8
GENETIC ALGORITHM

8.1 HISTORY

Darvin's theory of evolution by natural selection, or survival of fittest could be used as an optimization tool for engineering problems. This pioneering idea eventually led to the creation of a new discipline in engineering, known as *Evolutionary computation*. The concept behind evolutionary computation was to evolve a population (more then one) of candidate solutions to a given problem, using operators stimulated by natural genetic variations and natural selection.

The implementation of evolutionary algorithms can be classified into three strongly related, but independently developed methodologies:

- Evolutionary Strategies.
- Evolutionary Programming.
- Genetic Algorithm.

In this chapter, we are only concerned with the family of GAs and their application as an optimization tool to the problem of multi-user detection in CDMA based wireless communication systems.

8.2 INRODUCTION

The origin of GAs can be traced back to the 1960, when Holland and his students undertook the task of studying the phenomenon of adaptation, as it occurs in nature and then imported these adaptive mechanisms into artificial systems.

The basic approach of a GA employed in optimizing problem defined by an objective function is simple. The flowchart of a GA is shown in figure 8.1. First an initial population consisting of P number of so-called *individuals* is created in the 'initialization' block, where P is known as the population size. Each individual represents a legitimate solution to given optimization problem. An individual can be considered as a vector consisting of the decision variables to be optimized, as shown in figure 8.2. Here, we will regard the right most decision variable is referred to as the lth decision variable. Traditionally, the individuals in a GA population take the form of binary bit vectors. Hence if the decision variables to be optimized are not binary in nature, they have to be discretized and encoded to a bit vector, analogously to analogue –to-digital conversion. This initial population of individuals is usually generated randomly, although it does not necessarily have to be random specifically. If *explicit a priori* knowledge concerning the optimum vector is available, then this knowledge can also be used to generate the individuals of the initial population, in order to bias and expedite the search.

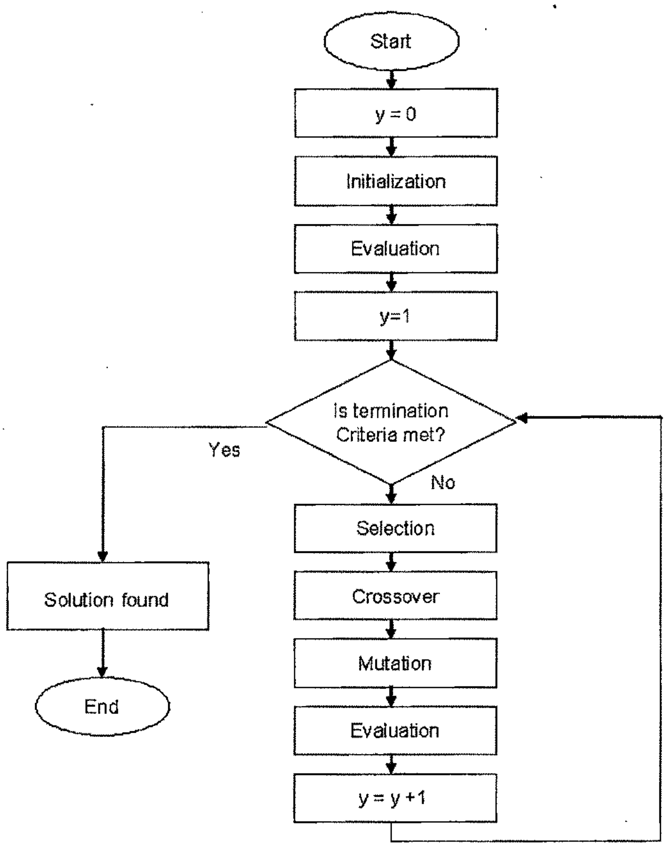


Figure 8.1 : GA Flowchart

1^{st} decision variable	2^{nd} decision variable	--	Lth decision variable

Figure 8.2 : Individual of length-l.

