Application of Game Theory and Markov Chains on English Premier League (EPL) Scorelines Analysis

Christogonus Ifeanyichukwu Ugoh a*, Chinwendu Alice Uzuke a, Obiora-Ilouno Happiness Onyebuchi a Obi-Okpala Chinelo Ijeoma a, Orji Gabriel Oyo a and Eze Theophine Chinaza a

a Department of Statistics, Nnamdi Azikiwe University, Awka Nigeria.

Authors’ contributions
This work was carried out in collaboration among all authors. All authors read and approved the final manuscript.

Abstract
The aim of this paper is to obtain the optimal strategies of two competitive players using Game Theorem and to make future predictions of games using Markov Chains involving the EPL. All the teams that have participated since 2005/2006 EPL season to EPL 2019/2020 season were considered and the method of proportion of wins was used to select five best teams. Linear programming was employed to select the optimal strategies, while the predictions for seasons 2020/2021 to 2023/2024 are obtained by Markov chain method. The results obtained revealed that Man U is the optimal strategy for Player A, and that Player A has to choose Man U to maximize his profit, meanwhile, Chelsea is the optimal strategy for Player B and he has to choose Chelsea to minimize his loss. The findings of the results also revealed that for Man U or Chelsea to win their home games, it will depend on their current home winning against the team they are playing with.

Keywords: Game theory; markov chains; scorelines; linear programming; payoff matrix; strategies; transition probability matrix.
1 Introduction

Scoreline is the score or final result achieved by the players in a game or competition. Game theory does not ascertain how game should be played, but rather tells us the procedures and principles by which the actions should be selected. Markov chain in the other hand is used to predict the chances of a team winning a particular game relied on its current and previous games.

English Premier League (EPL) is the top level of the English football league system, which is contested by twenty (20) clubs. It usually operates on promotion and relegation, where the last three clubs that finish on the bottom (that is 18, 19, and 20th position) relegate to EPL Championship, while the clubs that end first carries the league and is crowned champion of the EPL. EPL is the most viewed league in the world and the most rated.

Game theory is an important tool used in decision making among rational decision-makers [1-3] analyzed the scoreline of the EPL using Game theory and Markov Chains [4] carried out a prediction analysis on the outcome of an election in a given situation using the application of Markov chain [5] studied the performance of Markov chain model in predicting time series data [6] presented a machine learning algorithm for modelling route prediction using the application of Markov chains. [7] presented an approach in optimizing tactics and strategies in football and a novel model for making both pre-match and in-match decisions, using the game theory approach [8] analyzed the different methods and techniques used in solving game problems [9] analyzed the Transmission Cost in Slotted ALOHA and ZigZag Decoding using the application of Game theory and Markov chains.

2. Materials and Methods

2.1 Method of Team/Club Selection

Five teams from the English Premier League (EPL) within the period review, that is, from 2005/2006 season to 2019/2020 season for this study are selected from the list of the teams that have not relegated to Championship League using the method of proportion. This study relies only on the proportion of wins. The total games played by the twenty (20) teams of the EPL in each season is obtained, and the total number of wins (both home and away matches) is obtained, and the proportion of wins for each most occurred teams that has not relegated under the period review is obtained using equation (1).

\[ \text{Porportion of Club Wins} = \frac{\text{TCW}}{\text{TCP}} \]  

(1)

where TCW is the Total Club wins from 2005/2006 to 2019/2020, and TCP is the Total Games Played by the Club from 2005/2006 to 2019/2020

2.2 Game Theory

Game is a competitive situation among some finite number of persons or group of persons known as players conducted a prescribed set of rules with known payoffs [10], defined by the players, actions and outcomes [11].

Game theory is a theoretical framework to conceive social situations among competing players and produce optimal decision-making of independent and competing actors in a strategic setting. The actions and choices of all the competing actors (participants or players) affect the outcome of each other [12]. The competing actors within the game are assumed to be rational and will do anything to maximize their payoffs in the game.

2.2.1 Scoreline of a game

The scoreline of a football match or a game is the score or the final result of it. Scoreline, type of competition, venue of the game, time frame, and the competitor have great effect on sports performance [13-15].
2.2.2 Two-person zero-sum games

These games involve only two players. Two-Person Zero-Sum Games are called Zero-Sum Games if the sum of the payoffs to each player is constant for all possible outcomes of the game. A Two-Person Game is characterized by the strategies of each player and the payoff matrix. In the Two-Person Zero-Sum Games, the players are rational and greedy, meaning that they choose their strategies in their own interest. Hence, the winning of one player leads to the loss of the other player.

2.2.3 Payoff matrix

This matrix shows the gain (positive or negative) for the first player (Player A) that would result from each combination of strategies for two players (Player A and Player B), where the matrix for Player B is the negative of the matrix for Player A in a Zero-Sum Game. The entries payoff matrix can be in any units as long as they represent the value to the player.

