International Journal of Engineering and Technical Research (IJETR)
ISSN: 2321-0869, Volume-3, Issue-2, February 2015

Football Team Selection Using Genetic Algorithm

Vasudha Sarda, Prasham Sakaria, Prof. Khushali Deulkar

***Abstract*— The team selection process is a vital requirement of management in various fields and organizations such as in government agency, project management, industry, sports as well as in companies. The traditional team selection process is really lengthy, awkward and unclear due to manual process and personal judgment. Predicting the result of a football game is challenging due to the complexity and uncertainties of various attributes, which are involved in calculating the performance of each player. As there are conflicting constraints and personal judgments involved it is better to have software for an optimized result. This paper presents a solution to this problem with the help of genetic algorithm to find the optimal solution for the problem of football team selection and formation. In this paper we will be creating a model which combines the generally used quantitative approach with some new extensions such as attributes regarding personal performances, team performance along with the collaborative performance of a player in the presence of other players in the team. It will also be taking into consideration the budget of the football leagues and the cost of each player.**

***Index Terms*— Genetic Algorithm, Football, Team Selection, Multi-objective criteria.**

I. INTRODUCTION

American football is a game played by two teams having 11 players each on a rectangular field with goalposts on each end. The offense, team with the ball tries to advance down the field by running or passing the ball while the other team without the ball, defense tries to stop them by hitting the ball on the goalpost. There are many leagues like National Football League (NFL) the most famous league in United States requires team selection and formation before the matches start.

To make the team selection process precise, accurate and unbiased it is best to use team selection software which uses genetic algorithm to form a team with players having the best performance in the budget allotted. There can be an adaptation of island genetic algorithm to optimize the selection of multiplayer sports team having multiple conflicting constraints with mixed crossover where the fitness of common solution is used to drive the selection. We can use a new methodology for objective evaluation of American football team selection using a bi-objective genetic algorithm. An analysis the obtained trade-off solution has been shown to result in a preferred team that has been found to have better defending and offending averages of the winning teams in the National Football League. This can also help to include other factors and constraints to create a more practical team using a systematic process mentioned in the paper. By integrating this method along with statistical analysis by using a dynamic optimization technique to be applied in auction can develop a standard methodology for team selection for games having multiple players.

Abstract factors like team coordination etc. can also be used for decision-making process. Nevertheless, the procedure suggested in the paper clearly demonstrates the advantage of using a bi-objective computing methodology for the football team selection in major league tournaments. The issue is their time complexity, which makes it not suitable for the real time applications. But our method is just not specific for football team but it is converted to generic model for other multiplayer games.

Manuscript received February 12, 2015.

Vasudha Sarda, Computer Department, D.J.Sanghvi College of Engineering, Mumbai, India, 9833971010

Prasham Sakaria, Computer Department, D.J.Sanghvi College of Engineering, Mumbai, India, 9769055518,

Prof. Khushali Deulkar, Computer Department, D.J.Sanghvi College of Engineering, Mumbai, India, 9594049807.

II. METHODOLOGY

A. Genetic Algorithm

Genetic algorithm (GA) is a search heuristic that mimics the process of natural evolution and is used for optimization and search problems to come up with solutions. A genetic algorithm is a part of evolutionary algorithms, which use techniques such as mutation, selection, crossover to generate solutions for optimization problems. The individuals of Genetic Algorithm are encoded in the form of strings where a collection of strings is termed population. Initially, a random population is created, which represents different points in the search space. An objective function is associated with each string that represents the degree of the goodness of the string. It is based on the principle of survival of the fittest where only some strings are selected after crossover and mutation to create the next population from the parent string. The process of selection, crossover and mutation continues for a fixed number of generations or till a termination condition is met. Information entropy model is an effective performance evaluation tool. Pay attention to offending and defending the mainstream that the balance has already become sport development of current football. Eventually, a modular approach as shown in is proposed to adopt appropriate optimization technique considering different possibilities in aforesaid team selection problem. The processes of Genetic Algorithm are:

Initialization

In the start of the process there are many solutions, which are generated to form the initial population. The population size is decided by the nature of the problem, but usually contains several hundreds or thousands of possible solutions. Most of the times the population is generated randomly to allow a set of possible outcomes or solutions.

