Chaos, