

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

**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. 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 20<sup>th</sup> 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 [8,10]. [4] analyzed the scoreline of the EPL using Game theory and Markov Chains. [16] carried out a prediction analysis on the outcome of an election in a given situation using the application of Markov chain. [13] studied the performance of Markov chain model in predicting time series data. [15] presented a machine learning algorithm for modelling route prediction using the application of Markov chains.

[9] 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. [6] analyzed the different methods and techniques used in solving game problems. [1] 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).

$$\text{Porportion of Club Wins} = \frac{\text{TCW}}{\text{TCP}} \quad (1)$$

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 [11], defined by the players, actions and outcomes [2].

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 [7]. The competing actors within the game are assumed to be

rational and will do anything to maximize their payoffs in the game.

### 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 [3,5,12].

### 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.

### 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.

### 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.

**Table 1: Payoff Table for a Game Involving More Chances for One Player**

			Player B			
			$Q_1$	$Q_2$	.....	$Q_q$
Player A	$P_1$	$A_1$	$x_{11}$	$x_{12}$	.....	$x_{1q}$
	$P_2$	$A_2$	$x_{21}$	$x_{22}$	.....	$x_{2q}$
	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
	$P_p$	$A_p$	$x_{p1}$	$x_{p2}$	.....	$x_{pq}$

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

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

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

$x_{11}, x_{12}, \dots, x_{pq}$  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  $A_1, A_2, \dots, A_p$  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

$$\begin{aligned} &\text{Maximize } V = \text{Expected Payoff} \\ &\text{S.t } x_{11}P_1 + x_{21}P_2 + \dots + x_{p1}P_p \geq V \end{aligned}$$

$$\begin{aligned} &x_{12}P_1 + x_{22}P_2 + \dots + x_{p2}P_p \geq V \\ &\vdots \\ &x_{1q}P_1 + x_{2q}P_2 + \dots + x_{pq}P_p \geq V \\ &P_1 + P_2 + \dots + P_p = 1 \\ &\text{Where } P_1, P_2, \dots, P_p \geq 0 \end{aligned} \quad (2)$$

On the other hand, Player B will choose strategies  $B_1, B_2, \dots, B_q$  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

$$\begin{aligned} &\text{Minimize } V = \text{Expected Loss} \\ &\text{S.t } x_{11}Q_1 + x_{12}Q_2 + \dots + x_{1q}Q_q \leq V \\ &x_{21}Q_1 + x_{22}Q_2 + \dots + x_{2q}Q_q \leq V \\ &\vdots \\ &x_{p1}Q_1 + x_{p2}Q_2 + \dots + x_{pq}Q_q \leq V \\ &Q_1 + Q_2 + \dots + Q_q = 1 \\ &\text{Where } Q_1, Q_2, \dots, Q_q \geq 0 \end{aligned} \quad (3)$$

Dividing equation (2) and equation (3) by  $V$  which is the value of the game, then

Player A's new Linear Program becomes

$$\begin{aligned}
& \text{Minimize } \frac{1}{V} = \frac{P_1}{V} + \frac{P_2}{V} + \dots + \frac{P_p}{V} \\
& \text{S.t } x_{11} \frac{P_1}{V} + x_{21} \frac{P_2}{V} + \dots + x_{p1} \frac{P_p}{V} \geq 1 \\
& \quad x_{12} \frac{P_1}{V} + x_{22} \frac{P_2}{V} + \dots + x_{p2} \frac{P_p}{V} \geq 1 \\
& \quad \vdots \\
& \quad x_{1q} \frac{P_1}{V} + x_{2q} \frac{P_2}{V} + \dots + x_{pq} \frac{P_p}{V} \geq 1 \\
& \quad \frac{P_1}{V} + \frac{P_2}{V} + \dots + \frac{P_p}{V} = \frac{1}{V} \\
& \text{Where } \frac{P_1}{V}, \frac{P_2}{V}, \dots, \frac{P_p}{V} \geq 0
\end{aligned} \tag{4}$$

