each other to survive and to find a suitable mate. Therefore, the successive generations will be the combination of attributes of the fittest individuals in the population. The combination of good characteristics of the parents can sometimes produce an offspring that can be fitter than fittest parents.

GA has started forming its roots in late 1950s and early 1960s when biologists started developing biological evolution system through programs. In 1962, scientists like G. Friedman, Bledsoe and others had independently developed algorithms following biological evolution system. Late in 1962, Holland was the first to introduce the method of crossover. Genetic Algorithm work has mainly started after the publication of the book "Adaptation in natural and artificial system". After that Genetic Algorithm is being applied to abstract mathematics, pattern recognition and classification, and structural optimization.

A basic genetic algorithm comprises of three genetic operators-Selection, Crossover and mutation. The genetic algorithm uses these operators on the population to produce the next generation**.** It typically starts with a set of solutions called population; each individual in this population is represented by a chromosome [3]. Each chromosome is evaluated by its fitness [3], which is computed by the fitness function defined. The fitness of each individual serves as an important factor in its selection to form the next generation or the next possible set of solutions. This process continues until an optimal solution is reached. The fig 1 illustrates the basic flowchart of genetic algorithm.
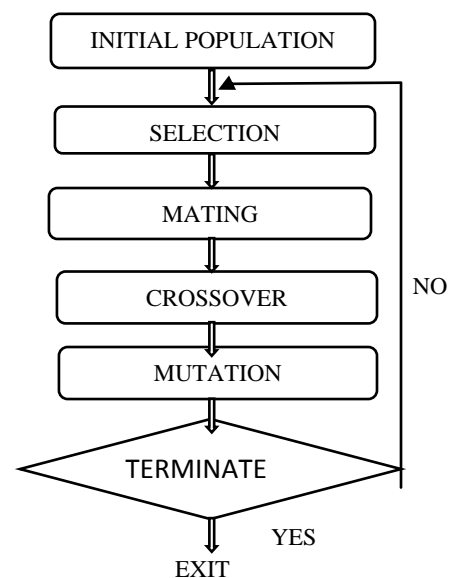


**Fig. 1. Basic Flowchart of Genetic Algorithm**

## 2. BACKGROUND
### 2.1 Selection Pressure

The selection pressure is the degree to which better individuals are favored [2].

The higher the selection pressure, the better will be the new generation. If the selection pressure is very low, the convergence rate will be slow, thus resulting in GA consuming longer time to find the optimal solution. A very high selection pressure, on the other hand, can lead to a premature convergence of GA which may result in an incorrect optimal solution.

### 2.2 Fitness Function

Fitness function is the measure to evaluate how good an individual is relative to the other individuals. Higher fitness value indicates better individuals.

## 2.3 Chromosome

The chromosome is encoded information about the properties of an individual.

## 3. SELECTION

The approach of selection mechanism is to select individuals from the current population to create better individuals or better solutions. Fitter the individual, better is the chance of its selection as selection procedure is based upon the strategy of survival of the fittest. There are various ways to perform selection of the individuals to create offspring. Some common ways of performing selection are: roulette wheel selection, tournament selection, rank selection, steady state selection, elitism etc. Each method has its own advantages and disadvantages.

This paper discusses roulette wheel selection, tournament selection and rank selection methods. Advantages and disadvantages of each have been discussed in detail. We have also proposed a new approach of performing selection. This paper shows a comparative study of the above mentioned approaches and the proposed approach. Each algorithm has been implemented with the help of a computer program in java and the results are shown for efficient comparison.

## 3.1 Roulette Wheel Selection

This is the simplest technique of selection, often called as stochastic sampling with replacement [1]. Each individual, based on fitness, is assigned a contiguous segment on the number line. A random number is generated and the segment where the number lies is selected as one of the parents. This process is repeated until all the parents are selected. This is analogous to a wheel where each individual is assigned a section in the wheel sized in proportion to individual's relative fitness. The fitter the individual is, the better probability it has to form the mating pool.

```
Declare mating pool size P
for each x in 1 to P do
for each y in 1 to N do
    Calculate percent fitness for each individual
end for
for each y in 1 to N do
     Calculate cumulative fitness for each individual
end for
Generate random number R.
for each y in 1 to N do
   if  R <= cumulative fitness of y
        return y
end for
Select y as a parent.
end for
```

### 3.1.1 Advantages
- The roulette wheel selection method is simplest and easiest to implement.
- It takes very less amount of time compared to other selection strategies.

