




Article

Node and Network Entropy—A Novel Mathematical Model for Pattern Analysis of Team Sports Behavior

Fernando Martins ^{1,2}, Ricardo Gomes ^{1,3,*}, Vasco Lopes ⁴, Frutuoso Silva ^{2,4}
and Rui Mendes ^{1,3,5}

¹ Instituto Politécnico de Coimbra, ESEC, UNICID-ASSERT, 3030-329 Coimbra, Portugal; fmlmartins@esec.pt (F.M.); rmendes@esec.pt (R.M.)

² Instituto de Telecomunicações, Delegação da Covilhã, 6201-001 Covilhã, Portugal; fsilva@di.ubi.pt

³ Instituto Politécnico de Coimbra, IIA, ROBOCORP, 3030-329 Coimbra, Portugal

⁴ Department of Informatics, Universidade da Beira Interior, 6201-001 Covilhã, Portugal; vascoferrinholopes@gmail.com

⁵ CIDAF, FCDEF, Universidade de Coimbra, 3040-248 Coimbra, Portugal

* Correspondence: rimgomes@esec.pt

Received: 28 July 2020; Accepted: 6 September 2020; Published: 9 September 2020



Abstract: Pattern analysis is a well-established topic in team sports performance analysis, and is usually centered on the analysis of passing sequences. Taking a Bayesian approach to the study of these interactions, this work presents novel entropy mathematical models for Markov chain-based pattern analysis in team sports networks, with Relative Transition Entropy and Network Transition Entropy applied to both passing and reception patterns. To demonstrate their applicability, these mathematical models were used in a case study in football—the 2016/2017 Champions League Final, where both teams were analyzed. The results show that the winning team, Real Madrid, presented greater values for both individual and team transition entropies, which indicate that greater levels of unpredictability may bring teams closer to victory. In conclusion, these metrics may provide information to game analysts, allowing them to provide coaches with accurate and timely information about the key players of the game.

Keywords: social network analysis; entropy; Markov chain; football

1. Introduction

Social Network Analysis (SNA) has increasingly gained attention in the research community over the past few years, emerging as a method of analysis of intra-team passing networks in team sports [1–5]. In football, passing networks between players (nodes) reflect how the entities are connected and bring light to how each entity within that group interacts with the others [1,2,4,6]. Group and individual metrics are used to describe the team's interactive behaviors using different degree measures to depict how connected the team is, as well as to identify the influential players within the team. By applying degree centrality and betweenness and closeness metrics to the global interaction of players, it is possible to assess the intermediary role of each player in distributing the ball during the game, allowing the identification of players with more relevant roles within the team [1,3,7–10].

uPATO is a dedicated software tool for network analysis in team sports [11,12]. In the last few years, the use of network analysis tools has been applied to team sports to understand how collective and individual performance may be optimized. These tools consider the network of passes between players during a game, which defines an adjacency matrix. Then, this matrix can be analyzed based on graph theory, which allows us to know which nodes (i.e., players) are more important and more involved in the game.

Another approach that can be performed to study ball passing networks created during the game is by using a Markov chain. According to this approach, the transition probability $P(j|i)$, represents a directed edge between player j to player i , and all the edges between possible player are given in advance. In this, every pass is independent of the previous sequence of passes [13,14]. Yamamoto [13], for example, has analyzed passing sequences in football using a Markov chain approach to estimate the predictability of certain passing courses and to determine when there is a chance of error in the information system—i.e., when a sequence of passes fails and the team loses possession.

Determining the degree of entropy of the probability matrices associated to Markov chains has not yet been explored in team sports analysis. According to entropy theory, higher levels of entropy show a greater unpredictability in how the system behaves. In our case, this means that the player (or the team) has a passing pattern that is more difficult to predict, and therefore more difficult to defend.

The aim of this paper is to present several novel mathematical models for pattern analysis in team sports, particularly football, to analyze the level of entropy in passing networks using a Markov chain approach. The novelty of this study relies in an innovative approach that proposes to estimate the likelihood of one player passing to, or receiving from, any other in the weighted directed network, as well as an estimation of the level of unpredictability of that player. In practical terms, higher levels of unpredictability will be reflected in a greater difficulty in predicting where to intercept the pass. Conversely, players with more predictable destinations will be more easily defendable and countered.

2. Novel Mathematical Models for Entropy of Nodes and Weighted Directed Networks

In this section, we present mathematical concepts that are based on information theory and probability theory that can be applied to networks when such networks can be considered weighted digraphs or weighted directed networks. The relation between the nodes of a network presented in this paper is described, in graph theory, by weighted digraphs [12,15,16]. Therefore, we consider that A_D^w is the weighted adjacency matrix of a weighted digraph with n nodes, G_D^w .

Based on the mathematical concept of conditional probability involving two random variables X and Y , such that the pair of transmitter and receiver (X, Y) is the joint distribution of the Markov chain associated with the weighted adjacency matrix, A_D^w , of a weighted digraph with n nodes, G_D^w , [17](p. 13)–[18–20] (p. 156), we can define the concept of the Markov chain transition matrix.

Definition 1. The $n \times n$ real matrix of order n , $M_T = [m_{ij}] \in R^{n \times n}$, is called the transition matrix of the Markov chain associated with the A_D^w of a G_D^w with n nodes. Each element of M_T is obtained by:

$$m_{ij} = p(Y = y_j | X = x_i) = \frac{w_{ij}}{\sum_{j=1}^n w_{ij}}, \tag{1}$$

where the two random variables X and Y are such that the pair of transmitter and receiver (X, Y) is the joint distribution of the Markov chain associated with the A_D^w of G_D^w , $i, j = 1, \dots, n$.