2.2.4 Saddle points and value of the game

The Saddle Point is that value in the payoff matrix which is both the minimum of the column maximums and the maximum of the row minimum. When there exists only one saddle point, the saddle point so obtained becomes the Value of the Game to Player A. Table 1 shows the payoff table format for a game involving more chances for one player.

| Player B | Player A |
|----------|----------|
|          | $Q_1$    | $Q_2$    | \ldots | $Q_q$ |
| $B_1$    | $x_{11}$ | $x_{12}$ | \ldots | $x_{1q}$ |
| $B_2$    | $x_{21}$ | $x_{22}$ | \ldots | $x_{2q}$ |
| \vdots   | \vdots   | \vdots   | \ddots | \vdots   |
| $B_p$    | $x_{p1}$ | $x_{p2}$ | \ldots | $x_{pq}$ |

where $A_1, A_2, \ldots, A_p$ are Player’s A Strategies; $B_1, B_2, \ldots, B_q$ are Player’s B Strategies

$P_1, P_2, \ldots, P_p$ are the probabilities of Player A choosing the respective Strategies $A_1, A_2, \ldots, A_p$

$Q_1, Q_2, \ldots, Q_q$ are the probabilities of Player B choosing the respective Strategies $B_1, B_2, \ldots, B_q$

$x_{11}, x_{12}, \ldots, x_{pq}$ are the Goal Difference between Player A and Player B in a Home or Away Games

2.3 Games and Linear Programming

When the Two-Person Zero-Sum Game involves more than two (2) choices and there is no evidence of Saddle Points, linear programming can be used to established the opposing player’s strategies. In Table 1, Player A will choose strategies $A_1, A_2, \ldots, A_p$ so as to receive as large an expected payoff as possible, no matter what Player B does.

Player A’s Complete Linear Program is given as

Maximize $V = $ Expected Payoff

S.t $x_{11}P_1 + x_{21}P_2 + \cdots + x_{p1}P_p \geq V$

$x_{12}P_1 + x_{22}P_2 + \cdots + x_{p2}P_p \geq V$

$x_{1q}P_1 + x_{2q}P_2 + \cdots + x_{pq}P_p \geq V$

$P_1 + P_2 + \cdots + P_p = 1$
Where \( P_1, P_2, \ldots, P_p \geq 0 \)

On the other hand, Player B will choose strategies \( B_1, B_2, \ldots, B_q \) so that his expected loss to Player A is as small as possible, no matter which strategy Player A chooses.

Player B’s Complete Linear Program is given as

Minimize \( V = \text{Expected Loss} \)

S.t \( x_{11} Q_1 + x_{12} Q_2 + \cdots + x_{1q} Q_q \leq V \)
\( x_{21} Q_1 + x_{22} Q_2 + \cdots + x_{2q} Q_q \leq V \)
\( \cdots \)
\( x_{pq} Q_1 + x_{pq} Q_2 + \cdots + x_{pq} Q_q \leq V \)

\( Q_1 + Q_2 + \cdots + Q_q = 1 \)

Where \( Q_1, Q_2, \ldots, Q_q \geq 0 \)

Dividing equation (2) and equation (3) by \( V \) which is the value of the game, then Player A’s new Linear Program becomes

Minimize \( \frac{1}{V} = \frac{P_1}{V} + \frac{P_2}{V} + \cdots + \frac{P_p}{V} \)

S.t \( x_{11} \frac{P_1}{V} + x_{12} \frac{P_2}{V} + \cdots + x_{1q} \frac{P_q}{V} \geq 1 \)
\( x_{21} \frac{P_1}{V} + x_{22} \frac{P_2}{V} + \cdots + x_{2q} \frac{P_q}{V} \geq 1 \)
\( \cdots \)
\( x_{pq} \frac{P_1}{V} + x_{pq} \frac{P_2}{V} + \cdots + x_{pq} \frac{P_q}{V} \geq 1 \)

\( \frac{P_1}{V} + \frac{P_2}{V} + \cdots + \frac{P_p}{V} = 1 \)

Where \( \frac{P_1}{V}, \frac{P_2}{V}, \ldots, \frac{P_p}{V} \geq 0 \)

Since in equation (4), Player A’s objective is to maximize the value of the game on his interest, it is equivalent to minimize \( \frac{1}{V} \).

Player B’s Linear Program is given as

Maximize \( \frac{1}{V} = \frac{Q_1}{V} + \frac{Q_2}{V} + \cdots + \frac{Q_q}{V} \)

S.t \( x_{11} \frac{Q_1}{V} + x_{12} \frac{Q_2}{V} + \cdots + x_{1q} \frac{Q_q}{V} \leq 1 \)
\( x_{21} \frac{Q_1}{V} + x_{22} \frac{Q_2}{V} + \cdots + x_{2q} \frac{Q_q}{V} \leq 1 \)
\( \cdots \)
\( x_{pq} \frac{Q_1}{V} + x_{pq} \frac{Q_2}{V} + \cdots + x_{pq} \frac{Q_q}{V} \leq 1 \)

\( \frac{Q_1}{V} + \frac{Q_2}{V} + \cdots + \frac{Q_q}{V} = 1 \)

Where \( \frac{Q_1}{V}, \frac{Q_2}{V}, \ldots, \frac{Q_q}{V} \geq 0 \)

Again, since Player B’s objective in equation (5) is to minimize his loss, which is equivalent to maximize \( \frac{1}{V} \).