### 3.1.2 Disadvantages
- It leads to premature convergence of the GA, thus, sometimes resulting in incorrect optimal solution.

## 3.2 Rank Selection

This method overcomes the scaling problem of roulette wheel selection method where the fitness differs very much. The fittest individual occupies the largest section on the wheel, thus further minimizing the probability of least fit individual. The rank selection method ranks the individual in increasing order of fitness from 1 to N. Therefore, even the worst individual in the population has rank 1. The probability that an individual is selected is directly proportional to its ranking, rather than its fitness values. Thus, if an individual is 10 times or 5 times better than the next individual, its selection probability will not differ that much. Each individual is weighted by its rank rather than its absolute fitness.

```
Sort population based on fitness value.
Declare mating pool size P
for each x in 1 to P do
Assign rank from 1 to N
for each y in 1 to N do
    Calculate cumulative rank for each individual
end for
Generate random number R>1 and < cumulative rank.
for each z in 1 to N do
   if R > cumulative[x] then
        R = R - cumulative[x]
   else
        Corresponding individual in the original unsorted
        array is the selected parent
end for
end for
```

**Fig. 2. Algorithm for Rank Selection**

### 3.2.1 Advantages
- The rank selection method maintains a constant pressure in the evolutionary search.
- This selection strategy is more robust than other proportional fitness mechanism.

### 3.2.2 Disadvantages
- It leads to a slower convergence as the fitness of the individuals does not differ much from each other.

## 3.3 Tournament Selection

Tournament selection is the most efficient selection method. In this scheme, a tournament is conducted among a few randomly chosen individuals from the population. The one with the best fitness is chosen as the winner and joins the mating pool. This process is repeated until all the parents are chosen. The number of individuals competing with each other in the tournament is referred to as the tournament size. The tournament size can vary but it is commonly set to 2, also called as binary tournament. The selection pressure can be adjusted by varying the tournament size. If the tournament size is large relative to the size of the population, it may lead to loss of diversity as the weaker individuals will have lesser probability of being the winner.

```
Declare tournament size N < population size
Declare mating pool size P
for each x in 1 to P do
Select N individuals randomly
Select fittest among N
end for
```

**Fig. 3. Algorithm for Rank Selection**

### 3.3.1 Advantages
- The selection pressure can be adjusted by changing the tournament size.
- It is well suited for parallel architecture.

## 3.4 Comparative study of existing selection algorithms

### 3.4.1 Why tournament is better than roulette wheel and rank selection method?

In tournament selection having tournament size two, we select two random individuals and choose the fittest individual among them. The worst case in this scenario will be when the two individuals selected for tournament are the ones with the least two fitness values. The winner of the tournament will be the second worst individual. Therefore, even in the worst case scenario, the second least fit individual will join the mating pool.

In tournament selection, the probability that the selected individual will have the fitness level which is at least equal to the second worst is 1.

While for Roulette wheel and Rank selection method, the probability of such case will be as follows:

In roulette wheel selection algorithm, we sum up the fitness value of all individuals and then assign each individual a slice on the wheel based on their percentage value in the total sum of fitness.

If the total sum of fitness is 'Φ' and fitness value of worst fit individual if 'μ', then the probability of selection of worst fit individual is

$$P = \mu/\Phi \qquad (1)$$

Therefore, the probability that the selected individual will have the fitness level which is at least second is

$$P = 1 - \mu/\Phi \qquad (2)$$

From the above equations Roulette wheel selection method will be equivalent to Tournament selection method if

Case 1: The difference between the fitness value of the worst fit and second worst fit individual is very large.

Case 2: The total sum of fitness value of individuals is much greater than the fitness value of the worst fit individual.

In Rank selection method, we sort the individuals and assign them rank from worst fit to best fit i.e. the individual with least fitness value will have rank 1 and the individual with highest fitness value will have rank N.

If the number of individuals is 'Ω', then the probability of selection of worst fit individual is

$$P = 1/((\Omega(\Omega-1))/2) \qquad (3)$$

So, the probability that the selected individual will have the fitness level which is at least second worst is

$$P = 1 - 1/((\Omega(\Omega-1))/2) \qquad (4)$$

Thus, Rank selection method becomes equivalent to Tournament selection method if the number of individuals is too large.