Considering the mathematical concept of n -step transition probabilities [20] (p. 157) and a weighted digraph with n nodes, we can define the concept of k -step node transition.

Definition 2. Given a $M_T = [m_{ij}] \in R^{n \times n}$, the transition matrix of the Markov chain associated with the weighted adjacency matrix, A_D^w , of a weighted digraph, G_D^w , with n nodes, each $m_{ij}^{(k)}$ element of matrix $(M_T)^K$ is called the k -step node transition, since they give the probabilities of transitions from x_i to x_j in k time steps, where $(M_T)^K$ is the power k of matrix M_T , $i, j = 1, \dots, n$ and $k \geq 1$.

Remark 1. In Definition 2, if $k = 1$, then $m_{ij}^{(1)} = m_{ij}$, and so m_{ij} is called node transition.

Based on Theorem 8.3 [20] (p. 156) and a weighted digraph with n nodes, we can define the probability of each node after the k times step in the network.

Definition 3. Given a $M_T = [m_{ij}] \in R^{n \times n}$, the transition matrix of the Markov chain associated with the A_D^w of a G_D^w with n nodes, the probability of all the nodes in the network after k times steps is obtained by:

$$p(k) = p(0) \times (M_T)^K, \tag{2}$$

where $p(0) = [p_i(0)] \in R^{1 \times n}$ and $p_i(0) = p(X = x_i)$, $i = 1, \dots, n$.

Considering the mathematical concept of entropy of a random variable, defined on the sample space of a random experiment and taking on a finite number [17] (p. 51) [21] (p. 19), and a weighted digraph with n nodes, we define the concept of node out-entropy.

Definition 4. Given a weighted digraph, G_D^w , with n nodes, $E^{out}(n_i)$ is called the node out-entropy of a node n_i , and is determined by:

$$E^{out}(n_i) = - \sum_{j=1}^n m_{ij} \log_2 m_{ij}, \tag{3}$$

where m_{ij} are the elements of M_T associated with the A_D^w of G_D^w , and $w_{ij} \in A_D^w$, $i, j = 1, \dots, n$.

Remark 2. In Definition 4, if we replace m_{ij} with m_{ji} , we obtain the concept node in-entropy, $E^{in}(n_i)$.

Based on the mathematical concept of conditional entropy [21] (p. 22) [22] (p. 70) and two random variables X and Y , such that the pair of transmitter and receiver (X, Y) is the joint distribution of the Markov chain associated with the weighted adjacency matrix of a weighted digraph with n nodes, we can define the concept of node transition out-entropy.

Definition 5. Given a weighted digraph, G_D^w , with n nodes, the $E_T^{out}(n_i)$ is called node transition out-entropy of a node n_i , and is determined by:

$$E_T^{out}(n_i) = E^{out}(n_i) \times \frac{\sum_{j=1}^n w_{ij}}{L_D^w}, \tag{4}$$

where $L_D^w = \sum_{i=1}^n \sum_{j=1}^n w_{ij}$ is the total links of G_D^w , and $E^{out}(n_i)$ is the node out-entropy of node n_i , $i, j = 1, \dots, n$.

Remark 3. In Definition 5, if we replace w_{ij} with w_{ji} , and m_{ij} with m_{ji} , we obtain the concept of node transition in-entropy, $E_T^{in}(n_i)$.

Considering the mathematical concept of relative entropy between two random variables X and Y , such that the pair of transmitter and receiver (X, Y) is the joint distribution of the Markov chain associated with the weighted adjacency matrix of a weighted digraph with n nodes [21] (p. 248) [23] (p. 25), we can define the concept of the relative out-entropy between two nodes.

Proposition 1. Given a weighted digraph, G_D^w , with n nodes and two distributions, $p_X(x_k) = \{m_{ik} : i, k = 1, \dots, n\}$ and $q_X(x_k) = \{m_{jk} : j, k = 1, \dots, n\}$, with two random variables X and Y , such that the pair of transmitter and receiver (X, Y) is the joint distribution of the Markov chain associated with the A_D^w of G_D^w , the relative out-entropy between two nodes n_i and n_j , $E_R^{out}(n_i||n_j)$, is obtained by:

$$E_R^{out}(n_i||n_j) = -E^{out}(n_i) - \sum_{k=1}^n m_{ik} \log_2 m_{jk}, \tag{5}$$

where $E^{out}(n_i)$ is the node out-entropy of a node n_i , and m_{ik} and m_{jk} are the elements of M_T associated with the A_D^w of G_D^w , $i, j = 1, \dots, n$.

Proof. The relative entropy between two distributions [23] (p. 247) is defined as:

$$E_R^{out}(n_i||n_j) = \sum_{k=1}^n p_X(x_k) \log_2 \left(\frac{p_X(x_k)}{q_X(x_k)} \right). \tag{6}$$

Considering $p_X(x_k) = \{m_{ik} : i, k = 1, \dots, n\}$ and $q_X(x_k) = \{m_{jk} : j, k = 1, \dots, n\}$, we can write:

$$E_R^{out}(n_i||n_j) = \sum_{k=1}^n m_{ik} \log_2 \left(\frac{m_{ik}}{m_{jk}} \right) = \sum_{k=1}^n m_{ik} \log_2 m_{ik} - \sum_{k=1}^n m_{ik} \log_2 m_{jk}. \tag{7}$$

By expression 3, we can write:

$$\sum_{k=1}^n m_{ik} \log_2 m_{ik} - \sum_{k=1}^n m_{ik} \log_2 m_{jk} = -E^{out}(n_i) - \sum_{k=1}^n m_{ik} \log_2 m_{jk}. \tag{8}$$

□

Remark 4. In Proposition 1, if we replace ik with ki and jk with kj , we obtain the concept relative in-entropy, $E_R^{in}(n_i||n_j)$.