2.4 Markov Chain

A Markov chain is a stochastic process \( X_n, n = 0, 1, 2, \ldots \) having the property that given the present state, the future is conditionally independent of the past. Markov chain is a discrete sequence of states, each drawn from a discrete state space (finite or not), and that follows the Markov property. Mathematically, Markov Chain can be denoted by
\[ X = (X_n)_{n \in N} = (X_0, X_1, X_2, \ldots) \]  

where at each instant of time the process takes its values in a discrete set \( E \) such that

\[ X_n \in E; \forall n \in N \]  

Suppose that whenever the process is in state \( i \), there is a fixed probability \( P_{ij} \) that it will be in state \( j \) next, then the property of Markov chains is defined as

\[ P(X_{n+1} = j | X_n = i, X_{n-1} = i_{n-1}, \ldots, X_0 = i_0) = P_{ij} \]  

for all states \( i_0, i_1, \ldots, i_{n-1}, i, j \) and all \( n \geq 0 \)

where \( j \) is the future state, \( i \) is the present or current state, \( i_{n-1}, i_{n-2}, \ldots, i_1, i_0 \) are the past states.

Since probabilities are nonnegative and the process must make a transition into some state, the sum of the probabilities will be equal to 1, for the probability of any transition is greater than zero (0). It is expressed as

\[ \sum_{j=0}^{\infty} P_{ij} = 1, \text{for } i = 0, 1, 2, \ldots \]  

\[ P_{ij} \geq 0, \text{for } i, j \geq 0 \]  

2.4.1 Transition Probability Matrix (TPM)

The state transition probability matrix of a Markov chain gives the probabilities of transitioning from one state to another in a single time unit. Let \( P \) represent the matrix of one-step transition probabilities \( P_{ij} \), with states 1, 2, \ldots, \( r \), which is expressed as

\[ P = \begin{pmatrix} p_{11} & p_{12} & \cdots & p_{1r} \\ p_{21} & p_{22} & \cdots & p_{2r} \\ \vdots & \vdots & \ddots & \vdots \\ p_{r1} & p_{r2} & \cdots & p_{rr} \end{pmatrix} \]  

For example, a one-step probability transition matrix with three states say Loss (L), Win (W) and Draw (D) is given as

\[ P = \begin{pmatrix} L & W & D \\ L & P_{LL} & P_{LD} \\ W & P_{WL} & P_{WD} \\ D & P_{DL} & P_{DD} \end{pmatrix} \]  

Suppose that a game leads to the outcome in equation (13), the chance of a team (club) winning their next game depends only upon whether or not they win their current game, they will win their next game with probability \( P_{WW} \); and if they do not win their current (that is, they lose their current game), their probability of winning their next game will be \( P_{WL} \). However,

\[ p_{LL} + p_{LW} + p_{LD} = 1, \; p_{WL} + p_{WW} + p_{WD} = 1 \]  

\[ p_{DL} + p_{DW} + p_{DD} = 1 \]  

Let the n-step transition probabilities of a Markov chain \( P_{(n)}^{ij} \) to be the probability that a process proceeds from state \( i \) to state \( j \) over \( n \) additional steps. This can be written as

\[ P_{(n)}^{ij} = P(X_{n+1} = j | X_n = i) \]
for \( n \geq 0, i, j \geq 0 \)

For a finite number \( n \) of possible states in \( i = \{i_1, i_2, \ldots, i_n\} \), the transition probability matrices up to \( m \) step transitions are given as

\[
\begin{align*}
(p^1)_{ij} &= P(X_{n+1} = j | X_n = i) \\
(p^2)_{ij} &= P(X_{n+2} = j | X_n = i) \\
(p^3)_{ij} &= P(X_{n+3} = j | X_n = i) \\
&\vdots \\
(p^m)_{ij} &= P(X_{n+m} = j | X_n = i)
\end{align*}
\]

(15)

The \( n \)-step transition probabilities can be computed or derived using Chapman-Kolmogorov equation as stated in equation (16)

\[
(p^n)_{ij}^{(n+m)} = \sum_{k=0}^{\infty} p^n_{ik} p^m_{kj}
\]

(16)

for all \( n, m \geq 0 \), and for all \( i, j \)

where \( p^n_{ik} p^m_{kj} \) represents the probability that starting in \( i \), the process will go to state \( j \) in \( n + m \) transitions through a path which takes it into state \( k \) at the \( nt \) transition.

2.5 Prediction Model

The prediction of matches over the seasons can be obtained using Stochastic Row Vector, where at the current season, \( x^0 = (0 \ 1 \ 0) \), which implies that the home Club has a probability of one (1), that is 100% chances of winning the visiting Club.