## 3.5 Alternis selection: proposed method

It is observed that the above described, commonly used selection mechanisms provide opportunity for each individual to get selected except the tournament selection method where the worst individual never gets a chance to join the mating pool. It has also been perceived that the mating pool may comprise of all the poorly fit individuals in the worst case. To take care of these two scenarios, a novel selection method has

been proposed, named 'Alternis'. It aims to overcome the drawback of tournament selection by providing a window to the worst individual to be a successful mating candidate. It further eradicates the probability of the mating pool to comprise of only the weak individuals.

In the proposed algorithm, the chromosomes are selected in such a way that the resulting generation has very less probability of deterioration. The population is sorted in descending order of the fitness value. As the name suggests, the individuals are then arranged in alternating fashion of their fitness value i.e. best fit will be placed right to worst fit and left to second worst fit which in turn will be placed left to second best fit. This process is repeated until are the individuals are arranged in this fashion.

**Case 1: N is Odd**

| $1^{st}$ worst | $1^{st}$ best | $2^{nd}$ worst | $2^{nd}$ best | … | $\frac{N+1}{2}$ th worst | $\frac{N+1}{2}$ th best |
|---|---|---|---|---|---|---|

**Case 2: N is even**

| $1^{st}$ worst | $1^{st}$ best | $2^{nd}$ worst | $2^{nd}$ best | …. | $\frac{N}{2}$ th best/worst |
|---|---|---|---|---|---|

**Fig. 4. Arrangement of chromosome in Alternis**

A random individual is chosen and put in the mating pool. Then the individual to its right as well as its left are chosen and put in the mating pool. This process is continued till the mating pool is filled.

If size of mating pool is X and number of individual is N, then the two cases shall be considered:

Case 1: X is even

The range of the random individual R is

$$\frac{X}{2} - 1 \leq R \leq N - \frac{X}{2}$$

i.e. $\frac{X}{2} - 1$ Individuals on left and $\frac{X}{2}$ on the right are excluded while selection of the first random individual R for mating pool. The others are chosen alternatively, one from right and one from left and so on till the mating pool is filled.

Case 2: X is odd

The range of the random individual R is

$$\frac{X-1}{2} \leq R \leq N - \frac{X-1}{2}$$

i.e. $\frac{X-1}{2}$ individuals on left and $\frac{X-1}{2}$ on the right are excluded while selection of the first random individual R for mating pool. The others are chosen alternatively, one from right and one from left and so on till the mating pool is filled.

```
Declare number of individuals N
Declare mating pool size X
Arrange the individuals in alternating fashion i.e.:
(worst_fit)(best_fit)(2nd_worst_fit)(2nd_best_fit).......

if X is even then
choose a random number R from X/2 − 1 to   N − X/2
Then choose X/2 − 1 individual from left and X/2 from right.
else
```

| |
|---|
| choose a random number R from $\frac{X-1}{2}$ to $N - \frac{X-1}{2}$ <br> Then choose $\frac{X-1}{2}$ from left and $\frac{X-1}{2}$ from right. <br> end if |

**Fig. 5. Algorithm for Alternis Selection**

*3.5.1 Advantages*

- The proposed selection method can create more diverse population.

- It is more flexible than proportional fitness selection methods.

- This is an unbiased selection strategy.

- It prevents premature convergence of the GA.

# 4. EXPERIMENTAL RESULTS

The discussed selection algorithms were implemented with a help of a computer program in Java. The program takes into consideration a population of 30 individuals, each composed of a single chromosome. Each chromosome has three characteristics: Anger, Shape and Color. Each characteristic is represented in the chromosome by its level from 1 to 20 with one being least and 20 being highest. Each characteristic in the chromosome is represented by 5 bits in binary making the complete chromosome of 15 bits.

The fitness function is calculated using simple multiplication of the respective levels of characteristics i.e.

Fitness(x) = anger(x)*shape(x)*color(x)

The mating pool size is fixed to three. The initial population goes through the basic steps of GA i.e. selection, crossover and mutation. The crossover technique used is two-point crossover. And mutation is carried out by flipping a bit randomly when the binary bits for any characteristic represents level greater than 20, as the maximum level should not exceed 20.

The program tests each selection algorithm individually for the initial population and is terminated when use of either selection algorithm results in population fulfilling the terminating condition and then compares the results achieved by all other selection algorithm and thus helps compare the performance of each selection algorithm when initial population, crossover technique and mutation technique is kept constant.