Considering the mathematical concept of conditional entropy between two random variables X and Y , such that the pair of transmitter and receiver (X, Y) is the joint distribution of the Markov chain associated with the weighted adjacency matrix of a weighted digraph with n nodes [21] (p. 22), [22] (p. 70), we can define the concept of network transition out-entropy.

Proposition 2. Given a weighted digraph, G_D^w , with n nodes and two random variables X and Y , such that the pair of transmitter and receiver (X, Y) is the joint distribution of the Markov chain associated with the A_D^w of G_D^w , E_{NT}^{out} is the network transition out-entropy of a G_D^w and is determined by:

$$E_{NT}^{out} = \sum_{i=1}^n E_T^{out}(n_i), \tag{9}$$

where $E_T(I)$ is the node transition entropy of a node I_i , $i = 1, \dots, n$.

Proof. Consider two random variables X and Y , such that the pair of transmitter and receiver (X, Y) is the joint distribution of the Markov chain associated with the A_D^w of G_D^w .

The conditional entropy [22] (p. 70) is defined as:

$$\begin{aligned} E_{NT}^{out} &= - \sum_{i=1}^n \sum_{j=1}^n p(X = x_i, Y = y_j) \log_2(p(Y = y_j | X = x_i)) = \\ &= - \sum_{i=1}^n \sum_{j=1}^n p(X = x_i) p(Y = y_j | X = x_i) \log_2(p(Y = y_j | X = x_i)) \\ &= \sum_{i=1}^n E_T^{out}(n_i) \times p(X = x_i) = \sum_{i=1}^n E^{out}(n_i) \times \frac{\sum_{j=1}^n w_{ij}}{L_D^w} = \sum_{i=1}^n E_T^{out}(n_i). \end{aligned} \tag{10}$$

□

Remark 5. In Proposition 2, if we replace ij with ji and ji with ij , we obtain the concept network transition in-entropy, E_{NT}^{in} .

Considering the mathematical concept of joint entropy [19] (p. 394) [21] (p. 22) [22] (p. 69) and a weighted digraph with n nodes, we can define the concept of total network entropy.

Proposition 3. Given a weighted digraph, G_D^w , with n nodes and two random variables X and Y , such that the pair of transmitter and receiver (X, Y) is the joint distribution of the Markov chain associated with the A_D^w of G_D^w , the total network entropy, E_N , is obtained by:

$$E_N = - \sum_{i=1}^n \sum_{j=1}^n \left(\frac{w_{ij}}{L_D^w} \right) \log_2 \left(\frac{w_{ij}}{L_D^w} \right), \tag{11}$$

where $L_D^w = \sum_{i=1}^n \sum_{j=1}^n w_{ij}$ is the total links of G_D^w , $i, j = 1, \dots, n$.

Proof. Consider two random variables X and Y , such that the pair of transmitter and receiver (X, Y) is the joint distribution of the Markov chain associated with the A_D^w of G_D^w . The joint entropy [19] (p. 236) is defined as:

$$\begin{aligned} E_N &= - \sum_{i=1}^n \sum_{j=1}^n p(X = x_i, Y = y_j) \log_2(p(X = x_i, Y = y_j)), \\ &= - \sum_{i=1}^n \sum_{j=1}^n p(X = x_i) p(Y = y_j | X = x_i) \log_2(p(Y = y_j | X = x_i) \times p(X = x_i)) \\ &= - \sum_{i=1}^n \sum_{j=1}^n \left(m_{ij} \times \frac{\sum_{j=1}^n w_{ij}}{L_D^w} \right) \log_2 \left(m_{ij} \times \frac{\sum_{j=1}^n w_{ij}}{L_D^w} \right) = - \sum_{i=1}^n \sum_{j=1}^n \left(\frac{w_{ij}}{L_D^w} \right) \log_2 \left(\frac{w_{ij}}{L_D^w} \right). \end{aligned} \tag{12}$$

□

Considering the mathematical concept of relative entropy between two random variables X and Y , such that the pair of transmitter and receiver (X, Y) is the joint distribution of the Markov chain associated with the weighted adjacency matrix of a weighted digraph with n nodes [21] (p. 247) [23] (p. 25), we can define the concept of network relative out-entropy.

Proposition 4. Given a weighted digraph, G_D^w , with n nodes and two random variables X and Y , such that the pair of transmitter and receiver (X, Y) is the joint distribution of the Markov chain associated with the A_D^w of G_D^w . The network relative out-entropy, E_{NR}^{out} , is obtained by:

$$E_{NR}^{out} = \sum_{i=1}^n \sum_{j=1}^n |E_R(n_i || n_j)|, \tag{13}$$

where $i, j = 1, \dots, n$.

Proof. Consider two distributions, $p_X(x_k) = \{m_{ik} : i, k = 1, \dots, n\}$ and $q_X(x_k) = \{m_{jk} : k, j = 1, \dots, n\}$, with two random variables X and Y , such that the pair of transmitter and receiver (X, Y) is the joint distribution of the Markov chain associated with the A_D^w of G_D^w . By the concept of relative entropy [21] (p. 249), we obtain:

$$E_{NR}^{out} = \sum_{i,k=1}^n \sum_{j,k=1}^n \left| p_X(x_{i,k}) \log_2 \left(\frac{p_X(x_{i,k})}{q_X(x_{j,k})} \right) \right| = \sum_{i=1}^n \sum_{j=1}^n |E_R(n_i || n_j)|. \tag{14}$$

□

Remark 6. In Proposition 4, if we consider $E_R^{in}(n_i, n_j)$, we obtain the concept network relative in-entropy, E_{NT}^{in} .

