# A Solution of a Mathematical Model Which Simulates Football Game as a Logistics Network

Kenan Mengüç[a]
*Beykent Unıversity, Applied Schools of Logistics, Akcaburgaz Neighborhood, 34522 Esenyurt, İstanbul*

Keywords: Transshipment Problem, Optimization, Spatio-temporal Data, Branch and Bound Algorıthm, Football, Logistic.

Abstract: Routing is very important in team sports as well as in many parts of daily life. On 01.07.2018, in the World Cup tournament organized in Russia, Spain lost the match with a penalty shootout although it performed a higher number of passes. As a result of the elimination of Spain despite its very high ball possession, the nature of modern football doesn't necessarily require the multiplicity of passes. It is important to know where the passes are done, rather than high amounts of them. This study has simulated a mathematical model by bringing a different interpretation to the categorization of pass throws.

## 1 INTRODUCTION

In recent years the interest of the world in football has been increasing. The economies formed in football organizations are developing many industrial sectors. Increased interest has enhanced the importance of every moment in the game. Many companies have made measurements about the tactics and the players within the football game. Since these measurement data is high confidence level, they give significant results for the game. In this study, a mathematical model was created in order to use these data more effectively. Similar scenarios are used between the logistics management and the football game for the model. The purpose of the study is that build a mathematical model by using this data set. Thus this mathematical model provide to determine the best attack route within the game for a team.

## 2 MANUSCRIPT PREPARATION

### 2.1 Literature

Over the last 50 years there has been a lot of publication on optimization in sport (Wright, 2009). The rules of the game in team sports were also effective in speeding up the game (Takeuchi et al., 2014). Football is a dynamic game as a course of its nature (Barghi, 2015). Football is a team game and the team's arrangement within the pitch is important. High level of communication among players is important for success (Grund, 2012). In addition the team arrangement in digital game industry and in robot game technology has been seen to have a serious impact on team results (Razykov, 2006).

The amount and importance of big data is increasing day by day. This big data take an important place in football. Many measurement companies make serious measurements of the game and players. Opta Sportsdata is an academically reliable company that makes these measurements (Liu et al., 2013). Opta Sportdata data was also used to determine a strategy for the football team (Stein et al., 2017). In addition, another study aimed at the best players selection for team squad (Tavana et al., 2013). In Gudmundsson and Horton's study, spatio-temporal data is used instead of classical percentage data (Gudmundsson and Horton, 2017). Yet another study has developed a clustering model for all time-spatial routes that have ended with a goal (Hirano and Tsumoto, 2005).

Passes between players takes an important place in football. In a study conducted for 4 major leagues in Europe, it was indicated that the teams with the high pass were in higher places in the score ranking (Cintia et al., 2015). The effect of the passes on the

[a] https://orcid.org/0000-0002-7536-2124



Mengüç, K.
A Solution of a Mathematical Model Which Simulates Football Game as a Logistics Network.
DOI: 10.5220/0007917801050112
In *Proceedings of the 7th International Conference on Sport Sciences Research and Technology Support (icSPORTS 2019)*, pages 105-112
ISBN: 978-989-758-383-4

Spain National Football team's game was told to be very important (Pena and Touchette, 2012). The passes also vary according to the way the ball tranships (Malqui, 2017). The problems of logistics and football are quite similar. Krishnan and Rao's transshipment model between two points in their study has been implemented in this study (Krishnan and Rao, 1965).

## 2.2 Model Creation

This study was started by reference to a transshipment problem model. After that the study was followed with regard to the steps below.

The data obtained were calculated in meters according to the measurements of the pitch and categorized for each player of Başakşehir F.K' squad.

↓

The measurement data obtained are separated into $x_1$, $y_1$, $x_2$ *and* $y_2$ coordinates.

↓

Four target directions have been determined and the data is categorized as forward, backward, right and left direction.

↓

Passes are categorized by a scoring method according to how far the ball is brought closer to the penalty zone.

↓

The general score has been created as a result of the analysis of all seasonal passes for each player in 4 different directions.

↓

General position of players in the pitch was determined according to the matches played by the players in a season.

↓

Player preferences have been determined for the squad of the team. According to the data we obtained in Step 6, players are assigned to zones.

↓

The two most successful options between possible players in the zone are determined to transshipment.

↓

In the scenario created, the model is intended to find the best combination of pass and player for the event of the ball passing through that zone during the match.

↓

The model is run for routes of three different positions of played match.