In all the experiments, two point crossover was used. If required, each child had to further undergo mutation by changing the bits of the chromosomes from 0 to 1 or vice versa. A comparison between the proposed selection method and the other selection methods are made and the results are presented with the help of the following graphs and tables:
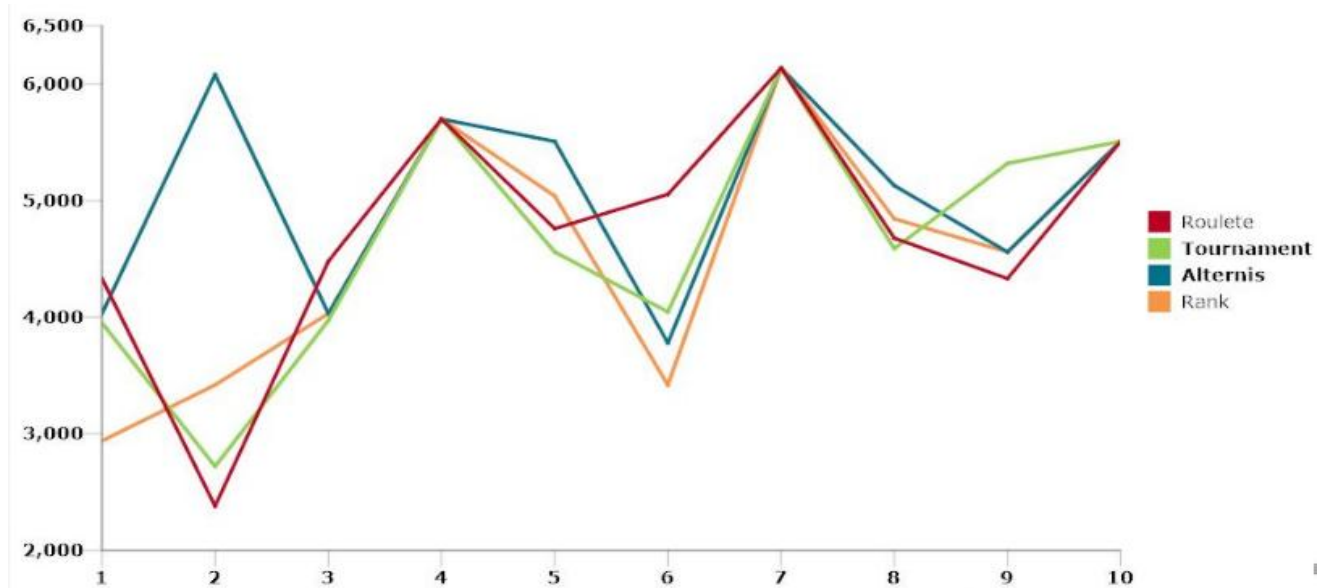


**Fig. 6. Graph for highest fitness comparision**

**Table:1 Highest Individual Fitness**

| RUN NO | 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|---|
| ROULETTE | 4332 | 2380 | 4480 | 5700 | 4760 |
| TOUNAMENT | 3952 | 2720 | 3974 | 5700 | 4560 |
| ALTERNIS | 4032 | 6080 | 4032 | 5700 | 5508 |
| RANK | 2940 | 3420 | 4032 | 5700 | 5040 |