3. A Case Study

The 2015/2016 Champions League final is presented as the case study for the presentation and interpretation of the metrics proposed. The two opposing teams were Real Madrid (RM) and Atletico Madrid (AM). After a 1-1 draw during regular time, RM became champion with a 5-3 win after a penalty shoot-out. For the notational analysis of each pass sequence (transition between nodes), the uPATO platform was used [11,12]. From this, adjacency matrices were computed to calculate the Markov chains, represented in Appendix A as the probability of state out-transition matrices—i.e., the probability of each player (node) passing (transitioning information) to any other player (node) present in the game (network). Note that, before computing the Markov chains, a Laplace Smoothing [18] was performed upon the weighted adjacency matrices. This smoothing technique serves the purpose of avoiding the pitfall of estimating chances that appear impossible. In our case, we use the special case of Laplace Smoothing, where 1 was added to every connection between two distinct nodes, also called add-one smoothing.

The RM starting 11 were the following: player 1: Keylor Navas (1); player 2: Dani Carvajal (15); player 3: Sergio Ramos (4); player 4: Pepe (3); player 5: Marcelo (12); player 6: Casemiro (14); player 7: Luka Modric (19); player 8: Toni Kroos (8); player 9: Karim Benzema (9); player 10: Gareth Bale (11); and player 11: Cristiano Ronaldo (7). The AM starting 11 constituted the following players: player 1: Jan Oblak (13); player 2: Juanfran (20); player 3: Stefan Savic (15); player 4: Diego Godin (2); player 5: Filipe Luís (3); player 6: Augusto Fernández (12); player 7: Koke (6); player 8: Gabi (14); player 9: Antoine Griezmann (7); player 10: Fernando Torres (9); and player 11: Saúl Ñíguez (17).

Tables A1 and A2, presented in the Appendix A section, show RM and AM's state out-transition probability matrices, there are two nodes with probabilities over 0.3. For RM, transitioning between player 2 and 10 had a 0.33 chance of occurring, and there was a 0.34 chance between player 3 and 5. This might mean that when the ball is in 2, the game is vertical through the center, in search for a creative midfielder. Conversely, player 3 searches for a supported, winged play. In AM's game, passing between 2 and 8 had a chance of 0.30, and between 5 and 11 has a chance of 0.37.

The state in-transition probabilities were also calculated (Tables A3 and A4), representing the probability of each player receiving the ball from each other player. Contrarily to the out-transition probabilities, these tables should be read column-wise. Here, the reception probabilities in RM's game only show player 5 receiving the ball from player 3, with a 0.32 probability of occurring. In AM's game, the same happened in only one case. In practical terms, for RM's game, a key interaction between two players (player 3 and 5) may have been identified. Specifically, player 5 predominantly receives the ball from 3, and the latter predominantly passes to 5. In AM's game, it is player 6 that has a greater chance of receiving the ball from 11.

Having this information might be useful for the opposing team—for example, in order to adapt their defensive tactics to reduce the chances of interaction between players. Knowing how team players interact with each other may identify the prime pathways to feed the attacking players and eventually reduce the chances of creating goal opportunities. This type of empirical observations requires, in our opinion, further analysis in order to produce ratios and cut-off values.

Applying a nonlinear-based approach to state transition matrices in football has not been done yet. Theoretically, higher levels of entropy reflect greater variability of node transitions. Node Transition Entropy was calculated for each node of the team to represent the degree of transition variability of each player, based on the probability of it happening. The more positive a value is, the more that node contributes to the overall entropy of the network when compared to the other nodes.

In the examples below (Tables A5 and A6), we selected the cut-off value of |0.90| for the Relative Out-Transition Entropies. RM showed more values above our stipulated limit, showing that there were more players with more chaotic passing probabilities. Regarding the AM relative entropy values,

only two players showed more chaotic behavior in their passing probabilities, and this difference is against the goalkeeper (node 1). This shows that AM players were more consistent, and therefore more predictable, in their passing patterns. It is important to note that the defined Relative Out/In-Transition Entropies can both have negative and positive values. Negative values indicate that the first node in the pair is more chaotic—i.e., $E_R^{out}(n_i||n_j) < 0$ means that n_i is more chaotic than n_j . The inverse also applies, where if $E_R^{out}(n_i||n_j) > 0$, n_i is less chaotic than n_j . In the case of $E_R^{out} = 0$, both nodes have the same chaoticity, meaning that they add the same value to the network as a whole.

Similarly, in-transition entropy reflects the degree of variability in receiving the ball (Tables A7 and A8). Applying the same criteria, we can note that for RM, two players (player 5 and 8) showed chaoticity in the probability of receiving passes from other teammates. In line with what has been previously observed, AM players also had smaller entropy values for this metric, again showing a more predictable behavior. In this case, no AM player presented values above the selected cut-off value.

The final step for this case study is to calculate team metrics for each team, presented in Table 1. Network Relative Out-Entropy (NROE) and Network Relative In-Entropy (NRIE) show the consistency of the network. Here, higher values of entropy reflect a smaller consistency of interactions between players or, in other terms, a more unpredictable passing pattern. In our football example, both values were higher for Real Madrid, which means that the team was more unpredictable both in passing and in receiving the ball from the other players. It has been shown that unpredictable passing patterns may create more goal-scoring opportunities and contribute to ball possession [1,24]. This may be due to the fact that the team has more players involved in each passing sequence, increasing the unpredictability of each play. In this case, success may be due to either the number of passes or the time it possessed the ball, and not to higher entropy values [25]. That was not the case, as RM had fewer passes than AM. We may regard entropy, therefore, as a positive feature in the game. Real Madrid’s unpredictability in passing may have increased the chances of goal-scoring opportunities, potentially having contributed to the win.