Associated with each individual in the population, there is a figure of merit or more commonly known in GA parlance as the *fitness* value. The fitness value is evaluated by substituting the candidate solution represented by the individual under consideration into the objective function, as indicated by the 'Evaluation' block of figure 8.1. Individuals having the T number of highest fitness value are then placed in a so called *matting pool,* where 2 T P. using a kind of natural selection scheme together with the genetically inspired operators of *crossover* and *mutation,* the individuals in the mating pool are then are evolved to a new population, as depicted in figure 8.3.

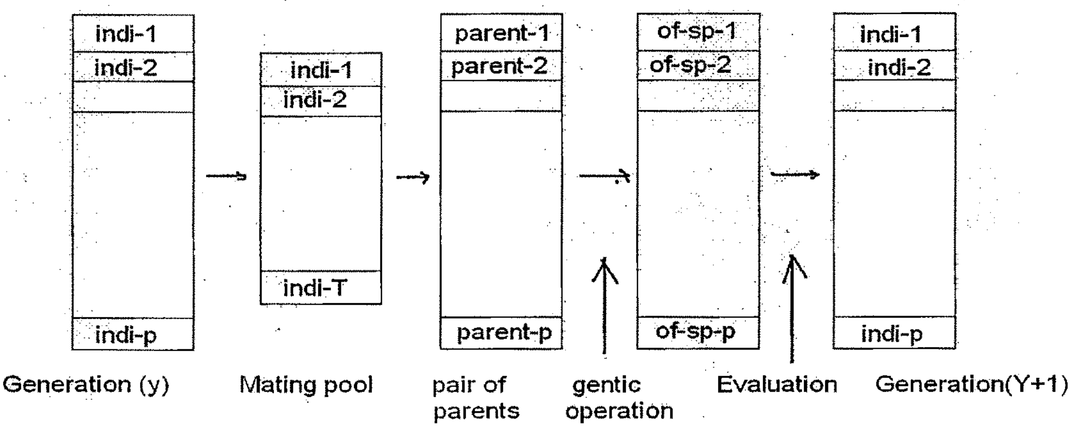


Figure 8.3: An example of a GA operation during a single cycle or generation.

Based solely on the fitness values of these individuals in the current mating pool, the *selection* process choose those individuals in the mating pool are then used by the crossover and mutation operations, in order to generate new individuals, which will form the new population for the next iteration. The selection process is invoked for improving the average fitness value of the population by giving the individuals of higher fitness values a higher probability to be reproduced in the new population. Hence it focuses the search on the promising regions in the search space, which might contain the optimum solution. Numerous selection schemes have been proposed in the GA literature. However, the selection does not alter the individuals. If the optimum solution is to be found, new individuals must be generated. The task of generating new individuals, using the individuals chosen by the selection process is accomplished by the crossover operation.

8.3 GENETIC ALGORITHM AT WORK

Consider an optimization problem, where the objective function is given by,

$$V(b). = 2\, b\, y - b\, R\, b^T \qquad (1)$$

Assuming that the values of y and R available, the goal of this optimization process is to find the decision variable vector b, which consist of l = 5 bits, that maximizes the objective function of equation 1.

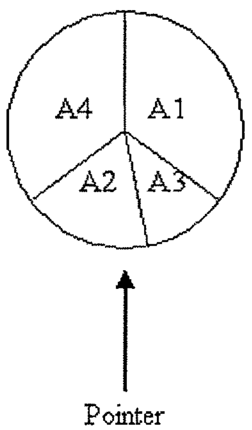
An example of the initial population consisting of P = 4 individuals, where the associated fitness values are evaluated according to the objective function given by equation following equation.

Individuals		Fitness (Fi)	Mapped fitness $F_i' = F_i + 10$	Selection probability $P_i = F_i' /\ {}_j^T F_i'$
A1	1 1 1 -1 1	9.06	19.06	0.3985
A2	1 1 -1 -1 -1	-1.056	8.944	0.1870
A3	-1 1 -1 -1 -1	-6.368	3.632	0.07594
A4	1 -1 -1 1 1	6.192	16.192	0.3385
				${}_j^P F_i' = 47.828$

Table 8.1.A maping function for Fitness proportionate slelection.

