Simulations of Nearest Teammate-Based Soccer Match-Plays with Different Formations

# Simulations of Nearest Teammate-Based Soccer Match-Plays with Different Formations

Francesco Scotognella

Dipartimento di Fisica, Politecnico di Milano, piazza Leonardo da Vinci 32, 20133, Milano, Italy

E-mail address: francesco.scotognella@polimi.it

## Abstract

Formation in soccer is among the most important tactical choices for a successful match.
Herein, the simulations of 420000 match-plays have been performed varying the formation, the number of opponents that are actively pressing the team, the speed of the opponents in attempting a pass interception. Dribbling has been neglected. The match-play ends either with a successful series of passes from a central back to the line of the strikers or with the opponents that steal the ball. In this work, I demonstrate that 3-4-3 formation, which is among the most employed formations, relates to the highest probability of success.

## Keywords

Network; Graph theory; Soccer; Performance analysis

## Introduction

Soccer, with around three billion fans all over the world, is nowadays the most popular sport. In soccer the formation describes the position of the players on the pitch and it is one of the most important tactical choices of the team's manager or coach.
Many reports on various tools to analyse soccer matches have been published [1–4]. In 2010, Tenga, Roglan, and Bahr examine the effectiveness of offensive match-plays in terms of scoring opportunities, score box possessions, and goals scored [5]. Rein and Memmert introduce the employment of big data technologies to analyse soccer matches and to support the development of theoretical models of tactical decision making [6]. Recently, graph theory has been successfully used to describe and analyse soccer matches. Network metrics have been used to analyse offensive match-plays in matches, for example to study node (i.e., player) connectivity [7]. Image-based representations to encode temporal network changes employed to analyse soccer matches have been developed by Uchoa Maia Rodrigues [8]. To mention other examples, in 2016 Gama et al. publish the work "Networks metrics and ball possession in professional football" [9] and in 2019 Korte et al. publish the work "Play-by-Play Network Analysis in Football" [10].
In this work, I simulate soccer match-plays with different formations, number of pressing opponents and number of attempts to pass the ball to the nearest neighbour teammate. I employ graphs in which the players are the hubs and the nearest neighbour pass trajectories are the links. The model is very light and allows a very large number of simulations. I find that the most efficient formation, in terms of highest probability to successfully pass the ball starting from a central back to the line of the strikers, is 3-4-3.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 27 August 2021 doi:10.20944/preprints202108.0519.v1

## Methods

In a certain formation the links between the nearest neighbour players are considered. In Figure 1, the players of the same team are depicted as circles and the links between the nearest neighbour teammates as dashed lines. The "X" sign corresponds to the position in which a pressing opponent can intercept the pass.

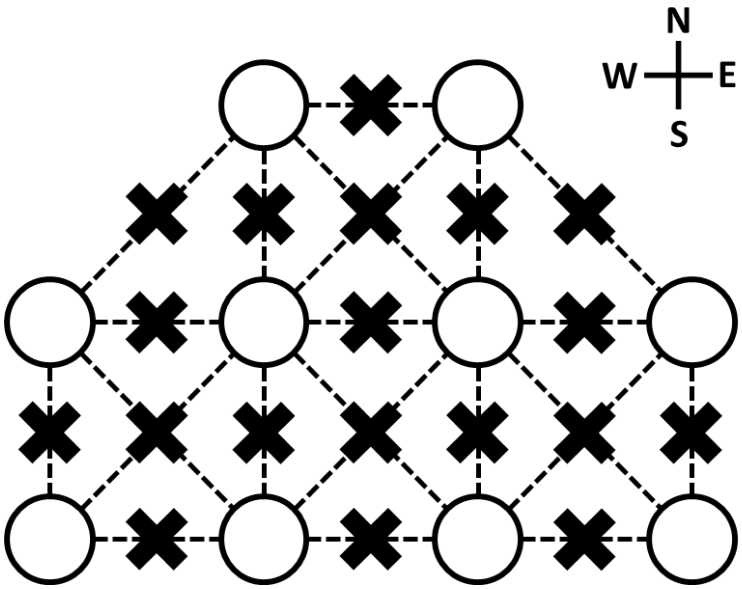


**Figure 1**. 4-4-2 formation. The players are depicted as circles. The dashed lines relate to the links between the nearest neighbour players. "X" sign corresponds to the position in which a pressing opponent can intercept the pass.