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

$$\begin{aligned}
& \text{Maximize } \frac{1}{V} = \frac{Q_1}{V} + \frac{Q_2}{V} + \dots + \frac{Q_q}{V} \\
& \text{S.t } x_{11} \frac{Q_1}{V} + x_{12} \frac{Q_2}{V} + \dots + x_{1q} \frac{Q_q}{V} \leq 1 \\
& \quad x_{21} \frac{Q_1}{V} + x_{22} \frac{Q_2}{V} + \dots + x_{2q} \frac{Q_q}{V} \leq 1 \\
& \quad \vdots \\
& \quad x_{p1} \frac{Q_1}{V} + x_{p2} \frac{Q_2}{V} + \dots + x_{pq} \frac{Q_q}{V} \leq 1 \\
& \quad \frac{Q_1}{V} + \frac{Q_2}{V} + \dots + \frac{Q_q}{V} = \frac{1}{V} \\
& \text{Where } \frac{Q_1}{V}, \frac{Q_2}{V}, \dots, \frac{Q_q}{V} \geq 0
\end{aligned} \tag{5}$$

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  $X_n, n = 0, 1, 2, \dots$  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

$$X = (X_n)_{n \in N} = (X_0, X_1, X_2, \dots) \tag{6}$$

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

$$X_n \in E; \forall n \in N \tag{7}$$

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 define as

$$P(X_{n+1} = j | X_n = i, X_{n-1} = i_{n-1}, \dots, X_0 = i_0) = P_{ij} \tag{8}$$

for all states  $i_0, i_1, \dots, i_{n-1}, i, j$  and all  $n \geq 0$

where  $j$  is the future state,  $i$  is the present or current state,  $i_{n-1}, i_{n-2}, \dots, i_1, i_0$  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 zero (0). It is expressed as

$$\sum_{j=0}^{\infty} P_{ij} = 1, \text{ for } i = 0, 1, 2, \dots \tag{9}$$

$$P_{ij} \geq 0, \text{ for } i, j \geq 0 \tag{10}$$

### Transition Probability Matrix (TPM)

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

$$P = \begin{pmatrix} p_{11} & p_{12} & \dots & p_{1r} \\ p_{21} & p_{22} & \dots & p_{2r} \\ \vdots & \vdots & \ddots & \vdots \\ p_{r1} & p_{r2} & \dots & p_{rr} \end{pmatrix} \tag{11}$$

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

$$P = \begin{matrix} & \begin{matrix} L & W & D \end{matrix} \\ \begin{matrix} L \\ W \\ D \end{matrix} & \begin{pmatrix} p_{LL} & p_{LW} & p_{LD} \\ p_{WL} & p_{WW} & p_{WD} \\ p_{DL} & p_{DW} & p_{DD} \end{pmatrix} \end{matrix} \tag{12}$$

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 loss their current game), their probability of winning their next game will be  $p_{LW}$ . However,

$$\begin{aligned}
p_{LL} + p_{LW} + p_{LD} &= 1, p_{WL} + p_{WW} + p_{WD} = 1 \\
p_{DL} + p_{DW} + p_{DD} &= 1
\end{aligned} \tag{13}$$

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

$$P_{ij}^{(n)} = P(X_{n+1} = j | X_n = i) \tag{14}$$

$$\text{for } n \geq 0, i, j \geq 0$$

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

$$\left. \begin{aligned} p_{ij} &= P(X_{n+1} = j | X_n = i) \\ (p^2)_{ij} &= P(X_{n+2} = j | X_n = i) \\ (p^3)_{ij} &= P(X_{n+3} = j | X_n = i) \\ &\vdots \\ (p^m)_{ij} &= P(X_{n+m} = j | X_n = i) \end{aligned} \right\} \quad (15)$$

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

$$P_{ij}^{(n+m)} = \sum_{k=0}^{\infty} P_{ik}^n P_{kj}^m \quad (16)$$

for all  $n, m \geq 0$ , and for all  $i, j$

where  $P_{ik}^n P_{kj}^m$  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  $n$ th 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

$$x^{(1)} = x^{(0)}P \quad (17)$$

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

$$x^{(2)} = x^{(1)}P = x^{(0)}P^2 \quad (18)$$

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

$$x^{(3)} = x^{(2)}P = x^{(0)}P^3 \quad (19)$$

In general, the Prediction Model is written as

$$x^{(n)} = x^{(n-1)}P = x^{(0)}P^n \quad (20)$$

Prediction for the games on more distant seasons change less and less on each subsequent season and tend towards a steady state vector [14] as defined in equation (21).

$$q = \lim_{n \rightarrow \infty} x^n \quad (21)$$

## 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**

		Player B				
		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

$$\begin{aligned}
 & \text{Maximize} && V = P_1 + P_2 + P_3 + P_4 + P_5 \\
 & \text{Subject to} && -P_1 - P_2 - P_4 - 2P_5 \leq 0 \\
 & && -3P_1 - P_2 - 2P_3 - 2P_4 - P_5 \leq 0 \\
 & && P_2 - P_3 + P_4 + 2P_5 \leq 0 \\
 & && 3P_2 - P_3 - P_4 - P_5 \leq 0 \\
 & && 2P_2 - P_3 - P_4 - P_5 \leq 0 \\
 & \text{where} && P_1 + P_2 + P_3 + P_4 + P_5 \geq 0
 \end{aligned} \quad (22)$$

season English Premier League (EPL) home and away games is shown in Table 5.

The optimal solutions of player A and player B with the value of the game for player A for 2005/2006

**Table 5: Optimal Solution for the Season 2005/2006 EPL Home and Away Games**

Value of the Game to Player A= 1.00					
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 6: Optimal Strategies for Player A and Player B for Seasons 2005/2006 to 2019/2020 EPL Home and Away Games**

Season	Player A		Player B		Value of the Game to Player A
	Strategy	Probability	Strategy	Probability	
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.

**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

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 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	1-Step TPM (P)
<b>Man U against Chelsea (when Man U is Home)</b>		
WDWWLWLLDDWWDW ⇒	$\begin{matrix} L & W & D \\ L & \begin{pmatrix} 0 & 1 & 1 \end{pmatrix} = 2 \\ W & \begin{pmatrix} 2 & 3 & 2 \end{pmatrix} = 7 \\ D & \begin{pmatrix} 0 & 3 & 2 \end{pmatrix} = 5 \end{matrix}$	$\begin{pmatrix} 0.000 & 0.500 & 0.500 \\ 0.290 & 0.420 & 0.290 \\ 0.000 & 0.600 & 0.400 \end{pmatrix}$
<b>Chelsea against Man U (when Chelsea is Home)</b>		
WDWDWDLWWDWDL ⇒	$\begin{matrix} L & W & D \\ L & \begin{pmatrix} 0 & 1 & 0 \end{pmatrix} = 1 \\ W & \begin{pmatrix} 0 & 3 & 5 \end{pmatrix} = 8 \\ D & \begin{pmatrix} 2 & 3 & 0 \end{pmatrix} = 5 \end{matrix}$	$\begin{pmatrix} 0.000 & 1.000 & 0.000 \\ 0.000 & 0.375 & 0.625 \\ 0.000 & 0.600 & 0.400 \end{pmatrix}$
<b>Man U against Man City (when Man U is Home)</b>		

$$DWLWVWVLLVWVLLV \Rightarrow \begin{matrix} L & W & D \\ L & \begin{pmatrix} 4 & 3 & 0 \end{pmatrix} = 7 \\ W & \begin{pmatrix} 2 & 2 & 1 \end{pmatrix} = 5 \Rightarrow \\ D & \begin{pmatrix} 1 & 1 & 0 \end{pmatrix} = 2 \end{matrix} \quad \begin{pmatrix} 0.570 & 0.430 & 0.000 \\ 0.400 & 0.400 & 0.200 \\ 0.500 & 0.500 & 0.000 \end{pmatrix}$$

**Man City against Man U (when Man City is Home)**

$$DLVLLDVWVWVLDVWV \Rightarrow \begin{matrix} L & W & D \\ L & \begin{pmatrix} 1 & 3 & 2 \end{pmatrix} = 6 \\ W & \begin{pmatrix} 4 & 1 & 0 \end{pmatrix} = 5 \Rightarrow \\ D & \begin{pmatrix} 2 & 1 & 0 \end{pmatrix} = 3 \end{matrix} \quad \begin{pmatrix} 0.170 & 0.500 & 0.330 \\ 0.800 & 0.200 & 0.000 \\ 0.670 & 0.330 & 0.000 \end{pmatrix}$$

**Man U against Liverpool (when Man U is Home)**

$$WVWVWVWVWVWVWVWV \Rightarrow \begin{matrix} L & W & D \\ L & \begin{pmatrix} 0 & 2 & 0 \end{pmatrix} = 2 \\ W & \begin{pmatrix} 2 & 6 & 2 \end{pmatrix} = 10 \Rightarrow \\ D & \begin{pmatrix} 0 & 1 & 1 \end{pmatrix} = 2 \end{matrix} \quad \begin{pmatrix} 0.000 & 1.000 & 0.000 \\ 0.200 & 0.600 & 0.200 \\ 0.000 & 0.500 & 0.500 \end{pmatrix}$$

**Liverpool against Man U (when Liverpool is Home)**

$$DVLVWVWVLDVLLDVWV \Rightarrow \begin{matrix} L & W & D \\ L & \begin{pmatrix} 2 & 2 & 1 \end{pmatrix} = 5 \\ W & \begin{pmatrix} 1 & 3 & 1 \end{pmatrix} = 5 \Rightarrow \\ D & \begin{pmatrix} 2 & 1 & 1 \end{pmatrix} = 4 \end{matrix} \quad \begin{pmatrix} 0.400 & 0.400 & 0.200 \\ 0.200 & 0.600 & 0.200 \\ 0.500 & 0.250 & 0.250 \end{pmatrix}$$

**Man U against Arsenal (when Man U is Home)**

$$WLVWVWVWVWVWVWVWV \Rightarrow \begin{matrix} L & W & D \\ L & \begin{pmatrix} 0 & 1 & 0 \end{pmatrix} = 1 \\ W & \begin{pmatrix} 1 & 4 & 4 \end{pmatrix} = 9 \Rightarrow \\ D & \begin{pmatrix} 0 & 3 & 1 \end{pmatrix} = 4 \end{matrix} \quad \begin{pmatrix} 0.000 & 1.000 & 0.000 \\ 0.110 & 0.440 & 0.440 \\ 0.000 & 0.750 & 0.250 \end{pmatrix}$$

**Arsenal against Man U (when Arsenal is Home)**

$$DVLWVWVLDVLDVWVWV \Rightarrow \begin{matrix} L & W & D \\ L & \begin{pmatrix} 0 & 3 & 1 \end{pmatrix} = 4 \\ W & \begin{pmatrix} 3 & 2 & 1 \end{pmatrix} = 6 \Rightarrow \\ D & \begin{pmatrix} 1 & 2 & 1 \end{pmatrix} = 4 \end{matrix} \quad \begin{pmatrix} 0.000 & 0.750 & 0.250 \\ 0.500 & 0.330 & 0.170 \\ 0.250 & 0.500 & 0.250 \end{pmatrix}$$

**Chelsea against Man City (when Chelsea is Home)**

$$WVWVWVWVWVLDVWVWV \Rightarrow \begin{matrix} L & W & D \\ L & \begin{pmatrix} 0 & 3 & 0 \end{pmatrix} = 3 \\ W & \begin{pmatrix} 2 & 5 & 2 \end{pmatrix} = 9 \Rightarrow \\ D & \begin{pmatrix} 1 & 1 & 0 \end{pmatrix} = 2 \end{matrix} \quad \begin{pmatrix} 0.000 & 1.000 & 0.000 \\ 0.220 & 0.560 & 0.220 \\ 0.500 & 0.500 & 0.000 \end{pmatrix}$$

**Man City against Chelsea (when Man City is Home)**

$$LVLVWVWVLDVWVWV \Rightarrow \begin{matrix} L & W & D \\ L & \begin{pmatrix} 3 & 2 & 1 \end{pmatrix} = 6 \\ W & \begin{pmatrix} 2 & 5 & 0 \end{pmatrix} = 7 \Rightarrow \\ D & \begin{pmatrix} 0 & 1 & 0 \end{pmatrix} = 1 \end{matrix} \quad \begin{pmatrix} 0.500 & 0.330 & 0.170 \\ 0.290 & 0.710 & 0.000 \\ 0.000 & 1.000 & 0.000 \end{pmatrix}$$

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

$$WVLDVLLDVLDVLDV \Rightarrow \begin{matrix} L & W & D \\ L & \begin{pmatrix} 2 & 2 & 1 \end{pmatrix} = 5 \\ W & \begin{pmatrix} 1 & 1 & 3 \end{pmatrix} = 5 \Rightarrow \\ D & \begin{pmatrix} 3 & 1 & 0 \end{pmatrix} = 4 \end{matrix} \quad \begin{pmatrix} 0.400 & 0.400 & 0.200 \\ 0.200 & 0.200 & 0.600 \\ 0.750 & 0.250 & 0.000 \end{pmatrix}$$

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

$$LVLDVWVLDVLLDVWV \Rightarrow \begin{matrix} L & W & D \\ L & \begin{pmatrix} 1 & 2 & 1 \end{pmatrix} = 4 \\ W & \begin{pmatrix} 1 & 2 & 2 \end{pmatrix} = 5 \Rightarrow \\ D & \begin{pmatrix} 1 & 2 & 2 \end{pmatrix} = 5 \end{matrix} \quad \begin{pmatrix} 0.250 & 0.500 & 0.250 \\ 0.200 & 0.400 & 0.400 \\ 0.200 & 0.400 & 0.400 \end{pmatrix}$$

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

$$WDLWVWVWVWVWVWV \Rightarrow \begin{matrix} L & W & D \\ L & \begin{pmatrix} 0 & 2 & 0 \end{pmatrix} = 2 \\ W & \begin{pmatrix} 2 & 5 & 3 \end{pmatrix} = 10 \Rightarrow \\ D & \begin{pmatrix} 0 & 2 & 0 \end{pmatrix} = 2 \end{matrix} \quad \begin{pmatrix} 0.000 & 1.000 & 0.000 \\ 0.200 & 0.500 & 0.300 \\ 0.000 & 1.000 & 0.000 \end{pmatrix}$$

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

$$LDVLLDVLDVLDVWV \Rightarrow \begin{matrix} L & W & D \\ L & \begin{pmatrix} 1 & 2 & 2 \end{pmatrix} = 5 \\ W & \begin{pmatrix} 2 & 0 & 2 \end{pmatrix} = 4 \Rightarrow \\ D & \begin{pmatrix} 2 & 2 & 1 \end{pmatrix} = 5 \end{matrix} \quad \begin{pmatrix} 0.200 & 0.400 & 0.400 \\ 0.500 & 0.000 & 0.500 \\ 0.400 & 0.400 & 0.200 \end{pmatrix}$$

**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  $x^0 = (0 \ 1 \ 0)$ , 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

$$\begin{aligned} x^{(1)} &= x^{(0)}P = (0 \ 1 \ 0) \begin{pmatrix} 0.000 & 0.500 & 0.500 \\ 0.290 & 0.420 & 0.290 \\ 0.000 & 0.600 & 0.400 \end{pmatrix} \\ &= (0.290 \ 0.420 \ 0.290) \end{aligned} \quad (23)$$

