International Journal of Sciences: Basic and Applied Research (IJSBAR)

ISSN 2307-4531
(Print & Online)

http://gssrr.org/index.php?journal=JournalOfBasicAndApplied

Optimal One Day International Cricket Team Selection by Genetic Algorithm

Isuru Kumarasiri[a*], Sanjeewa Perera[b]

[a]*Department of Information Technology, Sri Lanka Institute of Information Technology, Malabe 10115, Sri Lanka*

[b]*Research and Development Center for Mathematical Modeling, Department of Mathematics,University of Colombo, Colombo 00700, Sri Lanka*

[a]*Email: isurukuma@gmail.com*

[b]*Email: mscfm@sci.cmb.ac.lk*

Abstract

Cricket is one of the oldest and most popular games in the world. At present it is played professionally and none professionally around the world. Even though the origin of cricket is unknown, there are evidences that it has been played in 16th century. Since then, this remarkable game has developed massively and the first international test cricket match was played in 1877 between Australia and England. Generally there are three types of international cricket matches, which are Test match cricket, one day international cricket and 20-20 cricket. Since the popularity of cricket, it has an immense financial importance. Therefore, winning more matches consistently will generate more financial benefits for the country. Selecting the most suitable cricket team for the given scenario is very vital as far as winning matches are concerned. This system contains various individual characteristics such as players' batting average, bowling average, Number of wickets and so on. This study describes how to select the optimal one day international squad of Sri Lanka, which consists of 15 players out of 30 players, for the world cup 2015 which is played in Australia. By using Genetic Algorithm method and MATLAB 7 software a system has been constructed to implement the results.

Keywords: Genetic Algorithm; Optimal cricket team; Optimization.

---

* Corresponding author.

1. Introduction

Cricket is a team sport which is played between two teams. Since it is a popular sport in around the world, cricket has a financial worthiness. Wining more matches will pump more money for the winning countries. Selecting the exact combination for a particular match or a series is always a challenge. There are number of diverse factors which should be considered when selecting a squad of 15 players. At times, selection committee might involve with human errors and biased decisions. Therefore, it is necessary to have a system without human interference so that by considering all the possible factors, optimal cricket team is able to be selected. Even though some people have attempted to select optimal cricket team by using genetic algorithm before, there are some factors which can also be considered to improve the results such as players' performances in the given country and win loss ratio and so on."The authors in [4] have considered limited number of factors". The results are tested using winning Pakistan world cup team of 1992.

2. Genetic Algorithm

Genetic algorithm was discovered to illustrate some of the processes of nature and it is applied in different areas such as, biology, engineering, economics, chemistry, mathematics and etc... "The authors in [1] explain that Genetic Algorithm process was first described by American scientist, John Holland in early 1960's". Many complex real world optimization problems can be solved by using this method. The main principal of the concept of genetic algorithm is survival of the fittest. In genetic algorithm process it frequently alters the population of individual solutions. At each step genetic algorithm picks individuals at random from the existing population and uses them to create the next generation. After considerable number of generations, population converges towards the optimal solution. Generally genetic algorithm process begins with an initial set of random solutions and it is also called the initial population. Each set of random solutions of the population is called a chromosome. In most cases a chromosome is a string of numbers and these numbers alter throughout iterations. During the each iteration, chromosomes are evaluated by using a defined fitness function. Then the next population is selected after using crossover operator and mutation operator. In two points cross over method, two strings from the existing population are randomly selected. Then two cut points are also selected randomly and everything in between two cut points are swapped to construct modified chromosomes. In mutation operator a random number of a randomly selected string is changed. These operators are used to modify the existing population and fitness values of each chromosome are calculated after applying the above mentioned operators. The process of selection, cross over and mutation is the main three steps in genetic algorithm. This process continuous until it meets the defined stopping criteria and after considerable number of iterations this method converges to the optimal solution.

3. Methodology

First all 30 players in Sri Lankan one day international cricket team pool is categorized as batsmen, fast bowlers, spin bowlers, all-rounders and wicket keepers. Then players' fitness values are evaluated using variables such as batting average, batting strike rate, bowling average, number of wickets per match, win loss ratio, Experience. Overall performances and performances in Australia are considered when fitness function is defined. Ranking

method is used to evaluate the fitness values. For an example, if batting average is the variable, a selected player's mean batting average is evaluated by considering his overall batting average and batting average in Australia.

Similarly, mean batting average is calculated for all 30 players and rank them from 1 to 30. If the mean batting average is high of a selected player then his rank value is a small number. But as far as mean bowing average is concerned lesser mean bowling averages have smaller ranking values and it is the opposite ranking method from other variables.