Table 1. Team entropy metrics.

	Real Madrid	Atletico Madrid
Network Relative Out-Entropy	38.90	34.10
Network Relative In-Entropy	39.55	33.38
Network Transition Out-Entropy	3.03	2.93
Network Transition In-Entropy	3.00	2.93
Total Network Entropy	6.44	6.22

The concept of entropy to model social interactions has already been used [26–28]. Newman and Vilenchik (2019) used the concept of relative entropy to model the interactions of players passing the ball in football, having found that when comparing two opposing teams, higher entropy values lead to more chances of creating goal-scoring opportunities [27]. The authors limit their analysis to whether an interaction between players occurred, without defining the type of interaction—passing or receiving the ball. In other words, direction did not matter. In football, however, variability in passing patterns may come from these two actions and they may lead to different conclusions: a team can be more unpredictable in passing than in receiving the ball. Fewer players passing the ball to many players is different from many players passing the ball to a few. Concomitantly, game strategy must take this into account and adapt accordingly.

The division between the Network Transition Out-Entropy (NTOE) and Network Transition In-Entropy (NTIE) separates the teams’ out and in interactions. NTOE, therefore shows the entropy of

the network in sending the information to other nodes. Conversely, NTIE aims to show the entropy of receiving information from other nodes. For our football example, NTOE and NTIE are the entropy levels of the players' probability of passes and receptions, respectively.

NTOE and NTIE were slightly higher for Real Madrid, showing again that this team was more unpredictable in the interactions. Against an opponent, higher levels of entropy may prevent the defending team from learning the patterns the other team. It is worth noticing that the division in the out and in entropies allows, in our opinion, a deeper understanding of the team's dynamics by extending the network analysis to more than the existence (or not) of interactions between nodes, taking into account the direction of information and its probability of occurring.

Total Network Entropy (TNE) reflects the degree of variability and therefore the unpredictability of the team as a whole. Real Madrid was slightly more unpredictable than Atletico, with a TNE of 6.44 vs. 6.22. The unpredictability of passing patterns however, does not happen isolated from other actions; it is, rather, dependent on the interactions between rivals and teammates, space occupation, time, tactics, and game situation [10,29]. This leads to the need to constantly analyze how the team behaves during the different phases of the game, as they may behave differently. For instance, if, when winning, ball possession is kept using the same passing patterns or variability, the unpredictability decreases.

4. Conclusions

In summary, the proposed metrics present a new method to analyze the passing interactions in a football match. From a Bayesian perspective, knowing the probability of one player passing to other may provide the opposing team the information it needs to anticipate and create more opportunities for stealing ball possession. The innovation of this approach lies in the division between ball passing and reception probability patterns. An entropy-based analysis to in and out-transition matrices attempts to merge the Bayesian and non-linear approaches into one, increasing the tools available for match analysts. With this new approach, football analysts may also start providing information to the coach staff during the game, helping identify the players where passing or receiving is more likely to occur, as well as give the degree of unpredictability of each of the actions. Future studies should provide ratios and cut-of values for the state transition probabilities, helping decide when one passing pattern is predictable or not.

Author Contributions: F.M.: Conceptualization, Methodology, Formal Analysis, Writing, Project administration, Funding acquisition. R.G.: Conceptualization, Validation; Formal Analysis, Investigation, Writing, Visualization, Project Administration. V.L.: Methodology, Software, Validation, Visualization, Writing. F.S.: Resources, Data Curation, Writing, Funding Acquisition, Supervision. R.M.: Conceptualization, Resources, Supervision, Funding Acquisition. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by FCT/MCTES through national funds and when applicable co-funded EU funds under the project UIDB/50008/2020. The APC was funded by IPC/ESEC—UNICID.

Conflicts of Interest: The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analysis, or interpretation of data; in the writing of the manuscript; or in the decision to publish the results.

Appendix A

Table A1. Real Madrid’s matrix of state out-transition probabilities.

NODE	1	2	3	4	5	6	7	8	9	10	11
1	0	0.15	0.11	0.11	0.15	0.04	0.11	0.07	0.07	0.04	0.15
2	0.12	0	0.05	0.16	0.02	0.05	0.14	0.02	0.05	0.33	0.07
3	0.09	0.04	0	0.09	0.34	0.08	0.09	0.15	0.04	0.04	0.04
4	0.12	0.1	0.24	0	0.04	0.18	0.04	0.02	0.04	0.12	0.08
5	0.02	0.04	0.12	0.06	0	0.14	0.24	0.1	0.06	0.02	0.22
6	0.02	0.07	0.13	0.04	0.15	0	0.06	0.19	0.07	0.19	0.09
7	0.03	0.09	0.09	0.03	0.09	0.09	0	0.09	0.24	0.12	0.12
8	0.02	0.09	0.05	0.07	0.14	0.14	0.07	0	0.14	0.09	0.19
9	0.03	0.14	0.03	0.03	0.14	0.03	0.17	0.21	0.00	0.1	0.1
10	0.02	0.18	0.07	0.09	0.13	0.09	0.18	0.05	0.07	0	0.13
11	0.03	0.06	0.03	0.06	0.12	0.15	0.09	0.18	0.12	0.15	0