A common selection method used in GA's is the so called *fitness proportionate selection,* in which the probability of selection *Pi* of the i^{th} individual is equal to its fitness value Fi divided by the total fitness value of the mating pool. This method requires the fitness value to be positive for all combinations because a negative fitness value would yield a negative probability of selection. Hence a mapping function must be invoked, in order to ensure that the fitness values for all combinations become positive. A simple mapping function used is indicated in Table 8.1. Summing the mapped fitness values Fi' over all four individuals, we obtained total mapped fitness values of 47.828 and an average mapped fitness value of 11.957 for the initial population of table 8.1. The probability of selection for each individual is calculated in proportion to their individual fitness and the values are listed in table 8.1. We can see that individuals having higher fitness values are allocated a higher probability selection. Note that the probability of selection Pi of an individual is defined with respect to the average fitness of the current population. Hence, the probability selection of the same individual would be different in a different population.

A simple method of implementing the fitness proportionate selection is the so called *roulette wheel sampling* where each individual is allocated a slice of circular roulette wheel proportional in area to the individual's probability of selection. An example of a roulette wheel is shown in figure 8.4 for the population of table 8.1.



Individual	Probability of Selection
A1	0.3985
A2	0.1872
A3	0.07594
A4	0.3385

Figure 8.4 Roulette wheel sampling for selection

When roulette wheel is spun and the pointer comes to standstill on one of the wedge-shaped slices, the new corresponding individual will be selected as a parent. Each spin of the roulette wheel yields a new parent. Hence we can see here that by using this implementation, individuals with higher probability of selection have a higher chance of being selected as parents. Once pair of parents is selected, the cross over operation is then applied to this pair of parents as depicted in figure 8.3.

8.4 CROSSOVER

The crossover operation is process in which arbitrary decision variable is exchanged between a pair of selected parents, mimicking the biological recombination process between two single chromosome organisms. Hence the crossover operation creates two new individuals, known as offspring in GA parlance which have a high probability of having better fitness value than their parents. In order to generate P number of new offspring, P/2 numbers of crossover operations are required. A new pair of parents is selected from the mating pool for each crossover operation. The newly created offspring will form the basis of the new population.

Following is the example of cross over between the pairs of selected individuals of table 8.1.

Parents	Crossover Mask	Offspring
1 1 -1 1 1 -1 1 -1	0 0 1 0 1 0 1 1	1 1 1 1 -1 -1 1 1
1 -1 1 1 -1 1 -1 1		1 -1 -1 1 1 1 1 -1

Table 8.2 Cross over of the selected individuals of table 8.1

8.5 MUTATION

During the mutation operation, each decision variable in the offspring is perturbed, with a probability of *Pm*, by either a predetermined or a random value. This allows new areas in the search space to be explored. The mutation probability of a decision variable is usually very small, in the region of 0.1-0.01. However, the mutation operation is necessary in a GA, in order to prevent the phenomenon of so called premature convergence. Premature convergence refers to the loss of population diversity before the optimum solution has been found.

The size of the population in a GA is a major factor in determining the accuracy of convergence. As the population size increases, the GA has a better chance of finding the global optimum solution. Apart from the population size, a GA's performance will also depend subsequently on other factors, such as the choice of the selection method, the type of genetic operations employed, the parameter setting, for example, the value of T and *Pm* as well as the particular iteration termination criterion.

8.6 SELECTION

It is important that individuals having higher fitness values in a given mating pool must be given a better chance of reproducing in the subsequent generation, then the lower fitness individuals in the same mating pool. The task of choosing these individuals for reproduction is performed by the selection process. Numerous selection schemes have been proposed in the GA literature. We will highlight some of the more commonly used methods below. Here we will assume that the selection process is invoked in the mating pool, which contains T number of individuals associated with the highest fitness values in a given population.

8.6.1 Fitness-proportionate Selection

In *Fitness-proportionate Selection,* as invoked in example of table 2.1, the probability of selection Pi of the i^{th} individual is defined as:

$$Pi = \quad Fi \; / \quad {}_{j}^{T} Fj' \qquad (2)$$

Where, Fi is the fitness value associated with the i^{th} individual. However the fitness proportionate scheme has several deficiencies which may lead to a premature convergence.