Selection
During each successive generation, a part of the existing population is selected to generate a new generation. Individuals are selected via a fitness-based process, during which the solutions with high fitness value are more likely to be chosen. Certain selection methods rate the fitness of each solution and preferentially select the best solutions while there are some which try to save time by random sample of the population.

Crossover and Mutation
Crossover is applied on two randomly selected chromosomes from the initial population. After the selection process, it is time to generate the second-generation population of solutions from the first generation population using processes like crossover and mutation. To create a new solution from the pool of selected chromosomes using the mentioned processes to create a child from the parent. Further new parents will be selected for each new child, and the process continues until a new population of solutions of appropriate size is generated. The average fitness will be increased with each population since only the best organisms from the first generation are selected for breeding. Similarly Mutation is applied on a randomly selected chromosome with a mutation rate of 0.5.

Termination
This generational process is repeated until a termination condition has been reached. Usually the conditions are either a solution is found that satisfies minimum criteria or a fixed number of generations reached. Allocated budget such as the computation time reached or the highest ranking solution's fitness is reached or has reached a plateau such that successive iterations no longer produce better results.

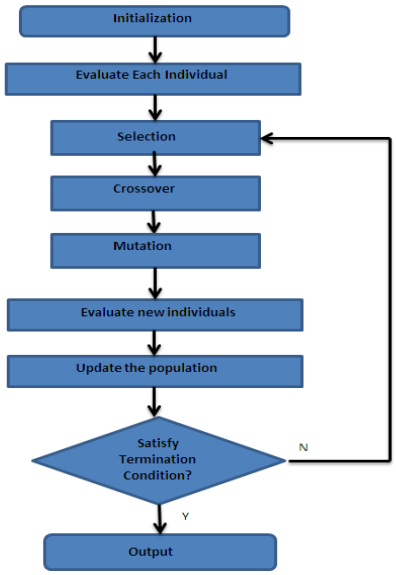

Figure: Flowchart for the working of Genetic Algorithm

B. NSGA 11-Non Dominating Genetic Algorithm

NSGA-II addresses those issues, and additionally, suggests a simple constraint-handling strategy for constrained multi-objective optimization that can be also used in any other Evolutionary Algorithm. The multi-objective programming process deals with optimization of the problems in the field of multi-criteria decision making involving more than one objective function to be optimized simultaneously. A solution is called non-dominated; Pareto efficient or non-inferior, if none of the objective functions can be improved in value without degrading some of the other objective values. Without additional subjective preference information, all Pareto optimal solutions are considered equally good. Researchers study multi-objective optimization problems from different viewpoints and, thus, there exist different solution philosophies and goals when setting and solving them. The goal may be to find a representative set of Pareto optimal solutions in order to quantify the trade-offs in satisfying the different objectives or it can be used in finding a single solution that will satisfy the subjective preferences of a decision maker.

C. INSGA 11-Improved Non Dominating Genetic Algorithm

Generally in the member selection problems the individual performance concerning a single candidate is used while collaborative performance concerning a single candidate is not taken into consideration. In this paper, as a solution to this problem, we propose a method for member selection of cross-functional teams, where both the individual performance of candidates and the collaborative performance between candidates are considered. A multi-objective 0–1 programming model is built to solve the member selection problem where the individual and collaborative performances are taken into consideration, which is an NP-hard problem. In order to solve such problems we can use the improved non-dominated sorting genetic algorithm II (INSGA-II). Furthermore, a real example is employed to illustrate the suitability of the proposed method. Additionally, extensive computational experiments to compare INSGA-II with the non-dominated sorting genetic algorithm II (NSGA-II) are conducted and much better performance of INSGA-II is observed. The NSGA-II significantly contributes to solving constrained multi-objective problems and it alleviates three difficulties in previous multi-objective evolutionary algorithms that use non-dominated sorting and sharing, namely, computational complexity, non-elitism approach and the need for specifying a sharing parameter. An INSGA-II based on NSGA-II is developed to solve model.