For the next season, that is at time \( n = 1 \), the prediction can be obtained using

\[
x^{(1)} = x^{(0)} p
\]

(17)

In the next two seasons, that is at time \( n = 2 \), the prediction can be obtained using

\[
x^{(2)} = x^{(1)} p = x^{(0)} p^2
\]

(18)

And in the next three, that is at time \( n = 3 \), the prediction can be obtained using

\[
x^{(3)} = x^{(2)} p = x^{(0)} p^3
\]

(19)

In general, the Prediction Model is written as

\[
x^{(n)} = x^{(n-1)} p = x^{(0)} p^n
\]

(20)

Prediction for the games on more distant seasons change less and less on each subsequent season and tend towards a steady state vector [16] as defined in equation (21).

\[
q = \lim_{n \to \infty} x^n
\]

(21)

3 Results and Discussion

The seven clubs that have never relegated to English Championship League from 2005/2006 to 2019/2020 seasons selected based on proportion of their wins are shown in Table 2.
Table 2. Proportion wins of English Premier League teams/clubs that have not been relegated to Championship between 2005/2006 to 2019/2020 seasons

| Team (Club) | Total Game Played | Total Win (both Home and Away) | Proportion Win |
|-------------|-------------------|---------------------------------|---------------|
| Man U (MU)  | 570               | 352                             | 0.62          |
| Chelsea (CH)| 570               | 346                             | 0.61          |
| Man City (MC)| 570              | 327                             | 0.57          |
| Liverpool (LP)| 570               | 321                             | 0.56          |
| Arsenal (AR)| 570               | 310                             | 0.54          |
| Tottenham (T)| 570               | 285                             | 0.50          |
| Everton (EV)| 570               | 227                             | 0.40          |

Man U, Chelsea, Man City, Liverpool and Arsenal are the teams with the highest proportion wins. However, they are the teams (clubs) of interest in this study. Table 3 shows the goal difference among the five selected clubs in the EPL, in the home and away games.

Table 3. 2005/2006 home and away results matrix

|        | Man U | Chelsea | Man City | Liverpool | Arsenal |
|--------|-------|---------|----------|-----------|---------|
| Man U  | -     | 1-0     | 1-1      | 1-0       | 2-0     |
| Chelsea| 3-0   | -       | 2-0      | 2-0       | 1-0     |
| Man City| 0-0    | 0-1     | -        | 0-1       | 1-3     |
| Liverpool| 0-0    | 1-4     | 1-0      | -         | 1-0     |
| Arsenal| 0-0   | 0-2     | 1-0      | 2-1       | -       |

Table 3 shows the match results between the competing teams in 2005/2006 English Premier League. For example, when Man U is home and Chelsea is Away, the resultant score is 1-0 in favour of Man U; when Man U is still Home and Man City is Away, the result score is 1-1, which shows a tie (draw); when Liverpool is Away and Man U still home, the resultant score is 1-0 in favour of Man U; and when Arsenal is Away, the resultant score is 2-0 in favour of Man U. However, a team cannot play with itself, in that case, a dash (–) is used instead.

Table 4 shows the goal difference of the games in Table 3, with two Players A and B using the clubs (or teams) as their strategies. A positive goal difference implies a win to the home teams and a loss to the away team. For example, when Arsenal is at Home, and Man U is Away, the goal difference is 0, which implies a draw game, each team (both Arsenal and Man U) goes home with a point; when Chelsea is Away, the goal difference is -2, which implies a loss to Arsenal and a win to Chelsea; when Man City is Away, the goal difference is 1, which implies a win to Arsenal and a loss to Man City; when Liverpool is Away, the goal difference is 1, which implies a win to Arsenal and a loss to Liverpool; and since Arsenal cannot play itself, the goal difference is 1.

Again, Player A chooses his strategies (clubs) so as to maximize his game, while in the other hand, Player B will choose his strategies (clubs) so as to minimize his loss.

Table 4. 2005/2006 home and away goal difference (payoff) matrix

|        | Man U | Chelsea | Man City | Liverpool | Arsenal |
|--------|-------|---------|----------|-----------|---------|
| Player A |       |         |          |           |         |
| Man U  | 1     | 1       | 0        | 1         | 2       |
| Chelsea| 3     | 1       | 2        | 2         | 1       |
| Man City| 0     | -1      | 1        | -1        | -2      |
| Liverpool| 0     | -3      | 1        | 1         | 1       |
| Arsenal| 0     | -2      | 1        | 1         | 1       |

Player A’s Linear Program is given as
\[
\begin{align*}
\text{Maximize} & \quad V = P_1 + P_2 + P_3 + P_4 + P_5 \nonumber \\
\text{Subject to} & \quad - P_1 - P_2 - P_4 - 2P_5 \leq 0 \\
& \quad -3P_1 - P_2 - 2P_3 - 2P_4 - P_5 \leq 0 \\
& \quad P_2 - P_3 + P_4 + 2P_5 \leq 0 \\
& \quad 3P_2 - P_3 - P_4 - P_5 \leq 0 \\
& \quad 2P_2 - P_3 - P_4 - P_5 \leq 0 \\
\end{align*}
\]

where \( P_1 + P_2 + P_3 + P_4 + P_5 \geq 0 \)\(^{(22)}\)

The optimal solutions of player A and player B with the value of the game for player A for 2005/2006 season English Premier League (EPL) home and away games is shown in Table 5.