**Table:2 Highest Individual Fitness**

| RUN NO | 6 | 7 | 8 | 9 | 10 |
|---|---|---|---|---|---|
| ROULETTE | 5054 | 6137 | 4680 | 4332 | 5508 |
| TOUNAMENT | 4046 | 6137 | 4590 | 5320 | 5508 |
| ALTERNIS | 3780 | 6137 | 5130 | 4560 | 5508 |
| RANK | 3420 | 6137 | 4845 | 4560 | 5508 |
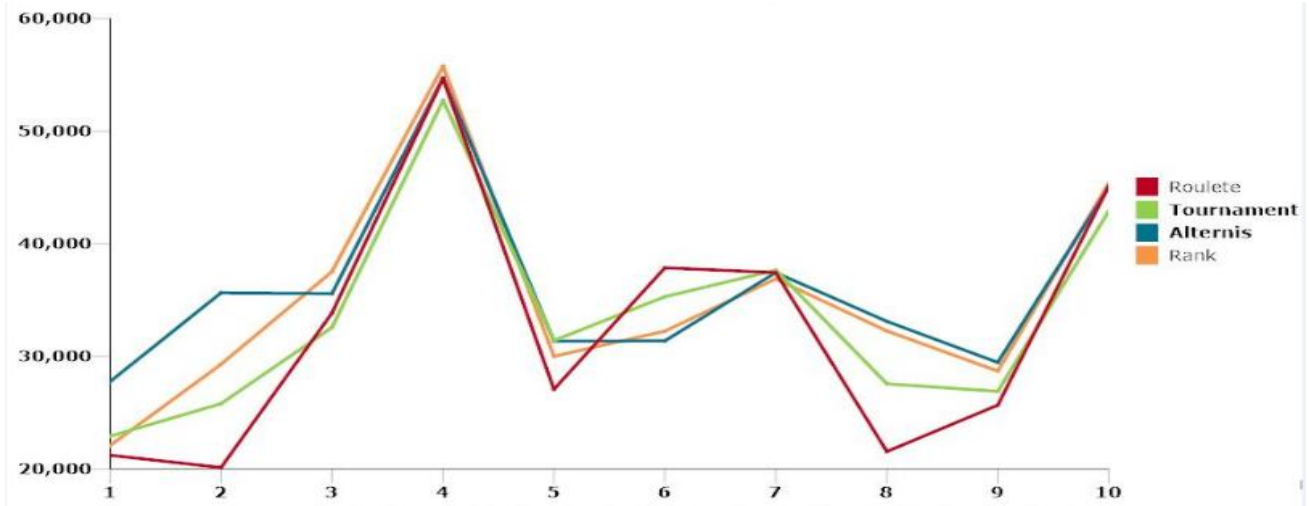
**Fig:7 Graph for cumulative fitness comparision**

**Table 3: Highest Cumulative Fitness**

| RUN NO | 1 | 2 | 3 | 4 | 5 |
|--------|------|------|------|------|------|
| ROULETTE | 21219 | 20138 | 33843 | 54667 | 27085 |
| TOUNAMENT | 22908 | 25808 | 32536 | 52729 | 31408 |
| ALTERNIS | 27716 | 35643 | 35577 | 54739 | 31355 |
| RANK | 22099 | 29294 | 37530 | 55765 | 30003 |

**Table 4:Highest Cumulative Fitness**

| RUN NO | 6 | 7 | 8 | 9 | 10 |
|--------|------|------|------|------|------|
| ROULETTE | 37867 | 37449 | 21558 | 25672 | 45103 |
| TOUNAMENT | 35293 | 37653 | 27576 | 26906 | 42883 |
| ALTERNIS | 31373 | 37449 | 33102 | 29471 | 45005 |
| RANK | 32241 | 36879 | 32263 | 28702 | 45383 |

Table I and II shows that out of 10 outcomes, Alternis is in the top two in 9. This shows that Alternis can be implemented as a very efficient selection mechanism. The experimental results prove that the new generation is always fitter than the previous generation. Since the individuals in the proposed algorithm are arranged in alternating fashion of most and least fit, therefore it enables the selection scheme to select varied individuals. Therefore, the mating pool never comprises of all the least fit individuals.

## 4.1 Efficiency Evaluation of Selection Methods

### 4.1.1 Highest Individual Fitness:
In this range, out of the 10 recorded readings, "Alternis" scores highest or combined highest in 6 readings while "Tournament" scoring highest in 4, "Rank" in 3 and "Roulette" in 6.

This concludes probability of occurrence of highest or combined highest fitness value in "Alternis" as 0.6, in "Tournament" as 0.4, in "Rank" as 0.3 and in "Roulette" as 0.6.

### 4.1.2 Highest Cumulative Fitness
According to the experimental reading of this section, "Alternis" is among highest or combined highest in 5 out of 10 results with, "Tournament" in 2, "Rank" in 3 and "Roulette" in 1.