Table A2. Atletico Madrid’s matrix of state out-transition probabilities

NODE	1	2	3	4	5	6	7	8	9	10	11
1	0	0.11	0.15	0.15	0.22	0.04	0.15	0.04	0.07	0.04	0.04
2	0.02	0	0.05	0.02	0.05	0.02	0.02	0.30	0.26	0.05	0.21
3	0.06	0.2	0	0.12	0.04	0.08	0.10	0.18	0.06	0.02	0.14
4	0.08	0.03	0.18	0	0.1	0.1	0.07	0.13	0.07	0.02	0.22
5	0.02	0.02	0.03	0.12	0	0.03	0.03	0.22	0.14	0.02	0.37
6	0.02	0.07	0.07	0.12	0.19	0	0.07	0.10	0.14	0.02	0.19
7	0.03	0.13	0.07	0.07	0.13	0.10	0	0.03	0.17	0.13	0.13
8	0.01	0.13	0.07	0.08	0.17	0.05	0.09	0	0.12	0.03	0.25
9	0.02	0.09	0.03	0.02	0.19	0.05	0.12	0.19	0.00	0.07	0.22
10	0.07	0.13	0.07	0.07	0.07	0.07	0.13	0.13	0.20	0	0.07
11	0.01	0.07	0.06	0.08	0.18	0.12	0.03	0.24	0.18	0.03	0

Table A3. Real Madrid’s matrix of state in-transition probabilities.

NODE	1	2	3	4	5	6	7	8	9	10	11
1	0	0.10	0.07	0.09	0.07	0.02	0.06	0.04	0.05	0.02	0.08
2	0.22	0	0.05	0.22	0.02	0.05	0.12	0.02	0.05	0.28	0.06
3	0.22	0.05	0	0.16	0.32	0.09	0.10	0.18	0.05	0.04	0.04
4	0.26	0.13	0.29	0	0.04	0.21	0.04	0.02	0.05	0.12	0.08
5	0.04	0.05	0.15	0.09	0	0.16	0.23	0.11	0.08	0.02	0.22
6	0.04	0.10	0.17	0.06	0.14	0	0.06	0.22	0.11	0.20	0.10
7	0.04	0.08	0.07	0.03	0.05	0.07	0	0.07	0.22	0.08	0.08
8	0.04	0.10	0.05	0.09	0.11	0.14	0.06	0	0.16	0.08	0.16
9	0.04	0.10	0.02	0.03	0.07	0.02	0.10	0.13	0	0.06	0.06
10	0.04	0.25	0.10	0.16	0.12	0.12	0.19	0.07	0.11	0	0.14
11	0.04	0.05	0.02	0.06	0.07	0.12	0.06	0.13	0.11	0.10	0

Table A4. Atletico Madrid’s matrix of state in-transition probabilities.

NODE	1	2	3	4	5	6	7	8	9	10	11
1	0	0.06	0.1	0.09	0.08	0.03	0.10	0.01	0.03	0.05	0.01
2	0.06	0	0.05	0.02	0.03	0.03	0.03	0.15	0.15	0.1	0.09
3	0.19	0.2	0	0.14	0.03	0.11	0.13	0.11	0.04	0.05	0.07
4	0.31	0.04	0.28	0	0.08	0.16	0.10	0.09	0.06	0.05	0.13
5	0.06	0.02	0.05	0.16	0	0.05	0.05	0.15	0.11	0.05	0.21
6	0.06	0.06	0.08	0.12	0.11	0	0.08	0.05	0.08	0.05	0.08
7	0.06	0.08	0.05	0.05	0.05	0.08	0.00	0.01	0.07	0.19	0.04
8	0.06	0.26	0.18	0.19	0.23	0.13	0.23	0	0.17	0.14	0.24
9	0.06	0.1	0.05	0.02	0.15	0.08	0.18	0.13	0	0.19	0.13
10	0.06	0.04	0.03	0.02	0.01	0.03	0.05	0.02	0.04	0	0.01
11	0.06	0.14	0.15	0.19	0.23	0.32	0.08	0.27	0.24	0.14	0

Table A5. Real Madrid’s individual Relative Out-Transition Entropy

NODE	1	2	3	4	5	6	7	8	9	10	11
1	0	0.24	0.06	0.22	-0.27	0.15	-0.37	-0.11	-0.03	-0.12	-0.14
2	0.54	0	0.94	0.05	1.31	0.64	0.07	0.76	0.73	-0.37	0.40
3	0.09	1.28	0	0.97	-0.43	0.29	0.07	-0.03	0.29	0.48	0.26
4	0.21	0.42	0.00	0	0.65	-0.13	-0.37	0.63	0.90	0.25	0.54
5	0.40	0.72	0.48	0.57	0	0.18	-0.91	0.10	0.38	0.27	-0.19
6	0.35	0.60	0.05	0.70	0.22	0	-0.07	-0.31	-0.07	-0.24	-0.05
7	0.41	0.69	0.71	0.83	0.64	0.38	0	-0.02	-0.46	0.14	0.00
8	0.23	0.37	0.56	0.23	-0.10	-0.13	-0.69	0	-0.04	-0.19	-0.41
9	0.20	0.39	0.32	0.95	0.01	0.15	-0.39	-0.34	0	-0.10	-0.10
10	0.02	0.14	0.29	0.35	-0.24	0.19	-0.68	0.04	0.02	0	-0.01
11	0.54	0.62	0.32	0.62	0.26	-0.36	-0.69	-0.39	-0.14	-0.12	0

Table A6. Atletico Madrid’s individual Relative Out-Transition Entropy.