In Figure 2, the flow chart of the algorithm is depicted. The considered pass is between a player and the nearest neighbour. The direction of the player's pass is related to a random number between 1 and 8. With the goalkeeper in the South (S) direction and the strikers in the North (N) direction, 1 means N direction, 2 Northeast (NE) direction, 3 East (E) direction, 4 Southeast (SE) direction, 5 S direction, 6 Southwest (SW) direction, 7 West (W) direction, 8 Northwest (NW) direction. If the opponent is in the direction of the pass, the player will try a new pass, but the counter of attempts (*Delay*) is increased by 1. If *Delay* reaches the parameter *MaxDelay* the opponent intercepts the pass and the match-play ends (unsuccessful match-play). If the players succeed in passing the ball up to the line of the strikers the match-play ends (successful match-play).

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 27 August 2021 doi:10.20944/preprints202108.0519.v1

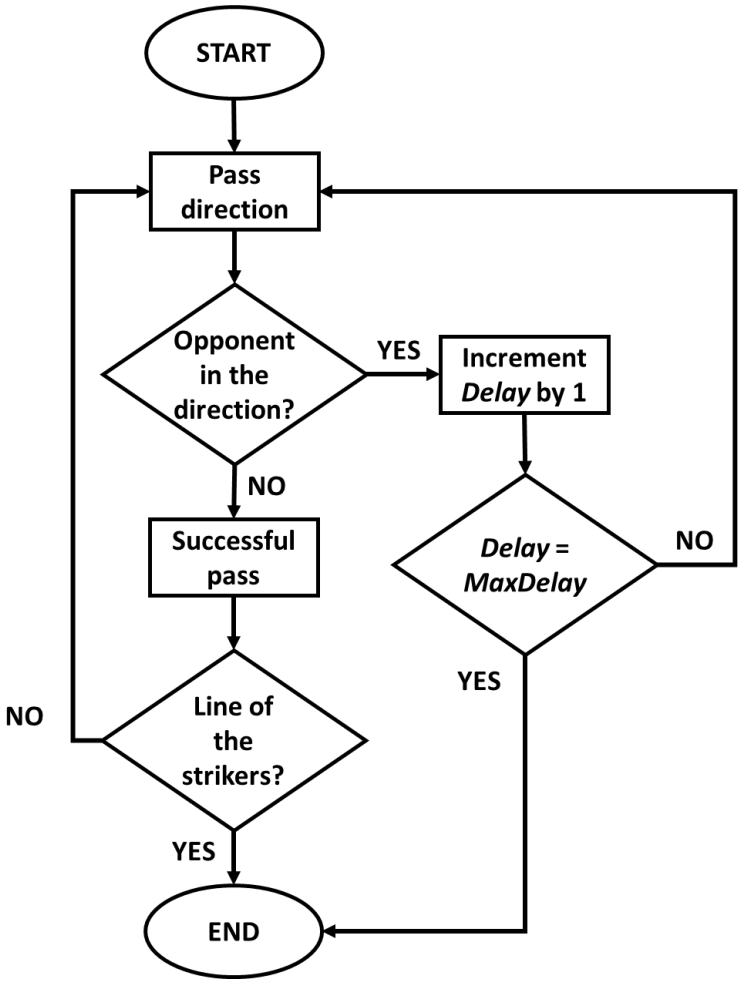


**Figure 2**. Flow chart of the algorithm. The player chooses a direction for the pass; with no opponent in such direction the pass is successful. If the ball reaches the strikers line, the algorithm ends. If the opponent is in the direction of player choice, the player can choose again, but a counter of attempts (*Delay*) increases. If *Delay* is equal to *MaxDelay* the algorithm ends (intercepted pass).