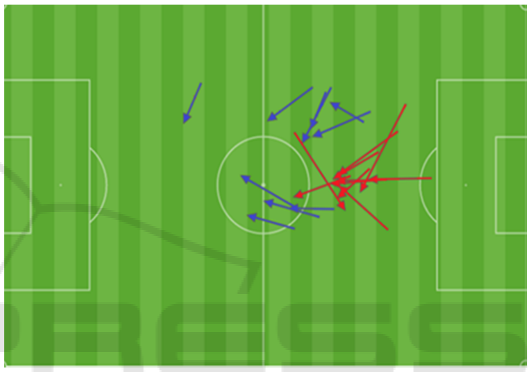

Figure 1: Passes of a player in a match.

Each pass in a season of Başakşehir F.K is analyzed according to $x_1$, $y_1$, $x_2$ *and* $y_2$ coordinates. Each player's passes are converted by using the formula 1, 2, 3 and 4 according to the 4 main directions that are left, right, back and forward.

$$x_2 - x_1 > 0 \Rightarrow f = x_1 - x_2 \tag{1}$$

$$x_2 - x_1 < 0 \Rightarrow b = x_1 - x_2 \tag{2}$$

$$y_2 - y_1 > 0 \Rightarrow r = x_1 - x_2 \tag{3}$$

$$y_2 - y_1 < 0 \Rightarrow l = x_1 - x_2 \tag{4}$$

$$P_\lambda = (|S_\lambda^1|) / (|S_\lambda^1 \cup S_\lambda^0|) \tag{5}$$

In this formula;
$P_\lambda$: Percentage of pass success for direction $\lambda$.

$\lambda=\{r, l, f, b\}$

$S_\lambda^1$: Set of successful passes.

$S_\lambda^0$: Set of unsuccessful passes.

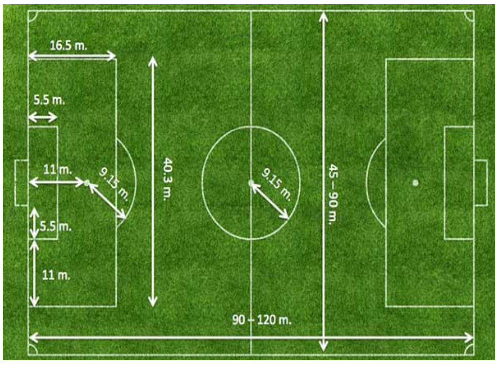


Figure 2: General pitch measurements.

Although there is not any standard measure of a pitch in the world, most of the stadium size in Turkey are 68 meters wide and 105 meters long. The previously acquired data are processed according to size of the pitch in Turkey such as step 4. According to this, the formula 6 is used for right and left direction passes and the formula 7 is used for forward and back direction passes.

$$a_{\,ij\,v} = \frac{\left|\frac{e}{2} - y_2\right|}{e} * |y_2 - y_1| \qquad (6)$$

$$a_{\,ij\,h} = \frac{x_2}{b} * |x_2 - x_1| \qquad (7)$$

**i:** Starting point of transshipment of ball for each pass.

**j:** Ending point of transshipment of ball for each pass.

$\boldsymbol{a_{ijv}}$**:** The success score of the vertical pass from **i** to **j**.

$\boldsymbol{a_{ijh}}$**:** The success score of the horizontal pass from **i** to **j**.

**e:** The length of the pitch where the match will be played.

**b:** The width of the pitch where the match will be played.

The 34 statement in the Formula 8 shows the center point of the pitch which is 68 meters wide. Therefore each ball transshipped to the center will have a high score. Similarly, the expression 105 in formula 9 shows the length of the field. Thus, each pass that makes the ball approach the opponent's goal will be more valuable.

$$a_{\,1\,2\,v} = \frac{|34 - y_2|}{34} * |y_2 - y_1| \qquad (8)$$

$$a_{\,1\,2\,h} = \frac{x_2}{105} * |x_2 - x_1| \qquad (9)$$

The red and blue arrows in figure 1 represent successful and unsuccessful passes. Therefore I categorized the previously passes analyzed as successful and unsuccessful. $a_{\,i\,j\,h}$, $a_{\,i\,j\,v}$ are categorized as $a_{\,i\,j\,l}$, $a_{\,i\,j\,r}$, $a_{\,i\,j\,b}$ and $a_{\,i\,j\,f}$.

$$\mu_\lambda = \frac{\sum_{i=1}^{k} d_{i\lambda}^{1} + \sum_{i=1}^{k} d_{i\lambda}^{0}}{n_\lambda} \quad (\forall\lambda) \qquad (10)$$

In this formula;

$\boldsymbol{\mu_\lambda}$**:** The average score of the pass done in the $\lambda$ direction according to $\boldsymbol{a_{\,i\,j\,l}}$, $\boldsymbol{a_{\,i\,j\,r}}$, $\boldsymbol{a_{\,i\,j\,b}}$ and $\boldsymbol{a_{\,i\,j\,f}}$

$\lambda=\{r, l, f, b\}$

$\boldsymbol{d_{i\lambda}^{1}}$**:** The score of $\boldsymbol{i}$ th successful pass in direction $\boldsymbol{\lambda}$.