III. Literature Review

For the success of a sports team it is very important to form a correct team. Team selection in most sports is a subjective issue involving commonly accepted notions to form a good team. Generally the subjective methods were used to select a team but now even the objective methods can be considered to form the best team for big leagues, which invest huge capital on their teams. In sports each player is paid a certain amount by the owner of the team that is called as the player's cost, which is either fixed or is decided through auction. **[1]**
Our research targets the 16 teams participated in Euro 2012. In one of the paper they had adopted group single round robin scoring rule where two teams winning out every group will become the final eight, then through playoffs, the final four, champion and runner-up teams will produced, which involve a total of 31 games. Along with this to calculate the technical

and tactical ability of a team, we need to consider the factors that reflect the team technical and tactical ability, and take the numbers of game played into consideration. **[3]**

Fitzpatrick and Askin (2005) point out that the formation of teams requires consideration of innate tendencies, interpersonal skills and technical skills. It is also seen that communication of team members has an effect on team performance (Campion et al., 1993). In order to avoid wasting time selecting a wrong team or wasting money it is important to ensure the success of cross functional teams. **[2]**

In order to predict sports performance more accurately, one of the papers used a hybrid prediction system based on genetic algorithm we investigated the recent algorithms, and proposed a hybrid prediction system based on genetic algorithm and artificial neural network (GANN). This is the first paper using the GANN to predict sport performance. Cricket as an example, they employed neural networks to predict each cricketer's performance in the future based upon their past performance. They classified cricketers into three categories, which are performer, moderate, and failure. The neural network models were progressively trained and tested using four sets of data. The trained neural network models were then applied to generate a forecast of the cricketer's near term performance. **[4]**

Another paper presents a new model for team formation based on group technology using parallel hybrid grouping genetic algorithm. Recently, several researchers have dealt with the problem of team formation without leaders in different scenarios, such as multi-functional skill requirements of the teams. In these situations, the makeup of the teams must take into account the personal or technical skills of the staff. In situations with very different staff skill categories, the problem may be difficult, and there is not a unified definition of the model to be used. In a novel approach to this multi-functional version of team formation is presented, and fuzzy planning to match customer's requirements and engineers' characteristics (skills or knowledge) is proposed. In this case, the hybrid-grouping genetic algorithm employed is modified to consider the peculiarities of the TFPGT. **[5]**

IV. EXPERIMENT

Data is collected of various football players based on various attributes are taken into consideration. The offensive attributes, which are taken into consideration, are goal, shoot, corner kick, offside, free kick, shoot kick, possession percentage and crossing. While the defensive attributes taken into consideration are fumbling, shoot and foul. Weights will be assigned to each of these attributes to calculate the performance score for defenders, offenders and midfielders. Each player is classified into three types forward, defender and midfielder which is shown as FOW, DEF and MID in the database. This is a sample database where the attribute common to all players is shown above. The database has two attributes, which combine to give the individual performance score. There are attributes, which will be used to calculate defender performance score, other attributes, which will be used to calculate the offender's performance, score as well as some attributes which will be used to calculate the midfielder's performance score. These attributes will be different from each other while there will be some common attributes such as fitness, stamina, strength, mental health, matches won and matches lost. The attributes related to the fitness of the players are taken from the fitness examiner based on the fitness test, which is taken before every match through surveys. The other attributes such as dribble, rate and shoot are taken from the previous matches played by the player. Along with this we need to consider the interpersonal skills between the players, which is really important attribute, which is generally not considered.

Figure: The GUI of the experimental model

The GUI for the football team selection software is shown where the input needs to be given which can run on any computer along with MATLAB. The budget for the team, which is the upper limit of the cost ready to be spent on the team, is mentioned. Along with this the number of players as forwards, midfielders and defenders need to be mentioned. On the right the user needs to mention if he would want the team to be defense oriented, offense oriented or an optimal one as the budget will be allotted accordingly. If the user has marked a defense-oriented team then strong defenders will be taken which have a high defense performance value and who will cost more. The output for the following input will be a list of the players which have been selected along with the total amount spent for the team which will be either equal to or less than the budget mentioned by the user, along with this the net performance of the team will be calculated and displayed. The net performance of the team will be the average value of the individual performance of the players, which have been selected. Weights will be assigned to the attributes based extend of influence the attribute has on the performance of the player. Then the performance score threshold will be assigned so that only the scores above the threshold are considered.