Table 5. Optimal solution for the season 2005/2006 EPL home and away games

| Player A’s Optimal Strategies | MU | CH | MC | LP | AR |
|--------------------------------|----|----|----|----|----|
| Probability                   | 0.50 | 0.50 | 0.00 | 0.00 | 0.00 |
| Player B’s Optimal Strategies | MU | CH | MC | LP | AR |
| Probability                   | 0.00 | 1.00 | 0.00 | 0.00 | 0.00 |

Table 5 shows the optimal solution of the LP in equation (17). For Player A to maximize his gain, he selects two strategies MU (Man U) or CH (Chelsea), both having a selection probability of 0.50. Meanwhile Player B minimizes his loss by selecting strategy CH (Chelsea) having a selection probability of 1.00. Again, Chelsea is considered the best club among others. The value of the Game to Player A is 1.00. Moreover, the analysis as in Table 3, Table 4, Table 5, and LP formation for Player A in equation (22) is also carried out for the seasons 2006/2007 to 2019/2020 Home/Away games, so as to obtain the optimal solutions, with the values of the game to Player A. The summary of the optimal strategies and values of the games for Player A is given in Table 6.

Table 5 shows the optimal solution of the LP in equation (17). For Player A to maximize his gain, he selects two strategies MU (Man U) or CH (Chelsea), both having a selection probability of 0.50. Meanwhile Player B minimizes his loss by selecting strategy CH (Chelsea) having a selection probability of 1.00. Again, Chelsea is considered the best club among others. The value of the Game to Player A is 1.00. Moreover, the analysis as in Table 3, Table 4, Table 5, and LP formation for Player A in equation (22) is also carried out for the seasons 2006/2007 to 2019/2020 Home/Away games, so as to obtain the optimal solutions, with the values of the game to Player A. The summary of the optimal strategies and values of the games for Player A is given in Table 6.

Table 6. Optimal strategies for player A and player B for Seasons 2005/2006 to 2019/2020 EPL home and away games

| Season      | Player A Strategy | Probability | Player B Strategy | Probability | Value of the Game to Player A |
|-------------|-------------------|-------------|------------------|-------------|-----------------------------|
| 2005/2006   | MU, CH            | 0.50 (each) | CH               | 1.00        | 1.00                        |
| 2006/2007   | MU                | 0.50        | MU, CH           | 0.50 (each) | 0.50                        |
| 2007/2008   | MU                | 0.71        | MU               | 1.00        | 1.00                        |
| 2008/2009   | LP                | 0.68        | MC               | 0.38        | 0.54                        |
| 2009/2010   | MU, MC            | 0.38 (each) | MU, CH           | 0.38 (each) | 0.25                        |
| 2010/2011   | MU                | 1.00        | LP               | 0.56        | 1.00                        |
| 2011/2012   | MC                | 1.00        | MC               | 0.67        | 1.00                        |
| 2012/2013   | MU, MC            | 0.50 (each) | MU, MC           | 0.50 (each) | 0.00                        |
| 2013/2014   | CH                | 1.00        | CH               | 1.00        | 1.00                        |
| 2014/2015   | CH                | 0.67        | CH               | 0.67        | 0.67                        |
| 2015/2016   | MU                | 0.59        | CH               | 0.37        | 0.58                        |
| 2016/2017   | LP                | 0.69        | LP               | 0.50        | 0.58                        |
| 2017/2018   | LP                | 0.67        | MU, CH, MC       | 0.33 (each) | 0.33                        |
| 2018/2019   | MC                | 1.00        | MC, LP           | 0.50 (each) | 1.00                        |
| 2019/2020   | LP                | 0.83        | MU, LP           | 0.50 (each) | 1.50                        |

The frequency (number of times) clubs are being selected by player A and player B, with the proportion (relative frequency) are shown in Table 7.

As given in Table 7 Player A should choose or purchase Man U so as to maximize his profit or maximize his winning potentials in EPL, meanwhile Player B has to choose or purchase Chelsea so as to minimize his loss or his losing deficiency in the EPL. Table 8 gives the game results sequences, matrix of flow and 1-step transition matrix for both home and away games for Man U and Chelsea.
Table 7. Number of times clubs are selected by player A and player B in season 2005/2006 to 2019/2020 EPL home and away games

| Clubs/Teams | Player A | | | Player B |
|-------------|----------|---|---|----------|
|             | Number of Selection | Proportion | Number of Selection | Proportion |
| Man U (MU)  | 7        | 0.47 | 6 | 0.40 |
| Chelsea (CH)| 3        | 0.20 | 7 | 0.47 |
| Man City (MC)| 4      | 0.27 | 5 | 0.33 |
| Liverpool (LP)| 4      | 0.27 | 4 | 0.27 |
| Arsenal (AR)| 0        | 0   | 0 | 0    |