$i \geq 0,\ \forall i = N,\ i \in (\, S_\lambda^1 \vee S_\lambda^0\,)$

$\boldsymbol{d_{i\lambda}^{0}}$**:** The score of $\boldsymbol{i}$ th unsuccessful pass in direction $\boldsymbol{\lambda}$.

$\boldsymbol{n_\lambda}$: Total number of passes made in direction $\boldsymbol{\lambda}$.

$n_\lambda \geq 0,\ n_\lambda = s(\, S_\lambda^1 \vee S_\lambda^0\,)$

Linear Approach:

The average score for the average distance is known. So we created a score data for any distance. Two different methods were used for this score. The linear approach formula is given below.

$$P_{\lambda k}^{l} = \frac{\mu_{\lambda k} * P_{\lambda k}}{d_{ik}} \qquad (11)$$

$$if\ P_{\lambda k}^{l} \leq 100, \qquad c_{\lambda k} = P_{\lambda k}^{l}/100 \qquad (12)$$

$$else\ P_{\lambda k}^{l} > 100, \quad c_{\lambda k} = 1 \qquad (13)$$

In this formula;

$\boldsymbol{P_{\lambda k}^{l}}$**:** The success percentage at the desired distance for the direction **λ** of the player $\boldsymbol{k}$.

$\boldsymbol{P_{\lambda k}}$**:** The success percentage at the average distance for a season for the direction **λ** of the player $\boldsymbol{k}$.

$\boldsymbol{\mu_{\lambda k}}$**:** The average pass distance for the direction **λ** of the player $\boldsymbol{k}$.

$\boldsymbol{d_{ik}}$**:** The desired distance for $\boldsymbol{i}$ th pass of the player **k**.

Exponential Approach:

A quadratic equation formula was used for the exponential approach. This formula is given below.

$$P^{\iota}_{\lambda k} = \frac{{\mu_{\lambda k}}^2 * P_{\lambda k}}{{d_{ix}}^2} \tag{14}$$

$$if\ P^{\iota}_{\lambda k} \leq 100, \qquad c_{\lambda k} = P^{\iota}_{\lambda k}/100 \tag{15}$$

$$else\ P^{\iota}_{\lambda k} > 100,\ \ c_{\lambda k} = 1 \tag{16}$$

$$c_{\lambda k} = \frac{1}{\sqrt{\frac{1}{{a_{\,i\,j\,h}}^2} + \frac{1}{{a_{\,i\,j\,v}}^2}}} \tag{17}$$

$$c_{ijk} = \frac{1}{\sqrt{\frac{1}{{a_{\,i\,j\,h}}^2} + \frac{1}{{a_{\,i\,j\,v}}^2}}} \tag{18}$$

$\boldsymbol{c_{\lambda k}}$**:** is the pass score at the desired distance for the direction λ of the player **k**.
$\boldsymbol{c_{ijk}}$**:** is the pass score at the desired distance done by player **k** from **i** to **j**.
The football ball will horizontally and vertically change its position against the opposing goal with the pass to be thrown. The distance of the pass is converted to score according to formulas 7 and 8.
The $\boldsymbol{a}_{\,i\,j\,h}$ *and* $\boldsymbol{a}_{\,i\,j\,v}$ scores in formulas 13 and 14 are as follows:
$\boldsymbol{a}_{\,i\,j\,h}$**:** Back direction passes or forward direction pass is score converted by formula 7, 8.
$\boldsymbol{a}_{\,i\,j\,v}$**:** Left direction passes or right direction pass is score converted by formula 7, 8.
The following model is the interpretation of a transportation model according to the football game.

**Objective Function:**
**Max**

$$\sum_{i=1}^{m}\sum_{j=1}^{m}\sum_{k=1}^{2} c_{ijk} * x_{ij} \tag{19}$$

Constraints:

$$x_{ijk} \neq x_{jzk} \qquad (\forall i, j, z, k) \tag{20}$$

$$x_{ij1} + x_{ij2} = 1 \quad (\forall i, j) \tag{21}$$

$$\textstyle\sum_{j=1,j\neq i}^{m} x_{ij} - \sum_{j=1,j\neq i}^{m} x_{ji} = 0$$
$$(\forall i = 1,2,..\,m \text{ ve } i \neq \gamma, F) \tag{22}$$

$$\textstyle\sum_{j=1,j\neq\gamma}^{m} x_{\gamma j} - \sum_{j=1,j\neq i}^{m} x_{j\gamma} = -1 \quad (\forall\gamma) \tag{23}$$

$$\left(\sum_{i=\theta}^{t}\sum_{j=1,j\neq i}^{m} x_{1j}\right) - \left(\sum_{i=\theta}^{t}\sum_{j=1,j\neq i}^{m} x_{ji}\right) = 1 \tag{24}$$

$$if \begin{cases} x_{ij} = 1, & c_{ijk} = 0 \\ x_{ij} = 0, & c_{ijk} = 1 \end{cases} \tag{25}$$

$$t, \theta > 0 \tag{26}$$

$$x_{ji} \in \{0, 1\} \quad i, j = 1,2,3,..m \tag{27}$$

19: The objective function is to maximize the sum of the scores according to the distance of **k** or **v** player's pass. The number 60 indicates that the football pitch was divided into 60 symmetric zones.
**m:** it's indicates the number of parts that form the pitch. The football pitch is divided into approximately 60 equal parts. Which corresponds to an area of approximately 1 square meter of each piece. **m** is 60 for this model
$\mathbf{x_{ij}}$: Indicates whether the ball has been transferred from zone i to zone j.

$\mathbf{c_{ijk}}$: Indicates the pass score of **k** player from **i** to **j**.
S: It is a set of players who play in the first 11 of a team. (k, v ϵ S)
20: A player cannot throw pass to himself.
21: One of the players **k** can throw pass from one zone.
22: The ball must exit from that zone if it has once entered the zone.
$\boldsymbol{\gamma}$**:** In a game played, the ball is a zone where the ball is taken from the opponent and where the attack organization begins. ( $\gamma = 1, 2 \ldots . m$)
**F:** is the region where the arc starting in the $\boldsymbol{\theta}$ region and ending in the t region is terminated. ($\theta$, $t$ ϵ F), {45, 46, 50, 51, 52, 53, 56, 57, 58, 59} ϵ F
23: The ball can enter the pitch only from certain zones.
24: The attack may only end in certain zones.
25: Sub tour constraint.
**θ:** The first of the effective regions for ending the attack. ($\theta$ ϵ F)
**t:** The last effective region for ending the attack ($t$ ϵ F )

## 2.3 Application of the Model

The match of Besiktas J.K with Basaksehir F.K in 2014/15 season was investigated for this study. The squad in of Başakşehir F.K. in the match was taken exactly and the players assignments as in Figure 3.

Table 1: The scores of the players.

| No | Score | Left | Right | Forward | Back |
|---|---|---|---|---|---|
| *#1* | Points | 8,810 | 7,827 | 16,88 | 0,042 |
| *#80* | Points | 4,797 | 2,096 | 6,447 | 3,539 |
| *#14* | Points | 7,006 | 5,387 | 6,661 | 4,628 |
| *#15* | Points | 5,832 | 6,285 | 4,479 | 3,674 |
| *#11* | Points | 2,472 | 4,043 | 7,312 | 3,960 |
| *#7* | Points | 5,156 | 5,399 | 5,578 | 3,261 |
| *#8* | Points | 4,838 | 5,774 | 5,298 | 3,459 |
| *#5* | Points | 6,662 | 7,359 | 6,156 | 3,075 |
| *#23* | Points | 5,119 | 4,888 | 4,443 | 3,422 |
| *#17* | Points | 5,078 | 5,076 | 5,234 | 3,579 |
| *#26* | Points | 3,202 | 4,491 | 3,962 | 3,172 |

| | | | | | |
|---|---|---|---|---|---|
| 6.Region | 5.Region *(1) | 4.Region *(1) | 3.Region *(1) | 2.Region *(1) | 1.Region *(11) |
| 12.Region *(80) | 11.Region *(1) *(80) | 10.Region *(1) | 9.Region *(1) | 8.Region *(15) | 7.Region *(11) |
| 18.Region *(80) *(15) | 17.Region *(15) *(11) *14 *(14) *(80) | 16.Region *(11) *(14) *(15) | 15.Region *(11) *(14) *(15) | 14.Region *(11) *(14) *(15) | 13.Region *(11) |
| 24.Region *(7) *(8) *(15) *(80) | 23.Region *(14) *(11)*(5) *(8) *(15)*(80) | 22.Region *(14) *(8) *(5) *(23) *(15)*(11) | 21.Region *(8) *(23) *(14) *(15) *(5) | 20.Region *(8) *(23) *(15) *(5) *(14) *(11) | 19.Region *(8) *(17) *(11) *(23) *(5) |
| 30.Region *(8) *(7) *(17) *(15) *(80) | 29.Region *(8) *(14) *(15) *(80) *(5) | 28.Region *(8) *(14) * (15) *(23) *(5) | 27.Region *(8) *(14) *(23) *(5) *(15) | 26.Region *(11) *(23) *(8) *(14) *(15) *(5) | 25.Region *(11) *(23) *(5) *(17) *(8) |
| 36.Region *(7) *(8) *(17) *(80) *(15) | 35.Region *(7) *(15) *(8) *(17) *(23) *(80) *(26) *(5) *(14) | 34.Region *(15) *(8) *(17) *(23) *(26) *(5) *(14) | 33.Region *15 *(8) *(17) *(23) *(26) *(5) | 32.Region *(15) *(8) *(17) *(23) *(26) *(5) *(11) | 31.Region *(15) *(8 )*(17) *(23) *(26) *(5) *(11) |
| 42.Region *(7) *(8) *(17) *(80) *(15) | 41.Region *(7) *(15) *(8) *(17) *(23) *(80) *(26)*(5) | 40.Region *(15) *(8) *(17) *(23) *(26) *(5) | 39.Region *(15) *(8) *(23) *(26) *(5) | 38.Region *(15) *(8) *(17) *(23)*(26) *(5) | 37.Region *(8) *(17) *(26) *(5) *(11) |
| 48.Region *(7) *(8) *(17) *(80) *(15) | 47.Region *(7) *(8) *(17) *(80) *(26) | 46.Region *(26) | 45.Region *(5) *(26) | 44.Region *(5) *(26) *(8) *(17) | 43.Region *(8 ) *(17) *(26) *(11) |
| 54.Region *(7) *(17) *(80) | 53.Region *(7) *(80) *(26) | 52.Region | 51.Region | 50.Region *(26) | 49.Region *(17) |
| 60.Region *(7) | 59.Region | 58.Region | 57.Region | 56.Region | 55.Region |

Figure 3: Regional assignment.

Table 1 shows the scores of the players of Başakşehir for each direction. Formula 8 and 9 were used for this scoring.

Table 2: The average pass success rates.

| No | Ratio | Left | Right | Forward | Back |
|---|---|---|---|---|---|
| *#1* | % | 56,9 | 61,3 | 56,9 | 100 |
| *#80* | % | 83,19 | 76,92 | 73,47 | 96,21 |
| *#14* | % | 83,3 | 83,5 | 79,9 | 95,4 |
| *#15* | % | 84,5 | 86,8 | 79 | 96,1 |
| *#11* | % | 53,5 | 87,4 | 64,4 | 95,7 |
| *#7* | % | 73,8 | 84,3 | 64,9 | 90,7 |
| *#8* | % | 82,9 | 81 | 75,5 | 88,5 |
| *#5* | % | 85,3 | 85,54 | 79,77 | 97,24 |
| *#23* | % | 80,3 | 81,5 | 72,02 | 96 |
| *#17* | % | 74,7 | 76,3 | 62,4 | 90,3 |
| *#26* | % | 77,34 | 76,31 | 63,57 | 88,43 |

The following figures show the Optasport routes and routes of the model.

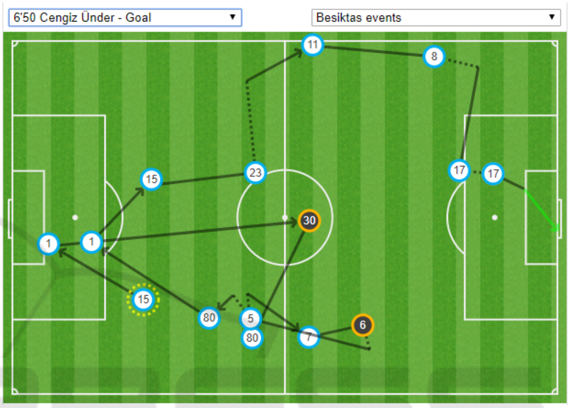


Figure 4: Attack at 6th minutes.

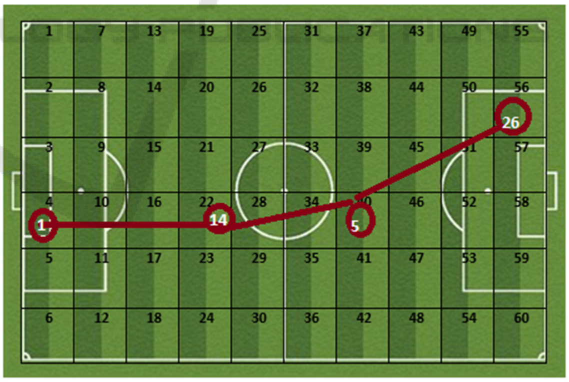

Figure 5: Exponential approach at 6th minutes.

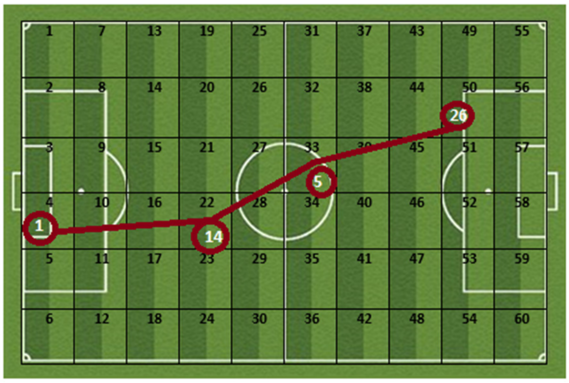


Figure 6: Linear approach at 6th minutes.

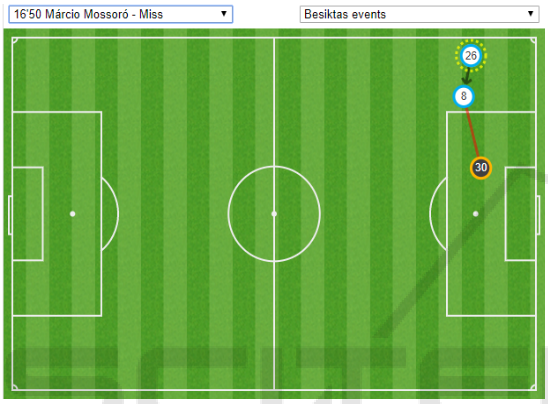


Figure 7: Attack at 16th minutes.

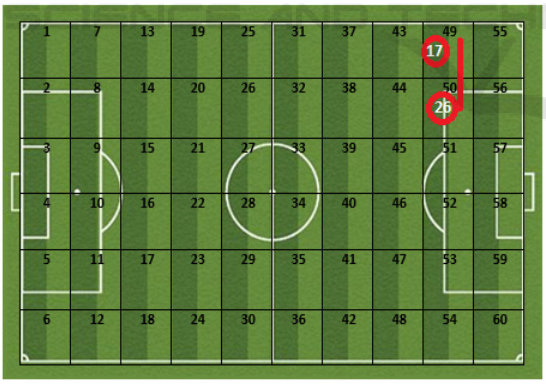


Figure 8: Exponential approach at 16th minutes.

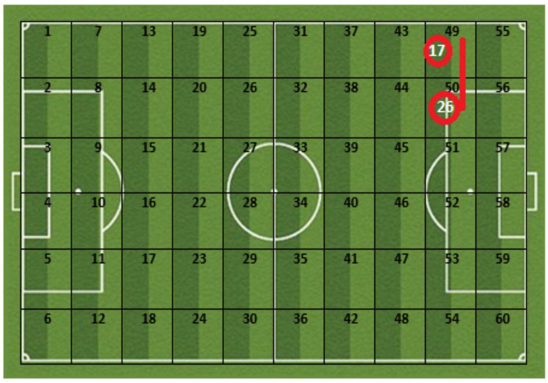


Figure 9: Linear approach at 16th minutes.

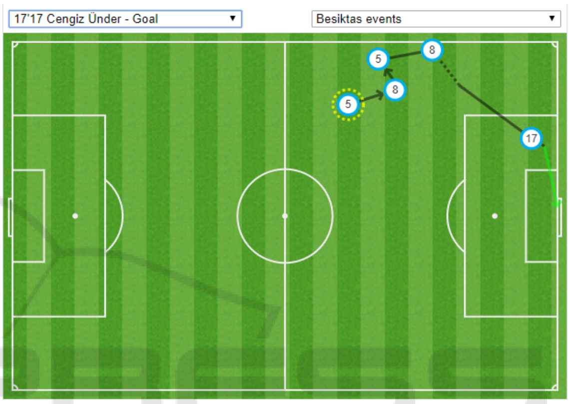


Figure 10: Attack at 17th minutes.