The parameters in the match-play simulations are the number of pressing opponents (6, 8, 10) and *MaxDelay* (10, 20). The pressing opponents are randomly distributed on the "X" and each "X" can only be occupied by one opponent. The algorithm is iterated 10000 times for a set of formation, number of opponents, and *MaxDelay*. With 6 different formations, the total number of simulations is 360000.

In Figure A1 in the Supporting information we show the convergence of the calculation. The probability of success to reach the line of the strikers, the average number of passes and its standard deviation starts to converge with more than 1000 simulations.

## Results and Discussion

In Table 1 the probability of success to reach the line of the strikers with a certain set of formation, number of pressing opponents, and *MaxDelay* is reported.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 27 August 2021 doi:10.20944/preprints202108.0519.v1

**Table 1**. Probability to reach the line of strikers with different formation, number of pressing opponents (# opponents) and maximum delay (MaxDelay).

*Formation: 4-4-2*

| # opponents / MaxDelay | 6 | 8 | 10 |
|---|---|---|---|
| 10 | 38.27 % | 20.41 % | 9.62 % |
| 20 | 71.82 % | 45.84 % | 25.38 % |

*Formation: 3-3-3-1*

| # opponents / MaxDelay | 6 | 8 | 10 |
|---|---|---|---|
| 10 | 15.95 % | 5.95 % | 1.90 % |
| 20 | 48.58 % | 21.47 % | 8.92 % |

*Formation: 4-2-4*

| # opponents / MaxDelay | 6 | 8 | 10 |
|---|---|---|---|
| 10 | 30.35 % | 15.37 % | 7.27 % |
| 20 | 55.99 % | 34.84 % | 17.16 % |

*Formation: 4-5-1*

| # opponents / MaxDelay | 6 | 8 | 10 |
|---|---|---|---|
| 10 | 18.50 % | 7.95 % | 3.72 % |
| 20 | 46.22 % | 22.93 % | 10.81 % |

*Formation: 4-3-3*

| # opponents / MaxDelay | 6 | 8 | 10 |
|---|---|---|---|
| 10 | 40.37 % | 22.25 % | 10.81 % |
| 20 | 71.32 % | 45.99 % | 25.55 % |

*Formation: 3-4-3*

| # opponents / MaxDelay | 6 | 8 | 10 |
|---|---|---|---|
| 10 | 46.47 % | 25.86 % | 13.42 % |
| 20 | 78.44 % | 53.74 % | 32.57 % |

*Formation: 3-3-4*

| # opponents / MaxDelay | 6 | 8 | 10 |
|---|---|---|---|
| 10 | 42.71 % | 24.08 % | 12.30 % |
| 20 | 74.88 % | 49.98 % | 28.67 % |

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 27 August 2021 doi:10.20944/preprints202108.0519.v1

The formation with the highest probability of success is the 3-4-3 formation. It is important to highlight that such formation has 3 strikers, and this increases the probability to reach the line of the strikers. Moreover, the network related to the 3-4-3 formation, in which the hubs are the players, has a relatively high number of links among the hubs (19). The 4-4-2 formation is related to a network with 19 links, as the 3-4-3 one, but with two strikers instead of three strikers. The 4-3-3 has three strikers, but 18 links instead of 19 links.
It is worth noting the 3-3-3-1 formation is characterized by a very low probability to reach the striker. With respect to a three-line formation this is a four-line formation, with an additional line of offensive midfielders, which are more upfield with respect to the line of defensive midfielders.

