Evolution of Sports through Operation Research

Group number: 4
Group members' names:

HARSH JAIN
ADHIP SAINI
MEHAK MALHOTRA
GAURAV BHARTIYA
HIMANSHU SACHDEVA

INTRODUCTION

OR in Sport is a field of research and analysis undertaken by hundreds across the world. This paper does not attempt to give a complete history of OR applied to sport, but instead, it aims to give the reader a flavor of what is now a large and diverse field. There is so much pressure and weight to these tournaments that they require long-term planning, strategy, organization and execution. This is where operational research comes in. It helps in decision-making, increases productivity, increases efficiency and coordination(Wright, 2009). In sports, operations research is mainly used for planning (between teams and between cities), forecasting and generating strategies to achieve optimization in sports.Thus, operational research could be used more effectively to reduce travel time, exhaustion and costs.The difference between sports and games cannot be clearly defined, but according to the working definition, a sport is something that is friendly, not related to the enemy, and involves some physical activity even to a small degree, and cannot be played at a distance. There are games like chess and poker that are the opposite of the above definition.OR can be used for different purposes and by different individuals such as team captain to select team players, sports management bodies, coaches, league association etc. The main purpose of applying operational research tools is to solve the problem of effective decision making either directly or indirectly. (Wright, 2009)Whether selecting a football team or managing a sports league, decisions are made for which OR is used consciously or unconsciously. In this article we can see the use of operations research in basketball, cricket, casino, football, tennis, formula 1, using its tools such as stimulation, dynamic programming, game matrix, game theory, pit stop strategy, energy and management strategy, mathematical models such as coin in per day and win per day, circle method and LPP.

LITERATURE REVIEW:

The Research paper states about application of Operational Research in various sports i.e., Football, Cricket, Tennis, Casino, Formula 1, Basketball. Multiple OR theories like Markovian decision process, OWA, LPP, Data Simulation, Ozceylan 2 phase approach, Monte Carlo Tree Search, Game Theory, Q-Mode, Reinforcement Learning, etc. To solve problems like pit-stop strategy for Formula 1, Thermal and energy optimization, Positioning of players in football, Ranking and qualification in tournaments, Fantasy team selection, probability of winning in casino, efficiently utilize the skill of players in basketball.

Football: In Football team line-up and players' position is allocated through application of OWA, the semi-markovian decision process, and Ozceylan 2 phase approach (Tukhachevskiy, 2019). **Formula 1**: Teams use data simulation and game theory to implement the pit-stop strategy in formula 1, which helps them to formulate correct time to pit-stop, saves fuel and time too(Coulthard, 2008)(Environment *et al.*, 2014). In Formula 1 itself OR is used to thermal and energy optimization of the cars, through Q mode, Monte Carlo Tree Search and Deep reinforcement Learning(Liu, Fotouhi and Auger, 2021). **Fantasy Cricket**: Application of Linear programming to select the team for fantasy cricket, which helps the participant to forecast and select the players (Analysis, Learning and Vision, no date). **Tennis**: use of Round robin scheduling method to conduct tournaments, which gave equal n fair opportunity to each and every competitor(Della Croce, Tadei and Asioli, 1999). **Casino**: probability of winning by analyzing Coin-in per day and win per day using Linear programming (Ghaharian and Singh, 2010).

KEY WORDS:

Operational Research, Sports, Monte Carlo Tree Search in Formula 1, LPP in fantasy cricket, OWA, Game Theory in Formula 1 and Basketball Round-robin scheduling in Tennis

FINDINGS & ANALYSIS:
Football

In every sport, whether it is football, basketball, or any other collective sport, the selection of human resources is fundamental. OR is used to analyze and select the best athlete according to the interest of the team.

Scientific literature describes the contributions of OR to sport. In the football industry, however, topics such as player selection and team formation are not properly studied. Due to the subjective nature of decisions relating to team formation, quantitative methodologies have the potential to overcome these issues and offer better sports results Thus, this study will review operations research techniques used for the formation of football teams.

1. This research aims to determine the correct timing of players transitioning during a football match. Carling Opta Football Yearbook data provides a wealth of information regarding players participating in the English Premier League in 1988-89.