If a selected player has played more number of matches then that player has more experiences than others, due to that reason his rank for experience is a small number. Fitness function is defined using following variables.

$$\frac{Overall\ batting\ average+Batting\ average\ in\ Australia}{2} = mean\ batting\ average \quad (1)$$

$$\frac{Overall\ bowling\ average+Bowling\ average\ in\ Australia}{2} = mean\ bowling\ average \quad (2)$$

$$\frac{Overall\ batting\ strike\ rate+Batting\ strike\ rate\ in\ Australia}{2} = mean\ strike\ rate \quad (3)$$

$$\frac{Overall\ wickets\ per\ match+wicket\ per\ match\ in\ Australia}{2} = mean\ wickets\ per\ match \quad (4)$$

$$\frac{Overall\ win\ loss\ ratio+win\ loss\ ratio\ in\ Australia}{2} = mean\ win\ loss\ ratio \quad (5)$$

By using the rank values of mean batting average, mean bowling average, mean strike rate, mean wickets per match, mean win loss ratio and experience, fitness function is defined. It is the average of above mentioned rank variables. Suppose that the optimal cricket squad of 15 players contains 10 specialist batsmen, then that squad is not an appropriate cricket squad.

To win matches a team requires batsmen as well as bowlers and all-rounders. In other words if the optimal squad provides 15 blowers then there are no batsmen in the team. So the team is unbalanced as a result of that it's impossible to win matches using that team.

Therefore by considering past data following constraints were defined and by doing so the appropriate combinations can be selected for the squad.

$4 \leq$ Number of Batsmen ≤ 6, $1 \leq$Number of Spinners ≤ 3, $3 \leq$ Number of Fast bowlers ≤ 5, $1 \leq$ Number of All-rounders ≤ 4, $1 \leq$ Number of Wicket Keepers ≤ 3

3.1. Representation

As it is mentioned above, every player in the 30 players' pool is numbered from 1 to 30. Then each number represents a particular player as in table 1.

Table 1: List of players in Sri Lankan one day international cricket pool

Number	Name	Number	Name
1	Kusal Perera	16	Nuwan Kulasekara
2	Mahela Jayawardene	17	Shaminda Eranga
3	Kithuruwan Vithanage	18	Nuwan Pradeep
4	Lahiru Thirimanne	19	Dhammika Prasad
5	Upul Tharanga	20	Farveez Maharoof
6	Dimuth Karunaratne	21	Thisara Perera
7	Ramith Rambukwella	22	Angelo Mathews
8	Ajantha Mendis	23	Tillakaratne Dilshan
9	Rangana Herath	24	Ashan Priyanjan
10	Sachithra Senanayake	25	Jeevan Mendis
11	Tharindu Kaushal	26	Seekkuge Prasanna
12	Lakshan Sandakan	27	Dilruwan Perera
13	Lahiru Gamage	28	Kumar Sangakkara
14	Lasith Malinga	29	Dinesh Chandimal
15	Suranga Lakmal	30	Niroshan Dickwella

After numbering each player, 100 random teams are constructed where each team contains 15 players. Therefore a team is represented as a string of 15 bits. In this scenario a chromosome is a team of 15 players and it can be represented as a bit string of length 15 as following,

t_i = 1 3 23 21 4 5 7 29 18 17 5 12 16 22 6

$t_i$is the i^{th} team of the initial population. Then out of these 100 teams 20 teams are randomly selected as the initial population, which satisfies the above mentioned constraints.

So that i varies from 1 to 20.Subsequently the fitness values of each chromosome which is the sum of fitness of each individual player are calculated. Crossover method and Mutation method are the next steps to follow in genetic algorithm process.

3.2. Crossover method

For this problem multiple cross over method is used. This method randomly cuts two cut points and switches the middle part of two chromosomes to generate new chromosomes.

After switching the middle part, new chromosomes should also satisfy the constraint. Chromosomes are selected randomly and crossover probability for this problem is kept as 0.6.