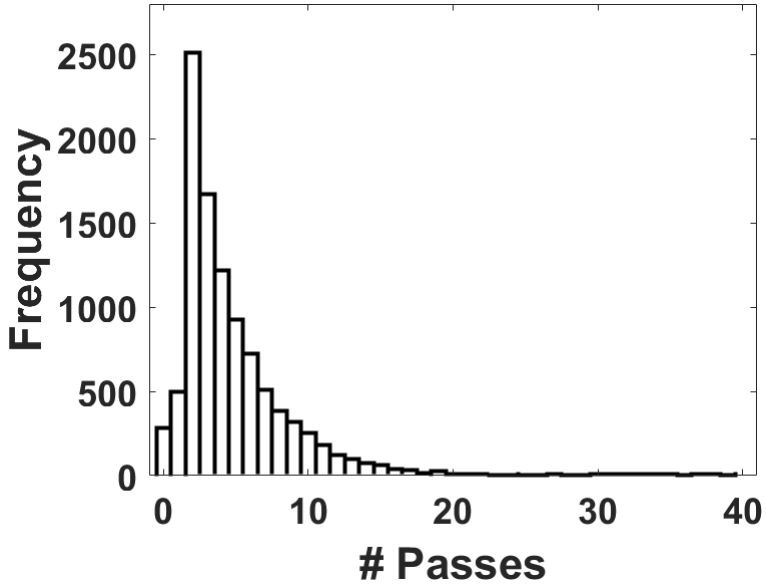


**Figure 3**. Distribution of the number of passes in a match-play for the case of the 3-4-3 formation, 6 pressing opponents, *MaxDelay* 20.

In Figure 3, the histogram with the distribution of number of passes in a match-play is given for the case of the 3-4-3 formation, 6 pressing opponents and maximum delay (*MaxDelay*) of 20. The average number of passes in this simulation with 10000 match-plays is 4.63. The distribution has a long tail with a maximum number of passes equal to 39.

## Conclusion

In this work, we have simulated 420000 different match-plays, varying formation, number of opponents that do pressing during the match-play, and maximum delay induced by the opponents. The match-play starts with the ball possessed by one of the central backs and ends either with the ball at the line of the strikers or with the ball intercepted by the opponents. Dribbling of opponents has not been considered. I demonstrated that the highest probability to reach the line of the strikers with the ball is achieved employing a 3-4-3 formation. A tactical choice with only passes to nearest neighbour teammates can be useful when fatigue is increasing and to support participation of all

players, since it has been observed that such participation decreases in the second half of the match [7]. The presented model can be improved by carefully taking into account the formation of the opponent team.

**Acknowledgement**

The author acknowledges Ilka Kriegel for the useful comments.

**References**

[1] M. Marchiori, M. de Vecchi, Secrets of soccer: Neural network flows and game performance, Computers & Electrical Engineering. 81 (2020) 106505. https://doi.org/10.1016/j.compeleceng.2019.106505.

[2] P. Halvorsen, S. Sægrov, A. Mortensen, D.K.C. Kristensen, A. Eichhorn, M. Stenhaug, S. Dahl, H.K. Stensland, V.R. Gaddam, C. Griwodz, D. Johansen, Bagadus: an integrated system for arena sports analytics: a soccer case study, in: Proceedings of the 4th ACM Multimedia Systems Conference, Association for Computing Machinery, New York, NY, USA, 2013: pp. 48–59. https://doi.org/10.1145/2483977.2483982 (accessed August 23, 2021).

[3] C. Perin, R. Vuillemot, J.-D. Fekete, SoccerStories: A Kick-off for Visual Soccer Analysis, IEEE Transactions on Visualization and Computer Graphics. 19 (2013) 2506–2515. https://doi.org/10.1109/TVCG.2013.192.

[4] A. Rehman, T. Saba, Features extraction for soccer video semantic analysis: current achievements and remaining issues, Artif Intell Rev. 41 (2014) 451–461. https://doi.org/10.1007/s10462-012-9319-1.

[5] A. Tenga, L.T. Ronglan, R. Bahr, Measuring the effectiveness of offensive match-play in professional soccer, European Journal of Sport Science. 10 (2010) 269–277. https://doi.org/10.1080/17461390903515170.

[6] R. Rein, D. Memmert, Big data and tactical analysis in elite soccer: future challenges and opportunities for sports science, SpringerPlus. 5 (2016) 1410. https://doi.org/10.1186/s40064-016-3108-2.

[7] F.M. Clemente, M.S. Couceiro, F.M.L. Martins, R.S. Mendes, Using Network Metrics in Soccer: A Macro-Analysis, Journal of Human Kinetics. 45 (2015) 123–134. https://doi.org/10.1515/hukin-2015-0013.