Table 8. Game results sequences, matrix of flow and 1-Step TPM for 2005/2006 to 2019/2020 EPL home and away games for Man U and Chelsea

| Game Results Sequences                      | Matrix of Flow |
|---------------------------------------------|----------------|
| Man U against Chelsea (when Man U is Home)  |                |
| WDWWLWWLDDDDWWDW ⇒                         | L 0 1 1 = 2    |
|                                             | W 2 3 2 = 7    |
|                                             | D 0 3 2 = 5    |
|                                             | 0.000 0.500 0.500 |
|                                             | 0.290 0.420 0.290 |
|                                             | 0.000 0.600 0.400 |
| Chelsea against Man U (when Chelsea is Home)| L 0 1 0 = 1    |
| WDWDDWDLWWDDWDL ⇒                         | W 0 3 5 = 8    |
|                                             | D 2 3 0 = 5    |
|                                             | 0.000 1.000 0.000 |
|                                             | 0.000 0.375 0.625 |
|                                             | 0.000 0.600 0.400 |
| Man U against Man City (when Man U is Home) | L 4 3 0 = 7    |
| DWLWWWLLLWDDLLWL ⇒                         | W 2 2 1 = 5    |
|                                             | D 1 1 0 = 2    |
|                                             | 0.570 0.430 0.000 |
|                                             | 0.400 0.400 0.200 |
|                                             | 0.500 0.500 0.000 |
| Man City against Man U (when Man City is Home) | L 1 3 2 = 6    |
| DLWLLDWLWWLWLDLWL ⇒                         | W 4 1 0 = 5    |
|                                             | D 2 1 0 = 3    |
|                                             | 0.170 0.500 0.330 |
|                                             | 0.800 0.200 0.000 |
|                                             | 0.670 0.330 0.000 |
| Man U against Liverpool (when Man U is Home) | L 0 2 0 = 2    |
| WWWLWWWLWWWLWDDDD ⇒                         | W 2 6 2 = 10   |
|                                             | D 0 1 1 = 2    |
Liverpool against Man U (when Liverpool is Home)

$$D L L W W W D W L W L D W W \Rightarrow$$

$$\begin{pmatrix}
0.400 & 0.400 & 0.200 \\
0.200 & 0.600 & 0.200 \\
0.500 & 0.250 & 0.250
\end{pmatrix}$$

$$L \begin{pmatrix} 2 \\ 1 \end{pmatrix} W \begin{pmatrix} 2 \\ 1 \end{pmatrix} D = 5$$

Man U against Arsenal (when Man U is Home)

$$W L W D W W W W D W D W D D \Rightarrow$$

$$\begin{pmatrix}
0.000 & 1.000 & 0.000 \\
0.110 & 0.440 & 0.440 \\
0.000 & 0.750 & 0.250
\end{pmatrix}$$

$$L \begin{pmatrix} 0 \\ 3 \end{pmatrix} W \begin{pmatrix} 2 \\ 1 \end{pmatrix} D = 9$$

Arsenal against Man U (when Arsenal is Home)

$$D W D W L W L D L W W \Rightarrow$$

$$\begin{pmatrix}
0.000 & 0.750 & 0.250 \\
0.500 & 0.330 & 0.170 \\
0.250 & 0.500 & 0.250
\end{pmatrix}$$

$$L \begin{pmatrix} 0 \\ 3 \end{pmatrix} W \begin{pmatrix} 2 \\ 1 \end{pmatrix} D = 4$$

Chelsea against Man City (when Chelsea is Home)

$$W W W W L W W W D W D L W L W \Rightarrow$$

$$\begin{pmatrix}
0.000 & 1.000 & 0.000 \\
0.220 & 0.560 & 0.220 \\
0.500 & 0.500 & 0.000
\end{pmatrix}$$

$$L \begin{pmatrix} 0 \\ 3 \end{pmatrix} W \begin{pmatrix} 2 \\ 1 \end{pmatrix} D = 3$$

Man City against Chelsea (when Man City is Home)

$$L L L W W W W L D W L W W \Rightarrow$$

$$\begin{pmatrix}
0.500 & 0.330 & 0.170 \\
0.290 & 0.710 & 0.000 \\
0.000 & 1.000 & 0.000
\end{pmatrix}$$

$$L \begin{pmatrix} 3 \\ 2 \end{pmatrix} W \begin{pmatrix} 2 \\ 1 \end{pmatrix} D = 6$$

Chelsea against Liverpool (when Chelsea is Home)

$$W W D L W L D W D L W D L \Rightarrow$$

$$\begin{pmatrix}
0.400 & 0.400 & 0.200 \\
0.200 & 0.200 & 0.600 \\
0.750 & 0.250 & 0.250
\end{pmatrix}$$

$$L \begin{pmatrix} 2 \\ 2 \end{pmatrix} W \begin{pmatrix} 2 \\ 1 \end{pmatrix} D = 5$$