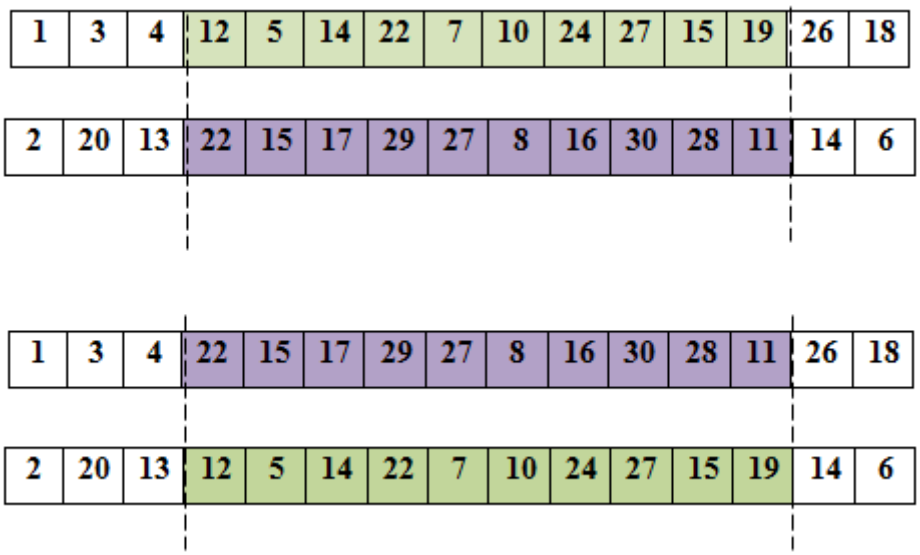

Figure 1: crossover method for chromosomes

Figure 1 demonstrates how to apply crossover method for two chromosomes of length 15. Third cell and the 14th cell have been selected as two cut points and every number in between those cells are switched to create new chromosomes.

3.3. Mutation method

In mutation method one cell is selected randomly and changes the number of that cell randomly to construct a new chromosome. After changing the randomly selected number, new chromosome should also satisfy the constraint. The mutation probability for this problem is kept at 0.05.

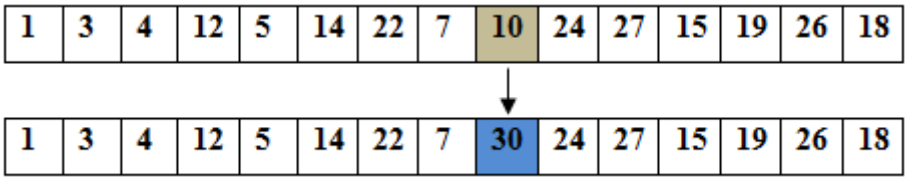

Figure 2: Mutation method for chromosomes

Figure 2 describes that 9th cell has been selected as the cell to change the number and before changing the value of that cell is 10 and after changing that value is 30. Consequently after applying mutation operator, fitness values of each team are calculated. Similarly this process continues until it meets the stopping criteria. The stopping criteria for this problem is, continue the above process until the difference of the fitness values of the best player and worst player will be less than 5. Here, 5 is the least value which gives the optimal solution. Selected team is also a better team where one player almost matches with the other 14 players in the squad and after considerable number of iterations it converges to the optimal team.

4. Results

4.1. MATLAB output of the Difference between highest and lowest fitness values vs. Iteration number

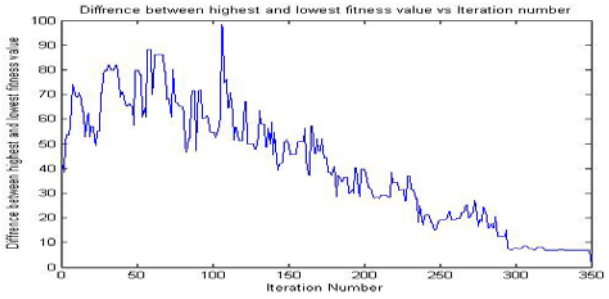

Figure 3: Difference between highest and lowest fitness values vs. Iteration number

Figure 3 is the MATLAB output of the Difference between highest and lowest fitness values vs. Iteration number which describes about the convergence of this process. After almost 350 iterations it converges to the optimal solution.

4.2. Selected Cricket team Squad

Out of 30 players pool 15 players are selected by using above mentioned process. MATLB 7 software is used to obtain results.

Table 2: Selected Sri Lankan Cricket Squad

Number	Name
1	Kusal Perera
2	Mahela Jayawardene
3	KithuruwanVithanage
4	Lahiru Thirimanne
5	Nuwan Kulasekara
6	Dhammika Prasad
7	Thisara Perera
8	Ajantha Mendis
9	Rangana Herath
10	Sachithra Senanayake
11	Angelo Mathews
12	Kumar Sangakkara
13	Tillakaratne Dilshan
14	Lasith Malinga
15	Suranga Lakmal

Table 2 shows the selected Sri Lankan one day international cricket squad, using genetic algorithm method for the world cup 2015, which is played in Australia. Figure 4 is obtained from MATLAB 7 and it explains that after 350 iterations this method gives the optimal cricket team and the fitness value of the team is 133.76. Fitness value of the squad is the sum of all fitness values of the optimal team players.

```
>> cric
   350

    19     9     2    23    14     8    22    10     3    15     1    28     4    16    21

  133.7600
```

Figure 4: Total fitness value of the optimal Sri Lankan cricket team

4.3. Testing Results

Results of the Optimal Sri Lankan cricket team can be tested using the world cup winning team of Pakistan in 1992, which was played in Australia. Following 15 players were selected as the optimal world cup team of Pakistan by Genetic Algorithm process. MATLB 7 software is used to obtain results.