2. The calculation of transition is done according to a Poisson regression model with data of games played during that season. The case study was done by simulating games between three teams, "A", "B", and "C", where the team "B" had the option to choose between two players: "P" and "Q". Match data was then entered into a dynamic programming model to identify the optimal replacement time between players, which aims at maximizing the number of points expected from the match. The results showed an increase in point expectation with the transition of player "P" By player "Q".(Tukhachevskiy, 2019)

3. Using the decision tree technique, Myers (2012) then developed a replacement strategy for improving match results.

4. The following variables were analyzed: goal differential before and after substitution, home advantage and time substitutions occurred. The results showed that when a team was losing before the first substitution, it was able to improve its result 41% of the time when the change was made before the 58th minute. In the second substitution, the model indicated that this should be done before the 72nd minute, with an improvement of 30% of the cases. Finally, the last substitution was indicated to happen previously to the 78th minute of the match, which was effective 24% of the time.

5. **Merig'o & Gil-Lafuente (2011)** developed new selection indices that eliminate preferences and subjectivities. The proposed indexes consist of the combination of the Hamming distance, the adequacy coefficient and the maximum and minimum level index with the **OWA**. (Tukhachevskiy, 2019)

6. **Jarvandi (2013)** states that failures may be related to the adaptation of the new player in a team. His model assists the clubs in quantifying compatibility effects, which helps in obtaining a better prediction of a new player's potential. The interactions between players are then analyzed through a **semi-Markovian decision process**.

Team lineup

A system was later introduced that was capable of assisting coaches to determine the optimum team lineup, At the first stage it determined all the qualities relevant to a team. The players are evaluated according to the qualities listed for each position. The importance of their qualities is also determined which reflects the game system and tactics that the coach prefers for his team. A mathematical programming model is used in which all possible choices for 11 positions and 26 individuals are calculated. It seeks to acquire a team with the maximum weights per player position.

In 2013 Tavan's research explored the interaction between players. The method is divided into two moments. Initially,rankings are made to choose the best team members on the basis of their performance on criteria and their respective positions. Then, in the combining phase, the fuzzy inference system is used to evaluate combinations between team members. The variables used in this model are provided by the coaches, based on combination factors i.e., number of matches that athlete played together and number of passes between them. The result of this model is that it gives the combination percentage of the best sets of players in each sector of the field.

Ozceylan two-phase approach.

● The first phase of the method is used for identifying the main attributes of a position, for which the AHP (Analytic Hierarchy Process) is used. In the second phase, integer programming is used to select the best set of players. Football Manager 2015 game data were used as criteria and performances of the evaluated players.

In 2017, l-Shboul used historical game data between teams and player statistics. His goals were: (i) To select the best combination of teams, according to their opponent; (ii) Predict probabilities of victory, according to the knowledge of the team. The result of the match is then compared with each player present, so that their contribution to final results can be evaluated. (Tukhachevskiy, 2019)The validity of the model is calculated by the ratio between correctly predicted results and the total data set, which generates an accuracy of 60%.

FANTASY CRICKET

Dream 11 is an Indian fantasy sports platform. Dream 11 users participate in contests with their teams. Each player is assigned a credit value, and they can be batsmen, wicketkeepers, bowlers or all-rounders.

Basic terms in LPP-

1. Objective function- the value of this function should either be minimized or maximized.
2. Decision variables- these are used as inputs to decide the final output
3. Constraints- these are limitations put on decision variables.

Connection between cricket and linear programming-(Analysis, Learning and Vision, no date)

To create a Fantasy-Team we need to select the best players and the final team should follow rules and regulations which are asked by the platform so this can be solved through linear programming.

PULP- it is an open-source python package which solves optimization problems so it works entirely within syntax of python language by providing python objects which represents optimization problems and decision variable and allows constraints to be expressed in a way similar to original mathematical expression Solving linear problem in python using PULP-

Following are the code snippets that will help to solve the problem

> 1. Package installations
> 2. Input data

To create a linear programming problem using following set of instructions-

> 1. Objective function and constraints require decision variables
> 2. Defining objective function- these are based on either each player points, percentage times player has been in best 11 teams, average points per match.
> 3. Adding constraints- these can be based on number of players, or players from team 1, or players in lineup

There are three types of information to solve LPP-

> 1. Objective function has to be maximized or minimized.
> 2. Constraints or rules need to be followed while maximizing or minimizing
> 3. Input variables on objective function and constraints have to be applied.