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

$$\begin{aligned} x^{(2)} &= x^{(1)}P \\ &= (0.290 \ 0.420 \ 0.290) \begin{pmatrix} 0.000 & 0.500 & 0.500 \\ 0.290 & 0.420 & 0.290 \\ 0.000 & 0.600 & 0.400 \end{pmatrix} \\ &= (0.122 \ 0.495 \ 0.383) \end{aligned} \quad (24)$$

**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

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

$$\begin{aligned} x^{(3)} &= x^{(2)}P \\ &= (0.122 \ 0.495 \ 0.383) \begin{pmatrix} 0.000 & 0.500 & 0.500 \\ 0.290 & 0.420 & 0.290 \\ 0.000 & 0.600 & 0.400 \end{pmatrix} \\ &= (0.144 \ 0.499 \ 0.358) \end{aligned} \quad (25)$$

In 2023/2024 Man U vs Chelsea Game (Man U is playing home), the prediction is obtained as

$$\begin{aligned} x^{(4)} &= x^{(3)}P \\ &= (0.144 \ 0.499 \ 0.358) \begin{pmatrix} 0.000 & 0.500 & 0.500 \\ 0.290 & 0.420 & 0.290 \\ 0.000 & 0.600 & 0.400 \end{pmatrix} \\ &= (0.145 \ 0.496 \ 0.360) \end{aligned} \quad (26)$$

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

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. 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.

#### References

[1] Admed Boujnoui, Abdellah Zaaloul and

- 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, Vol. 10, Pp. 197-207
- [2] Bhuiyan, B. A., An overview of game theory and some applications, *Philosophy and Progress*, 2016, LIX-LX, 113-114.
- [3] Clark, P. and O'Donoghue, P. G. (2013), 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, pp. 175-180
- [4] Etaga, H. O., Umeokeke, E. T., Nwosu, C. R., Etaga, N. C., Umeh, E. U., Awopeju, B., Eriobu., Okoye, V. C. and Omoruyi, F. A., *Stochastic Modeling/Game Theory Analysis*, *International Journal of Physical Science Research*, 2017, Vol. 1, No. 2, Pp. 1-13
- [5] Gomez, M. A., Lago-Penas, C. and Pollard, R. (2013), Situational variables, In McGarry, T., O'Donoghue, P. G. and Sampaio, J. (eds.), *Routledge Hand of Sport Performance Analysis*, London: Routledge, pp. 259-269
- [6] Issah, M., Douglas, K. B. and 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
- [7] Kuman, S. and Reddy, D. S. N., Graphical solution of nxm matrix of a game theory, *European Journal of Operational Research*, 1999, 112(2), 467-471
- [8] Ngonadi, L. O., Ezemma, G. C., Etaga, H. O., and Ugoh, C. I., English Premier League Scoreline Analysis: A Stochastic and Game Theory Approach, *American Journal of Theoretical and Applied Statistics*, 2021, 10(3), 136-145
- [9] Ryan Beal, Georgios Chalkiadakis, Timothy J. Norman and Sarvepali D. Ramchurn, Optimizing Game Tactics for Football, In Proc. Of the 19<sup>th</sup> International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2020), Auckland, New Zealand, May 9-13, 2020, IFAAMAS, 9 pages
- [10] Sindik, J. and 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, Vol. 6, No.1, Pp. 53-66
- [11] Singh, A. P., Optimal solution strategy for games, *International Journal of Computer Science and Engineering*, 2010, 2(9), 2778-2782
- [12] Taylor, J. B., Mellalieu, S. D., James, N. and Shearer, U. A., 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
- [13] Tie Liu, Application of Markov Chains to Analyze and Predict the Time Series, *Modern Applied Science*, 2010, Vol. 4, No.5
- [14] Van Kampen, N. G. (2007), *Stochastic processes in Physics and Chemistry*, NL: North Holland Elsevier, pp. 73-95
- [15] XipengWang, 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, Vol 53, Pp. 284-291
- [16] Yajie Wan, Using the Markov Chain to Predict the Outcome of an Election in a given Situation, *Journal of Pysical: Conference Series*, 1848(2021) 012051, doi: 10.1088/1742-6596/1848/1/012051