Liverpool against Chelsea (when Liverpool is Home)
\[ L \begin{pmatrix} W & D \end{pmatrix} = 4 \]
\[ L \begin{pmatrix} 1 & 2 & 1 \end{pmatrix} = 4 \]
\[ W \begin{pmatrix} 1 & 2 & 2 \end{pmatrix} = 5 \]
\[ D \begin{pmatrix} 1 & 2 & 2 \end{pmatrix} = 5 \]

**Chelsea against Arsenal (when Chelsea is Home)**

\[ LWDLWWWDDDDWW \Rightarrow \]
\[ \begin{pmatrix} 0.250 & 0.500 & 0.250 \\ 0.200 & 0.400 & 0.400 \\ 0.200 & 0.400 & 0.400 \end{pmatrix} \]

**Arsenal against Chelsea (when Arsenal is Home)**

\[ LDWLWDDLDDWWLWD \Rightarrow \]
\[ \begin{pmatrix} 0.200 & 0.400 & 0.400 \\ 0.500 & 0.000 & 0.500 \\ 0.400 & 0.400 & 0.200 \end{pmatrix} \]

### 3.1 Predictions on Man U vs Chelsea Games (when Man U is Playing Home) in EPL from 2020/2021 to 2023/2024

The prediction of Man U vs Chelsea in 2019/2020 EPL is the initial state vector \( x^0 = \begin{pmatrix} 0 & 1 & 0 \end{pmatrix} \), which implies that Man U has a 100% chances of winning Chelsea.

In 2020/2021 Man U vs Chelsea Game (Man U is playing home), the prediction is obtained as

\[ x^{(1)} = x^{(0)}p = \begin{pmatrix} 0.000 & 0.500 & 0.500 \\ 0.290 & 0.420 & 0.290 \\ 0.000 & 0.600 & 0.400 \end{pmatrix} \]
\[ = \begin{pmatrix} 0.290 & 0.420 & 0.290 \end{pmatrix} \] (23)

In 2021/2022 Man U vs Chelsea Game (Man U is playing home), the prediction is obtained as

\[ x^{(2)} = x^{(1)}p \]
\[ = \begin{pmatrix} 0.290 & 0.420 & 0.290 \end{pmatrix} \begin{pmatrix} 0.000 & 0.500 & 0.500 \\ 0.290 & 0.420 & 0.290 \\ 0.000 & 0.600 & 0.400 \end{pmatrix} \]
\[ = \begin{pmatrix} 0.122 & 0.495 & 0.383 \end{pmatrix} \] (24)

In 2022/2023 Man U vs Chelsea Game (Man U is playing home), the prediction is obtained as

\[ x^{(3)} = x^{(2)}p \]
\[ = \begin{pmatrix} 0.122 & 0.495 & 0.383 \end{pmatrix} \begin{pmatrix} 0.000 & 0.500 & 0.500 \\ 0.290 & 0.420 & 0.290 \\ 0.000 & 0.600 & 0.400 \end{pmatrix} \]
\[ = \begin{pmatrix} 0.144 & 0.499 & 0.358 \end{pmatrix} \] (25)

In 2023/2024 Man U vs Chelsea Game (Man U is playing home), the prediction is obtained as
\[ x^{(4)} = x^{(3)}p \]
\[ = (0.144 \ 0.499 \ 0.358) \begin{pmatrix} 0.000 & 0.500 & 0.500 \\ 0.290 & 0.420 & 0.290 \\ 0.000 & 0.600 & 0.400 \end{pmatrix} \]
\[ = (0.145 \ 0.496 \ 0.360) \]

(26)

The process in equations 25, 26, 27 and 28 is repeated for the games between Chelsea vs Man U (when Chelsea is playing home). Table 9 gives the predictions for 2020/2021, 2021/2022, and 2022/2023 EPL for Man U vs Chelsea, and Chelsea vs Man U.

Table 9. Predictions on Man U vs Chelsea, and Chelsea vs Man U Games for 2020/2021, 2021/2022 and 2022/2023 EPL

| Matches          | EPL Prediction |
|------------------|----------------|
|                  | 2020/2021 | 2021/2022 | 2022/2023 | 2023/2024 |
| Man U vs Chelsea | (0.290 0.420 0.290) | (0.122 0.495 0.383) | (0.144 0.499 0.358) | (0.145 0.496 0.360) |
| Chelsea vs Man U | (0.000 0.375 0.625) | (0.000 0.516 0.484) | (0.000 0.484 0.516) | (0.000 0.491 0.509) |

Table 9 shows that in 2020/2021 EPL season, as Man U is playing home Man U has 42% chances of winning Chelsea, 29% chances of losing to Chelsea and 29% chances of drawing with Chelsea; in 2021/2022 EPL season, Man U has 49.5% chances of winning Chelsea, 12.2% chances of losing to Chelsea and 38.3% chances of drawing with Chelsea; in 2022/2023 EPL season, Man U has 49.9% chances of winning Chelsea, 14.4% chances of losing to Chelsea and 35.8% chances of drawing with Chelsea, and in 2023/2024 EPL season, Man U has 49.6% chances of winning Chelsea, 14.5% chances of losing to Chelsea and 36% chances of drawing with Chelsea.