Results- it will try to solve the problem by identifying the best selection that will maximize the objective function and fulfill constraints that are set. If the case is that they are not able to find a feasible solution so status will be infeasible.(Analysis, Learning and Vision, no date)

Finally, after this method we can give a choice of constraints and click on the "create fantasy team" button and your fantasy team is ready to participate in the contest.

TENNIS

Tennis is a game played with two opposing players (singles) or pairs of players (doubles) using tautly strung rackets to hit a ball of specified size, weight, and bounce over a net on a rectangular court. In Tennis one of the big problems faced by cities is related to scheduling and organizing the game especially in winters as the courts need to be booked well in advance to utilize the time and court better. This leads to players having limited chances of playing matches hence they decide to not participate. Hence, to boost participation clubs decided to have a Round Robin tournament allowing every player to play one round every week with different opponents which enables all players to know other players.

ROUND-ROBIN METHOD

Here, using round robin method the tennis club provides a predetermined number of hours and court for the competition each week, and then the players must give a calendar of hours of each day they will be available or able to participate. Such tournaments can also encompass a wide range of ages, from children to adults. To accommodate both players' availability, each group of players is assigned not just a round but also a certain court and time for that round. Also, to be noted that availability of players might change weeks to weeks. In the world of tennis, this method is widely used towards the end of the year when organizing the ATP Finals and Next Gen Finals. (Della Croce, Tadei and Asioli, 1999)

Within a round robin schedule, many methods are utilized to calculate final rankings. Some competitions determine winners based on the number of matches won. When two players tie, the one who wins when they verse each other breaks the tie and takes the higher position. When more than one competitor is tied, the number of games won minus the number of games lost determines the winner. If the score is still tied, the winner is determined by points scored minus points against. A points system was another method for determining round robin schedule winners. Participants earn points for winning a game, a match, or earning more points. With this option, total points would decide the final standings.

OBJECTIVE-

The objective of scheduling or conducting a Round robin tournament is to maximize the no. of possible matches. A breakdown strategy was used, such as first designing a round robin competition to circumvent court hour and availability restrictions, and then assigning or allocating rounds to the weeks following players' availability. (Della Croce, Tadei and Asioli, 1999)Then two different approaches or procedures were used-

> 1) GPI Procedure
> 2) Tabu search Procedure

Following a thorough examination of the GPI procedure, it was discovered that the format under consideration was repetitive, implying that the pattern was only useful for rounds of the same series, so the second procedure was implemented to address this issue by focusing on only one series at a time.

CASINO

As we all know, a casino is a live dealer game or a place where visitors come and play games(bet) in which they may win or lose money. Basically, in casino's there are combinations of slot machines placed on the floor, but many casino owners or managers don't follow a formal or systematic process for placing the slot machine and rely on their customer's feedback or inputs. So here we used linear programming to design the slot floor to reach an optimal solution.

TWO MATHEMATICAL MODELS WERE DESIGNED

The two models were -
1)Coin-in Per Day
2)Win Per Day
This Casino, like all its competitors obtains the majority of its revenue from slot machines. The three criteria that were considered in determining the hypotheses were Coin-in per unit per day, Win per unit per day and Promo Liability Per Unit Per Day. (Ghaharian and Singh, 2010)
In the first criteria the total number of coin-in (total amount of money played into the money) generated was divided by total no of days that machine was on the floor during that 6-month period.
 In the second criteria the total win (the total amount of money that the casino has after all players having paid their bets) was divided by total no of days that machine was on the floor during that 6-month period.
 In the third criteria Promo Liability Per Unit Per Day refers to a sum of money which the player is handed back and its calculation is pretty complex.

EXCEL 2007's SOLVER ADD-IN

After designing the models and stating the problem, the issue was resolved using Excel 2007's Solver Add-In (tool used to solve equations using excel). There were also various software's available, but Excel Solver was chosen as the best fit because it not only creates the answer report but also does a sensitivity analysis to determine the impact of modifications to the objective function, coefficients, and restrictions. After interpreting both the reports, the casino owners can now check which machines are generating more profit or revenue and have the best combination of slot machines on the floor in a more accurate and systematic way.