[8] D.C. Uchoa Maia Rodrigues, F.A. Moura, S.A. Cunha, R. da S. Torres, Graph visual rhythms in temporal network analyses, Graph. Models. 103 (2019) 101021. https://doi.org/10.1016/j.gmod.2019.101021.

[9] J. Gama, G. Dias, M. Couceiro, T. Sousa, V. Vaz, Networks metrics and ball possession in professional football, Complexity. 21 (2016) 342–354. https://doi.org/10.1002/cplx.21813.

[10] F. Korte, D. Link, J. Groll, M. Lames, Play-by-Play Network Analysis in Football, Front. Psychol. 0 (2019). https://doi.org/10.3389/fpsyg.2019.01738.

*Preprints* (www.preprints.org) | NOT PEER-REVIEWED | Posted: 27 August 2021 doi:10.20944/preprints202108.0519.v1

**Appendix**

*Convergence*: In this work, for each set of formation, number of pressing opponents and value of *MaxDelay*, 10000 match-play simulations have been performed. In order to assess the reliability of this number of simulations, the probability of success to reach the line of the strikers, the average number of passes and the related standard deviation have been analysed as a function of match-play simulations for each formation discussed in the paper. In Figure A1 the aforementioned 3 values as a function of the match-play simulation (in log10 scale) is reported for the case of 3-3-3-1 formation, 6 pressing opponents and value of *MaxDelay* equal to 20. Above 1000 simulations the 3 values show small variations. Thus, an analysis with 10000 match-play simulations is reasonably reliable.

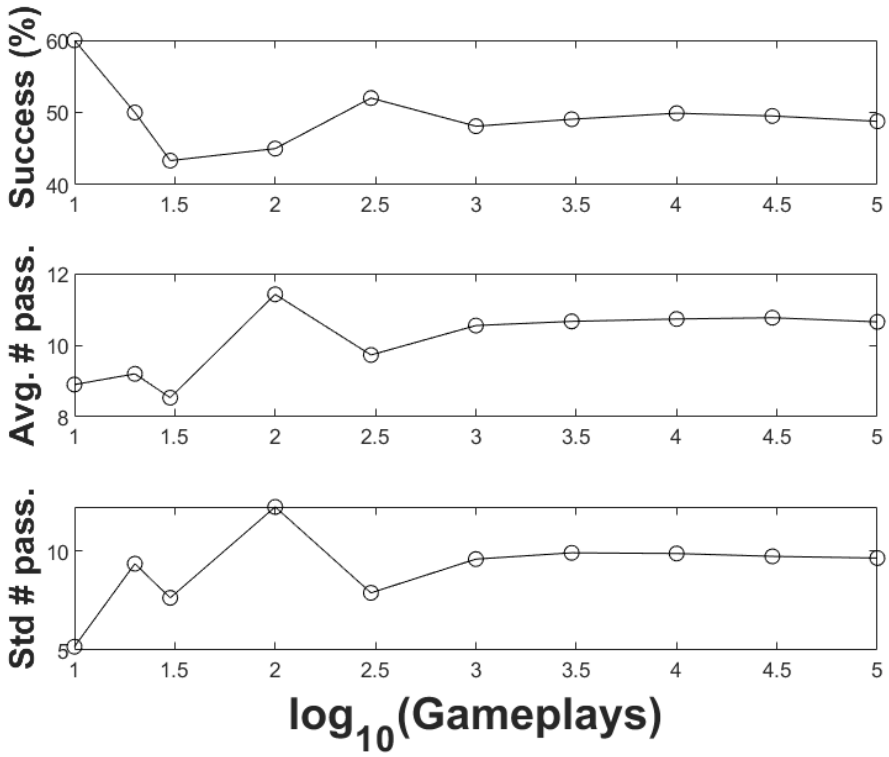


**Figure A1**. Probability of success to reach the line of the strikers, average number of passes and related standard deviation as a function of the simulated match-plays for the case of 3-3-3-1 formation, 6 pressing opponents and *MaxDelay* 20.