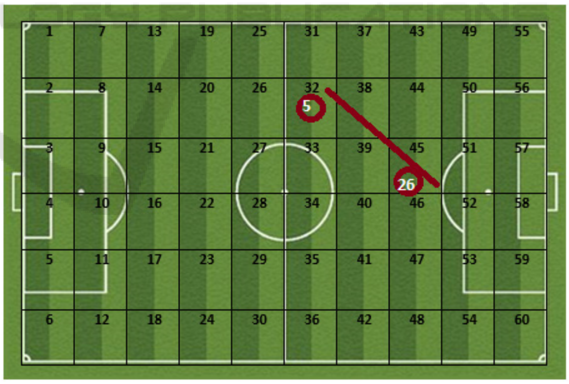


Figure 11: Exponential approach at 17th minutes.

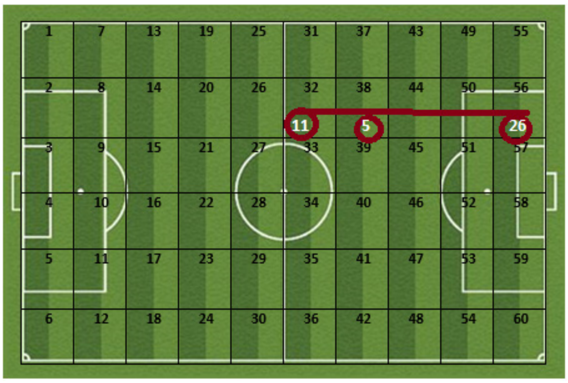

Figure 12: Linear approach at 17th minutes.

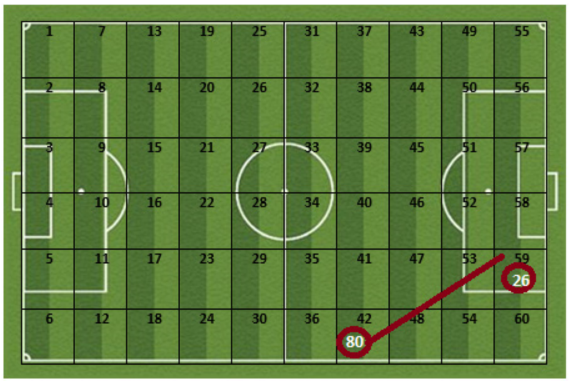

Figure 15: Linear approach at 26th minutes.

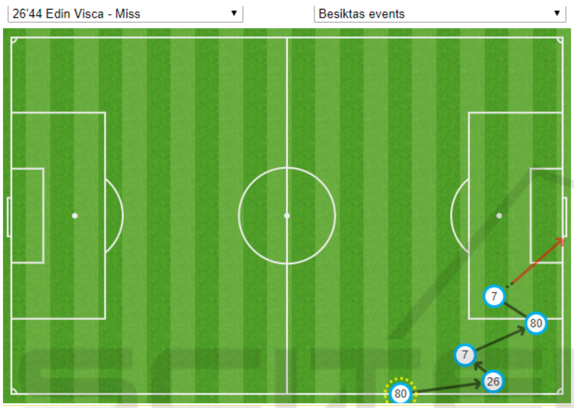


Figure 13: Attack at 26th minutes.

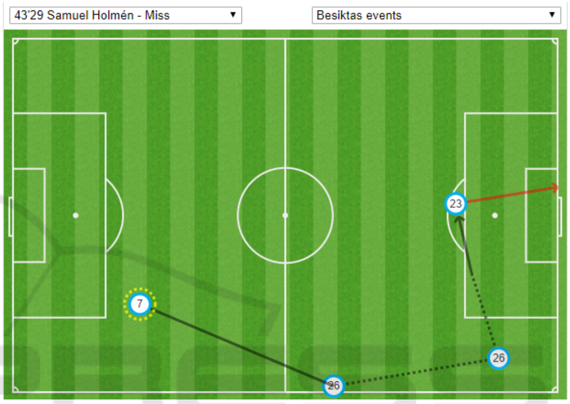


Figure 16: Attack at 43th minutes.

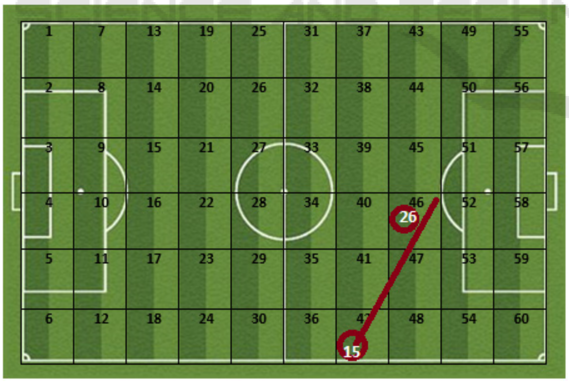

Figure 14: Exponential approach at 26th minutes.