Formula 1

Formula 1 as a sport is very high-paced and highly competitive, involving huge investment and stake too, **Pit-stop strategy, Energy management and Thermal management** are the 3 were very important places where we found application of OR. in which even a bit of time will be a big issue, so it is very crucial to know which time would be best for pit-stop and what amount of fuel and energy is required at what time. Its importance was increased in 2003 when the competitors had to start the race with the same amount of fuel, pit-stops, and refueling were also implemented in Formula 1. There was a requirement for the formulation of a systemized and logical way to manage these. Hence, they needed to optimize energy, gain time and neutralize temperature, for which there was much juggling between different methods and techniques.

A. PIT STOP STRATEGY: This strategy first came into the picture when in the **Canadian Grand Prix** there had been the application of the split strategy between the red-bull drivers under the strategic operations' head "**Neil Martin**" one-stop strategy and 2 stop strategy(Coulthard, 2008), the one who used one-stop strategy finished 3rd, which aspired the other drivers like Toyota to use the one-stop strategy wherein they used to predict the correct time to pit-stop according to the requirement.

  The teams started using the Data fed by drivers to analyze the accurate time to pit-stop through the application of the "**Game-theory** "(Environment *et al.*, 2014), which refers to the strategy of analyzing data and having a relative comparison between rational competition on different factors like tire wear, Fuel, engine-components, etc.

This strategy brought a huge transition in Formula 1 as a sport as this didn't remain the competition between drivers anymore, it became the competition between the two sides, which side will be able to analyze the data provided to them accurately, and which driver would be able to implement the instructions-conveyed better. This strategy also resulted in a new perspective in the world, that the quality of the car doesn't have much effect if the data and interpretation are done appropriately, it also made the sport more technical rather than just being seen as a game of speed. It made the pit-stop strategy very technical rather than vaguely wasting the time.

B. ENERGY MANAGEMENT STRATEGY:

Researchers juggled many techniques to solve the above problem. Lime beer et al. studied energy management strategy for a Hybrid system. Herrmann optimized electric usage for an autonomous car. Liu studied both energy and battery thermal constraints. This problem became more difficult to solve as the dynamic nature was exposed in the multi-lap race especially. By taking all the constraints in mind, Monte Carlo methods were expected to serve the best as it was already used in many other racing games at that time, but its disadvantages were that it was not much reliable to solve this problem as it is dependent on brute computation force, hence to increase efficiency and accuracy, Liu proposed to solve it by **Monte Carlo Tree Search (MCTS)**, it focuses on the relatively more relevant and promising actions. Further MCTS was found to be subject to improvisation as it took 10 seconds to make decisions, which was a huge time considering the highly dynamic nature of the sport. To overcome this there was a formulation of a new method through "**Deep Reinforcement Learning**" (it is a branch of machine learning algorithm designed to optimize actions in an environment in the form of a Markov decision process to maximize the collected reward)(Liu, Fotouhi and Auger, 2021). In the case of Formula 1, time was Taken as a reward, integrated with constraint violations as penalties. Then, this method was tested in different cases (temperature, car type, etc.), each time making an addition to the method on the basis of different constraints. Like the /-learning method, the **Deep Deterministic Policy Gradient Method compares algorithms of DRL, A2C, and Architecture – Ape-X**.

Deep deterministic policy gradient method under **Ape-X** architecture was suitable, Reward shaping with soft penalties benefited performance of **RL algorithm**. The continuous-action-based method successfully found superior strategies of 3-4 seconds faster than the discrete-action-based method. This whole study gave an optimal and efficient way to manage energy. The backhand team itself would take care of it and convey the instructions to drivers, which makes drivers' life simpler and reduces their headaches.

C. THERMAL MANAGEMENT STRATEGY

Tremlett et al tried studying the Optimal tire usage of Formula 1 cars by involving a **thermal-dynamic model** in OCP, Andon the other hand, for Thermal management, **Smooth pedal operation** proved to be an effective way to neutralize the temperature. This smooth pedal was subjected to continuous action, which is referred to as **Q-Mode**.(Liu, Fotouhi and Auger, 2021)

This somehow minimized the risk involved of breakdown of the car due to high temperature, which was one of the huge reasons for accidents in Formula 1 earlier. Hence, now Formula 1 became safer as a sport as risk of accidents minimized. This also would minimize the impact of these uncontrollable factors on the race.