8.6.2 Sigma scaling

The *sigma scaling* selection scheme was proposed in order to make the GA less sensitive to premature convergence. Under this scheme, the probability of selection of the i^{th} individual is a function of several variables, namely that of its fitness value, the mating pool mean fitness and the mating pool fitness standard deviation as given by .

$$\begin{aligned} Pi &= (1/T)\,(1 + (Fi-F')/2\sigma) \quad \text{if } \sigma \;\; 0 \qquad \text{or} \qquad (3) \\ &= 1/T \qquad\qquad\qquad\qquad \text{if } \sigma = 0, \end{aligned}$$

Where Fi is the fitness value of the i^{th} individual, F' is the mean fitness of the mating pool and σ is the standard deviation of the mating pool's fitness value. If the standard deviation of the mating pool fitness is high, individuals having high fitness values will not be assigned a high probability of selection. Hence the individuals having lower fitness values are given a fair chance of reproducing. On the other hand, if the individuals in the mating pool are similar, resulting in a low standard deviation, then the individuals exhibiting higher fitness values will be assigned a higher probability of selection.

8.6.3 Linear Ranking Selection

The *linear ranking selection* is an alternative method of preventing premature convergence. According to this method, the individuals in the mating pool are ranked according to their associated fitness values, such that the rank T is assigned to the individual associated with the highest fitness value in the mating pool, while rank first is assigned to the individual exhibiting the lowest fitness value in the mating pool. Similarly, the remaining individuals in the mating pool are ranked accordingly. The i^{th} individual will then be assigned its probability of selection Pi, based on its specific ranking *ranki*, in the mating pool, as given by.

$$Pi = 1/T\,[\eta^- + (\eta^+ - \eta^-)\ (ranki - 1)/(T-1)] \qquad (4)$$

Where η^- / T is the probability of selection assigned to the individual associated with the lowest fitness value and η^+ / T the probability of selection assigned to the individual having the highest fitness value.

8.6.4 Tournament Selection

According to the *Tournament Selection* scheme, t number of individuals are chosen randomly from the mating pool, where $t < T$ is referred to as the tournament size. The individual associated with the highest fitness value out of these t preferred individuals will be selected as a parent. This process is repeated for another set of individuals, in order to form a pair of parents for the cross over operation.

8.7 TERMINATION CRITERION

The exact structure of the search space is often unknown in optimization problems. Hence in search algorithms, with the exception of an exhaustive search, it is typically infeasible to ensure that the optimum solution can be found. There are numerous ways of determining the termination criterion for GA's.

The GA assisted search can be terminated, if there are no further improvements in the maximum fitness value after several consecutive generations. In this case, the time required for GA to reach a decision is uncertain. On the other hand, if the structure of the search space is time-invariant, then it is possible to set a threshold, such

that the GA assisted search is terminated, once the fitness value of an individual is found to exceed this threshold. Unfortunately neither of these termination criteria can be applied to GA assisted CDMA multi-user detection, since typically a fixed implementation complexity is also required and also the search space is time variant due to noise and fading imposed by the transmission channel.

Hence in our application, we will terminate the GA assisted search at the Yth generation and the individual associated with the highest fitness assisted search at the Yth generation and the individual associated with the highest fitness value at this point will be detected solution. By specifying the exact number of generations, the computational complexity of GA can be decreased.

8.8 SUMMARY

In this chapter we have presented a brief overview of GA's. Specifically, we introduced terminologies and procedures of GA with an example. We revived some of the more commonly used GA based operations such as cross over, mutation and selection schemes as well as implementation strategies. There are many ways of implementing GA, using different combinations of selection schemes, cross-over and mutation operations. There is no definite theoretical justification as to which combination is the optimum performance, since different combinations work best for different problems. The best way of identifying the specific combination of operations that most suitable for the problem at hand is to critically appraise and adapt these combinations to the problem.