NODE	1	2	3	4	5	6	7	8	9	10	11
1	0	0.86	0.04	0.00	0.26	0.12	-0.45	0.12	0.39	0.31	0.37
2	1.46	0	0.57	0.52	0.01	0.45	0.65	-0.47	-0.48	0.37	-0.39
3	0.47	-0.12	0	0.09	0.58	0.07	-0.24	-0.30	0.23	0.17	-0.03
4	0.42	0.47	-0.29	0	0.22	-0.05	-0.17	-0.06	0.16	0.44	-0.23
5	1.58	0.25	0.39	0.00	0	0.33	0.59	-0.33	0.10	0.79	-0.67
6	0.30	0.23	0.16	-0.16	-0.42	0	-0.13	-0.32	-0.23	0.28	-0.42
7	0.49	0.01	0.52	0.57	0.38	0.43	0	0.24	-0.12	-0.14	0.04
8	0.47	0.25	0.23	0.13	-0.02	-0.08	-0.49	0	-0.11	0.39	-0.38
9	0.78	0.33	0.36	0.37	-0.11	0.15	-0.18	-0.43	0	0.32	-0.30
10	0.13	-0.02	-0.08	0.15	0.46	-0.02	-0.52	-0.19	-0.30	0	0.08
11	0.67	0.34	0.36	0.27	0.06	-0.12	-0.03	-0.34	-0.11	0.22	0

Table A7. Real Madrid’s individual Relative in-Transition Entropy.

NODE	1	2	3	4	5	6	7	8	9	10	11
1	0	-0.06	-0.23	-0.58	0.96	0.10	0.68	1.15	0.88	0.35	0.84
2	0.50	0	0.22	0.00	0.20	-0.33	0.10	0.21	-0.08	-0.23	0.04
3	0.40	0.21	0	-0.37	0.49	-0.70	0.53	0.92	0.58	0.32	0.29
4	0.14	-0.38	0.19	0	0.25	0.19	0.08	0.36	0.44	0.03	0.24
5	0.40	0.56	-0.36	0.33	0	-0.12	0.47	-0.15	0.37	0.48	0.58
6	0.58	0.42	0.17	-0.37	0.18	0	0.64	0.45	0.61	0.54	-0.06
7	0.59	-0.03	0.21	0.06	-0.32	0.12	0	0.31	0.19	0.24	-0.01
8	0.78	0.59	0.07	0.64	-0.18	-0.46	0.25	0	0.01	0.54	0.12
9	0.74	0.17	0.35	0.37	0.22	-1.03	-0.33	-0.10	0	0.12	-0.20
10	0.46	-0.37	0.67	0.11	1.22	-0.32	0.55	0.75	0.44	0	0.40
11	0.59	0.06	0.15	0.07	-0.30	-0.48	-0.11	-0.03	0.05	0.37	0

Table A8. Atletico Madrid's individual Relative in-Transition Entropy.

NODE	1	2	3	4	5	6	7	8	9	10	11
1	0	0.54	−0.42	−0.49	0.75	−0.20	0.26	0.35	0.58	0.63	0.30
2	0.39	0	−0.35	0.13	0.34	−0.34	−0.21	−0.19	0.15	0.26	0.07
3	−0.01	0.53	0	−0.58	0.08	−0.27	0.21	0.09	0.31	0.54	−0.01
4	0.42	0.35	−0.04	0	−0.24	−0.27	0.25	−0.08	0.26	0.55	−0.13
5	0.53	0.15	0.18	0.16	0	−0.47	0.04	−0.19	−0.24	0.26	−0.27
6	0.60	0.46	−0.10	−0.18	0.21	0	0.63	−0.28	0.14	0.59	−0.59
7	0.08	0.12	−0.17	0.06	0.09	0.03	0	−0.24	−0.14	0.23	0.15
8	0.70	0.26	0.32	0.39	0.14	0.17	0.89	0	−0.16	0.55	−0.45
9	0.66	−0.12	0.25	0.11	0.00	−0.29	0.44	−0.30	0	0.23	−0.45
10	0.54	−0.09	0.36	0.58	0.09	−0.59	−0.41	0.19	−0.46	0	−0.07
11	0.67	0.49	0.34	0.11	−0.26	0.20	0.32	−0.35	−0.01	0.65	0