Budget Analysis:

For selection of a football team for any league a budget is assigned from the start. Every player has a particular price and a team has a budget. There needs to be a proper tradeoff between the budgets assigned as well as the performance of

the players. Both of them need to be considered and given equal weightage which considering a team. There can be times when you spend a huge amount of the budget on one star player but there isn't enough money to be spent on the other good players and the overall net performance of the team is very low. A football match cannot be one by one very good player but it is won by taking decently good players playing together. Similarly assigning the entire budget to defenders will give you very good defenders but a very weak offense will lead to a disaster and the match will be lost. At the end of the game the amount of the budget spent and overall net performance score of the team should be given which can be used later on when your selecting teams in future or it can be compared with another possible team to get a proper tradeoff between the two. The budget constraint affects defenders dominant team's more than offender-dominant teams. This is because the price difference among defender with a high defender average and those having a low average is significant. The same effect does not exist among offender and hence the minimum offender average region of the trade-off front is fairly unchanged due to increase or decrease in the budget value. Such an analysis can provide the team managers an idea about how much gain in defending and offending averages is possible by spending an increase in the amount of money in selecting a team. Interestingly, for the chosen player's performance statistics and price, a large amount of investment need not make a better offending team; however, a better defending team is possible to be formed by investing more money.

Knee Point Approach:

Obtained trade-off front comprises of a set of points in the objective space representing various teams, which will include their performance as well as the budget. For the given problem, we prefer selecting the team represented by a knee point present in the knee region of the Pareto-optimal front. Such a region is preferred because deviating from the knee region by a small change such as the defender performance in the value of one of the objectives will compromise a large change in the value of the other objectives which is the defenders which are selected or it can also be the change in the budget allotted can lead to change in the team performance score. There are many ways of finding the knee region in a two objective scenario like the team selection process. But we need to include criteria such as the team expenditure; brand value of players, midfielder's performance and the success rate of the manager while carrying out the analysis from the selector's point of view.

V. Future Work And Conclusion

Here for football team selection to calculate the collaborative performance of the team players, the influence of a player's presence on the other player is one of the constraints, which needs to be overcome. Along with selecting players for a team even selecting a manager for the team needs to be considered and calculated within the budget given for the team. When we are considering the manager it is important to note the number of matches won and lost under the leadership of a manager as well as the number of players who get along with a manager, this coupled with the experience the manager has. The results for the proposed multi-objective optimization and decision-making methodologies, which was applied for football game, can be applied to game of cricket, applied to IPL Twenty20 format with some minute changes. It can be applied to any team formation process where a few players need to be selected from a pool of players considering their past performance and present fitness tests. The team representation scheme, which was applied for football team along with the operators used in genetic algorithm, can be applied to other sports like basketball, hockey and other similar games. Thus, with a little modification, our representation scheme and proposed genetic operators can be extended to a team selection for Indian cricket like the IPL. The obtained trade-off teams can then be analyzed to identify high-performing players for different positions. Along with application in sports it can be applied to the teams, which are formed for various projects in companies. Here the team members with different skill sets are selected to work together for a particular project. In this case the collaborative performance between the team members are also considered for the work to get done efficiently.

References

[1] Faez Ahmed, Kalyanmoy Deb, Abhilash Jindal. "Multi-objective optimization and decision making approaches to cricket team selection" Applied Soft Computing Volume 13, Issue 1, January 2013, Pages 402–414.

[2] Bo Feng, Zhong-Zhong Jiang, Zhi Ping Fan. "A method for member selection of cross-functional teams using the individual and collaborative performances" European Journal of Operational Research Volume 203, Issue 3, 16 June 2010, Pages 652–661.

[3] Minjian Wang. "Evaluating Technical and Tactical Abilities of Football Teams in Euro 2012" Advanced Science and Technology Letters Vol.53 (AITS 2014), pp.477-483.

[4] Biao Xu. "Prediction of Sports Performance based on Genetic Algorithm and Artificial Neural Network" International Journal of Digital Content Technology and its Applications (JDCTA) Volume6, Number22, December 2012.

[5] 5. Luis E. Agustín-Blas, Sancho Salcedo-Sanz, Emilio G. Ortiz-García. "Computers & Operations Research" Volume 38, Issue 2, February 2011, Pages 484–495.