BASKETBALL
A basketball game involves two teams of five active players, each trying to score points against the other by throwing a ball into a 300cm high hoop (the basket). Both teams strive to protect their own goal while attempting to score on the other team's side, to analyze the strengths and weaknesses of the other teams, NBA (National Basketball Association) coaches and personnel analyze past games and statistics. In order to achieve the best results, they suggest strategies for teams to adopt that can be backed up quantitatively.(Srikantha and Kundur, 2017) It used to be difficult to determine which strategy would be optimal for the offensive team and defensive team in the current game state. Few strategies were introduced to help the team perform better during the game

GAME THEORY MODEL

With this model, for each match-up, there are 25 possible levels of aggression for one shot, using an offensive team, a defensive team and a handler. Nevertheless, guards can only defend other guards in this model, so there are only two ways to assign that match-up and six ways to assign the front court.

GAME MATRIX

Using each player's moves, we create a game matrix and calculate the ability of a specific defender against a particular shooter as well as the ability of the shooter to shoot from that zone and the difficulty of passing into it.

1. Then we create a matrix from the data and solve for the optimal strategies. The best strategy will be a combination of each player's skill set and the goals that must be met during the game.
2. Determining those each player is assigned with each strategy or combination of players perform on certain strategies together. In this situation, the game isplayed according to the strategies. Therefore, the following are the results:
3. The game is won by the team.
4. Each player of the team was able to use their skill set in the best way possible.
5. The team was able to counterattack the opponent's strategies.
6. Strategic goals were met as planned.

CONCLUSION
We have come to the conclusion that Operations Research has enhanced and optimized the sports industry in a positive way. Operation Research is flexible and can be applied in various areas of sports industry. We focused on various areas like scheduling

and organizing of sports events, selecting the best athletes for the team, find optimal solution of a problem of creating a fantasy team, design a slot floor in casino. All sports require scheduling and use Operations research for the same as they are inter-city tournaments on a large scale. Finding optimal solution with the help of operation Research is crucial and important in sports as helps the team having an advantage to their opponent. This explains how beneficial the application of operations research is to the sports industry and how it can draw commonalities and differences at the same time.

REFERENCES

1. Analysis, F., Learning, D. and Vision, C. (no date) 'Lakshmi Ajay 791', pp. 5–6.
2. Coulthard, D. (2008) 'Computers are Formula One's wizards of winning', pp. 21–23.
3. Della Croce, F., Tadei, R. and Asioli, P.S. (1999) 'Scheduling a round robin tennis tournament under courts and players availability constraints', *Annals of Operations Research*, 92, pp. 349–361. Available at: https://doi.org/10.1023/a:1018999101596.
4. Environment, M.C. *et al.* (2014) 'F1 strategy and pit stops', (February), pp. 1–9.
5. Ghaharian, K.C. and Singh, A.K. (2010) 'A mathematical approach for optimizing the casino slot floor: A linear programming application', *ProQuest Dissertations and Theses*, p. 58. Available at: https://www.proquest.com/dissertations-theses/mathematical-approach-optimizing-casino-slot/docview/852517101/se-2?accountid=188653.
6. Liu, X., Fotouhi, A. and Auger, D.J. (2021) 'Formula-E race strategy development using distributed policy gradient reinforcement learning', *Knowledge-Based Systems*, 216. Available at: https://doi.org/10.1016/j.knosys.2021.106781.
7. Srikantha, P. and Kundur, D. (2017) 'A Game Theoretic Approach to Real-Time', 13(3), pp. 1006–1016.
8. Tukhachevskiy, M.N. (2019) 'Tactics and Strategy', *The Soviet Art of War*, pp. 44–45. Available at: https://doi.org/10.4324/9780429314681-7.
9. Wright, M.B. (2009) '50 Years of or in Sport', *Journal of the Operational Research Society*, 60(SUPPL. 1), pp. 1–7. Available at: https://doi.org/10.1057/jors.2008.170.