References

- Clemente, F.M.; Martins, F.M.L.; Kalamaras, D.; Wong, D.P.; Mendes, R. General network analysis of national soccer teams in FIFA World Cup 2014. *Int. J. Perf. Anal. Sport* **2015**, *15*, 80–96. [CrossRef]
- Clemente, F.M.; Couceiro, M.; Martins, F.M.L.; Mendes, R. Using Network Metrics in Soccer: A Macro-Analysis. *J. Hum. Kinet.* **2015**, *45*, 123–134. [CrossRef] [PubMed]
- McLean, S.; Salmon, P.M.; Gorman, A.D.; Read, G.J.M.; Solomon, C. What's in a game? A systems approach to enhancing performance analysis in football. *PLoS ONE* **2017**, *12*, e0172565. [CrossRef] [PubMed]
- Passos, P.; Davids, K.; Araujo, D.; Paz, N.; Minguens, J.; Mendes, L. Network as a novel tool for studying team ball sports as complex social system. *J. Sci. Med. Sport* **2011**, *14*, 170–176. [CrossRef] [PubMed]
- McLean, S.; Salmon, P.M.; Gorman, A.D.; Stevens, N.J.; Solomon, C. A social network analysis of the goal scoring passing networks of the 2016 European Football Championships. *Hum. Mov. Sci.* **2018**, *57*, 400–408. [CrossRef] [PubMed]
- Ribeiro, J.; Silva, P.; Duarte, R.; Davids, K.; Garganta, J. Team Sports Performance Analysed Through the Lens of Social Network Theory: Implications for Research and Practice. *Sports Med.* **2017**, *47*, 1689–1696. [CrossRef] [PubMed]
- Lusher, D.; Robins, G.; Kremer, P. The Application of Social Network Analysis to Team Sports. *Meas. Phys. Educ. Exerc.* **2010**, *14*, 211–224. [CrossRef]
- Aquino, R.; Carling, C.; Palucci, L.H.; Martins, G.; Jabor, G.; Machado, G.; Santiago, P.; Garganta, J.; Puggina, E. Influence of Situational Variables, Team Formation, and Playing Position on Match Running Performance and Social Network Analysis in Brazilian Professional Soccer Players. *J. Strength Cond. Res.* **2020**, *34*, 808–817. [CrossRef] [PubMed]
- Castellano, J.; Echeazarra, I. Network-based centrality measures and physical demands in football regarding player position: Is there a connection? A preliminary study. *J. Sports Sci.* **2019**, *37*, 2631–2638. [CrossRef] [PubMed]
- Arriaza-Ardiles, E.; Martín-González, J.M.; Zuniga, M.D.; Sánchez-Flores, J.; de Saa, Y.; García-Manso, J.M. Applying graphs and complex networks to football metric interpretation. *Hum. Mov. Sci.* **2018**, *57*, 236–243. [CrossRef] [PubMed]
- Martins, F.M.L.; Silva, F.; Clemente, F.; Gomes, A.J.P.; Correia, A.; Nguyen, Q.; Sequeiros, J.B.; Ribeiro, J.S.; Lopes, V.F. Ultimate Performance Analysis Tool (uPATO). 2018. Available online: <http://uPATO.it.ubi.pt> (accessed on 12 March 2020).
- Silva, F.; Nguyen, Q.; Correia, A.; Clemente, F.; Martins, F.M.L. *Ultimate Performance Analysis Tool (uPATO): Implementation of Network Measures Based on Adjacency Matrices for Team Sports*; Springer International Publishing: Dordrecht, The Netherlands, 2019. [CrossRef]
- Yamamoto, K.; Narizuka, T. Examination of Markov-chain approximation in football games based on time evolution of ball-passing networks. *Phys. Rev. E* **2018**, *98*, 052314. [CrossRef]
- Narizuka, T.; Yamamoto, K.; Yamazaki, Y. Statistical Properties of Position-Dependent Ball-Passing Networks in Football Games. *Phys. A* **2014**, *412*, 157–168. [CrossRef]

15. Wasserman, S.; Faust, K. *Social Network Analysis: Methods and Applications*; Cambridge University Press: New York, NY, USA, 1994.
16. Clemente, F.; Martins, F.; Mendes, R. *Social Network Analysis Applied to Team Sports Analysis*; Springer International Publishing: Dordrecht, The Netherlands, 2016.
17. Manning, C.D.; Raghavan, P.; Schütze, M. *Introduction to Information Retrieval*; Cambridge University Press: Cambridge, UK, 2008.
18. Bishop, C.M. *Pattern Recognition and Machine Learning*; Springer: Dordrecht, The Netherlands, 2006.
19. Shannon, C. A Mathematical Theory of Communication. *Bell Syst. Tech. J.* **1948**, *27*, 379–423. [[CrossRef](#)]
20. Tuckwell, H. *Elementary Applications of Probability Theory*; Chapman and Hall Ltd.: New York, NY, USA, 1988.
21. Han, T.S.; Kobayashi, K. *Mathematics of Information and Coding*; American Mathematical Society: Providence, RI, USA, 2002; ISBN 978-0-8218-4256-0.
22. Rao, C.R.; Gudivada, V.N. *Computational Analysis and Understanding of Natural Languages: Principles, Methods and Applications*; Elsevier: North Holland, The Netherlands, 2018.
23. Marinescu, D.C.; Marinescu, G.M. *Classical and Quantum Information*; Academic Press: Burlington, VT, USA, 2011.
24. Pina, T.J.; Paulo, A.; Araújo, D. Network Characteristics of Successful Performance in Association Football. A Study on the UEFA Champions League. *Front. Psychol.* **2017**, *8*, 1173. [[CrossRef](#)]
25. Neuman, Y.; Israeli, N.; Vilenchik, D.; Cohen, Y. The Adaptive Behavior of a Soccer Team: An Entropy-Based Analysis. *Entropy* **2018**, *20*, 758. [[CrossRef](#)]
26. Neuman, Y.; Vilenchik, D.; Israeli, N. From physical to social interactions: The relative entropy model. *Sci. Rep.* **2020**, *10*, 1–8. [[CrossRef](#)] [[PubMed](#)]
27. Neuman, Y.; Vilenchik, D. Modeling Small Systems through the Relative Entropy Lattice. *IEEE Access* **2019**, *7*, 43591–43597. [[CrossRef](#)]
28. Martínez, J.H.; Garrido, D.; Herrera-Diestra, J.L.; Busquets, J.; Sevilla-Escoboza, R.; Buldú, J.M. Spatial and Temporal Entropies in the Spanish Football League: A Network Science Perspective. *Entropy* **2020**, *22*, 172. [[CrossRef](#)]
29. Martín González, J.M.; de Saá Guerra, Y.; García Manso, J.M.; Arriaza, E. Design and flow in basketball. *Int. J. Heat Technol.* **2016**, *34*, 51–58. [[CrossRef](#)]



© 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).