When Chelsea is playing against Man U, while at home, in 2020/2021 EPL season Chelsea has 37.5% chances of winning Man U, 0% chances of losing to Man U and 62.5% chances of drawing with Man U; in 2021/2022 EPL season, Chelsea has 51.6% chances of winning Man U, 0% chances of losing to Man U and 48.4% chances of drawing with Man U; in 2022/2023 EPL season, Chelsea has 48.4% chances of winning Man U, 0% chances of losing to Man U and 51.6% chances of drawing with Man U, and in 2023/2024 EPL season, Chelsea has 49.1% chances of winning Man U, 0% chances of losing to Man U and 50.9% chances of drawing with Man U.

4 Conclusion

The aim of this paper is to obtain the optimal strategies and to make future predictions based on the game theory and Markov chains involving the EPL. After examining and analyzing the scorelines, it is best that Player A should purchase Man U for profit maximization and Player B should purchase Chelsea for loss minimization. Man U has a greater winning chances against Chelsea while playing home and a lesser chance of losing to Chelsea, while Chelsea in the other hand, has a very low chance of low of losing to Man U, but a greater chance of drawing with Man U, which implies that Man U is better at home than Chelsea. However, based on the findings, the winning potential of Chelsea against Man U increases as EPL Season increases, while Man U’s winning potential fluctuates.

Competing Interests

Authors have declared that no competing interests exist.

References

[1] Ngonadi LO, Ezemma GC, Etaga HO, Ugoh CI. English Premier League Scoreline Analysis: A Stochastic and Game Theory Approach, American Journal of Theoretical and Applied Statistics. 2021; 10(3):136-145.
[2] Sindik J, Vidak N. Application of Game Theory in Describing Efficacy of Decision making in Sportman’s Tactical Performance in Team Sports, Interdisciplinary Description of Complex System. 2008;6(1):53-66.

[3] Etaga HO, Umeokeke ET, Nwosu CR, Etaga NC, Umeh EU, Awopeju B, et al. Stochastic Modeling/Game Theory Analysis, International Journal of Pysical Science Research. 2017;1(2):1-13.

[4] Yajie Wan, Using the Markov Chain to Predict the Outcome of an Election in a given Situation, Journal of Pysical: Conference Series. 2021;1848:012051. DOI: 10.1088/1742-6596/1848/1/012051

[5] Tie Liu, Application of Markov Chains to Analyze and Predict the Time Series, Modern Applied Science. 2010;4:5.

[6] Xipeng Wang, Yuan Ma, Junru Di, Yi L. Murphey, Shiqi Qiu, Johannes Kristinsson, Jason Meyer, Finn Tseng, Timothy Feldkamp, Building Efficient Probability Transition Matrix using Machine Learning from Big Data for Personalized Route Prediction, Procedia Computer Science. 2015;53:284-291.

[7] Ryan Beal, Georgios Chalkiadakis, Timothy J. Norman and Sarvepali D. Ramchurn, Optimizing Game Tactics for Football, In Proc. Of the 19th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2020), Auckland, New Zealand. 2020, IFAAMAS, 9.

[8] Issah M, Douglas KB, Baba S. A comprehensive review of solution methods and techniques for solving games in game theory, Journal of Game Theory. 2020;9(2):25-31.

[9] Admed Boujnoui, Abdellah Zaaloul, Abdelkrim Haqiq, Mathematical model based on Game theory and Markov chains for analyzing the transmission cost in SA-ZD mechanism, International Journal of Computer Information Systems and Industrial Management Applications. 2018;10:197-207.

[10] Singh AP. Optimal solution strategy for games, International Journal of Computer Science and Engineering. 2010;2(9):2778-2782.

[11] Bhuiyan BA. An overview of game theory and some applications, Philosophy and Progress. 2016;LIX-LX:113-114.

[12] Kuman S, Reddy DSN. Graphical solution of nxm matrix of a game theory, European Journal of Operational Research. 1999;112(2):467-471.

[13] Clark P, O’Donoghue PG. Scoreline effect on work-rate in English FA Premier League Soccer, In Nunomw, H., rust, B. and Dawson, B. (Eds.), Science and Football VII, London: Routledge. 2013;175-180.

[14] Gomez MA, Lago-Penas C, Pollar R. Situational variables, In McGarry T, O’Donoghue PG, Sampaio J. (eds.), Routledge Hand of Sport Performance Analysis, London: Routledge. 2013;259-269.

[15] Taylor JB, Mellalieu SD, James N, Shearer UA. The influence of match location, quality of opposition and match status on technical performance in professional association football, Journal of Sports Sciences. 2008;26:885-895.

[16] Van Kampen NG. Stochastic processes in Physics and Chemistry, NL: North Holland Elsevier. 2007;73-95.