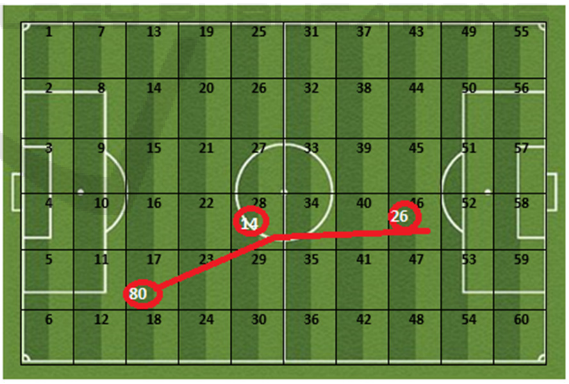

Figure 17: Exponential approach at 43th minutes.

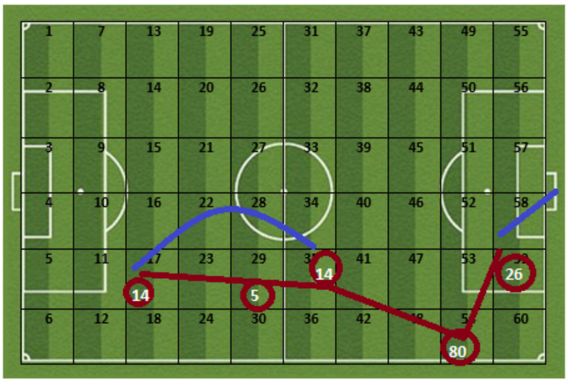


Figure 18: Linear approach at 43th minutes.

## 3 CONCLUSIONS

The reason for selecting these special minutes in the match is the success of the attack occurs in these minutes. In particular, the route in the 26th minute of match and route of the model using the linear data is quite similar. Apart from this, the linear and exponential approach for the 16th position of the match showed the same result. The first and end points of the position at 17th minute of the match are quite similar to the results of the linear approach. In general, the linear approach shows quite similar preferences with actual routing, whereas the exponential approach has chosen to reach the effective zones with less passes such as counter attack.

## REFERENCES

Barghi, A. R. (2015). Analyzing dynamic football passing network. PhD Thesis. Université d'Ottawa/University of Ottawa.

Cintia, P., Giannotti, F., Pappalardo, L., Pedreschi, D., and Malvaldi, M. (2015). The harsh rule of the goals: Data-driven performance indicators for football teams. Paper presented at *the 2015 IEEE International Conference on Data Science and Advanced Analytics (DSAA).*

Grund, T. U. (2012). Network structure and team performance: The case of English Premier League soccer teams. 34(4), 682-690. doi:10.1016/j.socnet.2012.08.004

Gudmundsson, J., and Horton, M. J. A. C. S. (2017). Spatio-temporal analysis of team sports. 50(2), 22.

Hirano, S.,and Tsumoto, S. (2005). A clustering method for spatio-temporal data and its application to soccer game records. Paper presented at the *International Workshop on Rough Sets, Fuzzy Sets, Data Mining, and Granular-Soft Computing.*

Krishnan, K., and Rao, V. J. J. o. I. E. (1965). Inventory control in N warehouses. In (Vol. 16, pp. 212-&).

Liu, H., Hopkins, W., Gómez, A. M., and Molinuevo, S. J. J. I. J. o. P. A. i. S. (2013). Inter-operator reliability of live football match statistics from OPTA Sportsdata. 13(3), 803-821.

Malqui, J. L. S. (2017). A visual analytics approach for passing strateggies analysis in soccer using geometric features.

Pena, J. L., and Touchette, H. J. a. p. a. (2012). A network theory analysis of football strategies.

Razykov, S. (2006). Optimal offensive player positioning and collaboration in a digital soccer game. *School of Interactive Arts and Technology-Simon Fraser University,*

Stein, M., Janetzko, H., Seebacher, D., Jäger, A., Nagel, M., Hölsch, J., and Grossniklaus, M. J. D. (2017). How to make sense of team sport data: From acquisition to data modeling and research aspects. 2(1), 2.

Takeuchi, J., Ramadan, R., and Iida, H. J. I. P. S. o. J. (2014). Game refinement theory and its application to Volleyball. 2014, 1-6.

Tavana, M., Azizi, F., Azizi, F., and Behzadian, M. J. S. M. R. (2013). A fuzzy inference system with application to player selection and team formation in multi-player sports. 16(1), 97-110.

Wright, M. B. (2009). 50 years of OR in sport. *Journal of the Operational Research Society*, 60(sup1), S161-S168.