This shows that there is 0.5 probability that "Alternis" will be among the highest for cumulative fitness value with "Tournament" as 0.2, "Rank" as 0.3 and "Roulette" as 0.1.
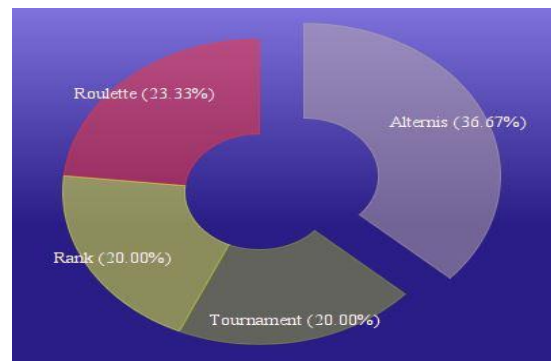


**Fig:8 Pie Chart depicting the probability of best occurences in both individual and cumulative Fitness**

## 5. CONCLUSION
In this paper, a novel and efficient selection method in GA is introduced. The performance of each selection method in terms of highest individual fitness and cumulative fitness is weighed through the implementation of a program in Java. A comparison between the results of the proposed method and other selection methods are made. It is observed that on an average, the proposed selection method produces better results than other selection methods. The fitness of the new generation is the key factor of a GA. Mostly, the proposed selection method outperforms all the selection methods to produce a healthier generation. The experimental results and the comparative study of the various selection methods and the proposed selection method clearly depicts that the proposed selection method has a greater potential to improve the performance of GA as a whole.

# 6. REFERENCES

[1] Baker, J. E.: Reducing Bias and Inefficiency in the Selection Algorithm. in [Grefenstette, J. J.: Proceedings of the Second International Conference on Genetic Algorithms and their Application, Hillsdale, New Jersey, USA: Lawrence Erlbaum Associates, 1987.], pp. 14-21, 1987.

[2] Goldberg, D. E., and Miller, B. L., "Genetic Algorithms, Tournament Selection, and the Effects of Noise", Complex Systems, 9 (1995) 193- 212

[3] Kumar, R, and Jyotishree, "Blending Roulette Wheel Selection & Rank Selection in Genetic Algorithms", International Journal of Machine Learning and Computing, Vol. 2, No. 4, August 2012

[4] J.E. Baker, Adaptive selection methods for genetic algorithms, in Proceedings of the First International Conference on Genetic Algorithms, Lawrence Erlballm Associates, Hillsdale, NJ, pp. 101-111, 1985Tavel, P. 2007 Modeling and Simulation Design. AK Peters Ltd.

[5] Goldberg, D. E., and Deb, K., "A Comparative Analysis of Selection Schemes used in Genetic Algorithms," Foundations of Genetic Algorithms, 1 (1991) 69-93 (also TCGA Report 90007.

[6] Blickle, T, and Thiele L, "A Comparison of Selection Schemes used in Genetic Algorithms", TIK-Report, Nr. 11, December 1995, Version 2 (2. Edition)

[7] Beasley D, Bull D. R. and Martin R.R, "An overview of Genetic Algorithms: Part 1, Fundamentals", University Computing, 1993, 15(2) 58 69.

[8] O. A. Jadaan, L. Rajamani, C. R. Rao, "Improved Selection Operator for GA," Journal of Theoretical and Applied Information Technology, 2005

[9] B. A. Julstrom, It‟s All the Same to Me: Revisiting Rank-Based Probabilities and Tournaments, Department of Computer Science, St. Cloud State University, 1999

[10] J. Zhong, X. Hu, M. Gu, J. Zhang, "Comparison of Performance between Different Selection Strategies on Simple Genetic Algorithms," Proceeding of the International Conference on Computational Intelligence for Modelling, Control and automation, and International Conference of Intelligent Agents, Web Technologies and Internet Commerce, 2005.

[11] Handbook of Evolutionary Computation, IOP Publishing Ltd. and Oxford University Press, 1997

[12] K. S. Goh, A. Lim, B. Rodrigues, Sexual Selection for Genetic Algorithms, Artificial Intelligence Review 19: 123 – 152, Kluwer Academic Publishers, 2003

[13] J. H. Holland, Adaptation in natural and artificial systems, The University of Michigan press, 1975

[14] P.G. Bachhouse, A.F. Fotheringham, and G. Allan, A comparison of a genetic algorithm with an experimental design technique in the optimization of a production process, Journal of Operational Research Society, Vol. 48, pp. 247-254, 1997

[15] T. Bii.ck, Selective pressure in evolutionary algorithms: a characterization of selection mechanisms, in Proceedings of the First IEEE Conference on Evolutionary Computation, IEEE Press, Orlando, FL, pp. 57-62, 1994