Table 3: Selected Pakistan Cricket Squad

Number	Name
1	Aaqib Javed
2	Javed Miandad
3	Ijaz Ahamed
4	Zakir Khan
5	Saijad Akbar
6	Mushtaq Ahamed
7	Saeed Anwar
8	Saleem Malik
9	Moin Khan
10	Aamer Sohail
11	Imran Khan
12	Inzamam-ul-haq
13	Waqar Younis
14	Wasim Akram
15	Rameez Raja

Table 3 shows the selected Pakistan one day international cricket squad, using genetic algorithm method for the world cup 1992, which was played in Australia. According to Table 3 there are 13 players are in the original

Squad. Figure 7 is obtained from MATLAB 7 and it explains that after 222 iterations this method gives the optimal cricket squad and the fitness value of the squad is 128.0833.

```
>> cric
   222

    9    18    13    26     7    10     8    27    20    21     5     4    23    30    28

  128.0833
```

Figure 5: Total fitness value of the optimal Pakistan cricket team

5. Discussion

It was found that 12 players in this squad are in the original squad. Dhammika Prasad, Ajantha Mendis and Kithruwan vithanage are not in the original squad [vi]. But some selections from the cricket board are controversial decisions and the best way to check this by testing these results with a previous world cup. These results can be tested using 1992 world cup results, which was also held in Australia and the winning team was Pakistan. According to the players obtained for the Pakistan cricket team by Genetic Algorithm process, there are two players, who are not in the original squad. They are Iqbal Sikander and Wasim Haider. "The Data set in [6] clearly shows, by 1992 these two players had played only few matches." So the available international data was very limited. As a result of that, fitness values become very high. Therefore Iqbal Sikander and Wasim Haider have a small probability of getting selected for the optimal cricket team.

The system has considered all possible factors which can be explained by statistical data of the players'. But there are certain factors such as mentality of the players, confidence level, injuries, which can never be measured statistically. Even though the numbers of catches caught or dropped are available for an each player, that variable is not considered due to the reason that generally a player is not thrown away from the team by considering his fielding ability. If some injuries occur, then another set of players as the replacement, can also be selected by using this system. As it is mentioned in the introduction results can be improved by considering players' performances in Australia. Even though some people have tried this process before some important factors such as performances in playing country and selecting bowlers as fast bowlers as well as spin bowlers are missing in their implementation. (Omar &Verma, n.d.).This method can be improved by considering league and club level matches. Then there will be a probability for debutants and players like Iqbal Sikander, Wasim Haider for getting selected to the optimal cricket squad.

6. Conclusion

Using Genetic Algorithm method a system was constructed to select optimal cricket team of Sri Lanka for 2015 world cup. Testing part has been done by using Pakistan 1992 world cup winning squad and majority of players in the original Squad are in the Pakistan original squad. So it can be concluded that Sri Lankan squad using genetic algorithm will be a better squad than the original squad.

7. Recommendations

Genetic Algorithm process more often provides you the optimal solution and after some iteration it converges to the optimal solution. Therefore this method can be used in many other sports, such as football, baseball, basketball etc., to find the optimal team. Also this process can be used for any opponent cricket playing nation in order to find the best team against them.

References

[1]. M.D.I Song and M. Treiber. "An Introduction to Genetic Algorithms and Evolution Strategies." Technical Report, University of Waterloo, Canada, 2002.

[2]. N.V. Wankhade and S. P. Deshpande, (2012, Sept.). "Data Mining System for Selection of Best Basket Ball Team." International Journal of Engineering and Science. [On- line]. 1(3), pp.05-07. Available: http://www.researchinventy.com/papers [Jan. 1,2015].

[3]. S.S. Satya and S. Jamal. "Appling genetic Algorithm to select an optimal cricket team." In Proc. Proceeding of international conference on advances in computing, communication and control.2009, pp. 43-47.

[4]. S.N. Omkar and R. Verma, "CRICKET TEAM SELECTION USING GENETIC ALGORITHM." In Proc. International Congress on Sports Dynamics.2003, pp. 295-300.

[5]. "Genetic Algorithm." http://in.mathworks.com/discovery/genetic-algorithm.html, [Jan. 01,2015].

[6]. " player list : Sri Lanka." http://www.espncricinfo.com/, [Jan 01, 2015]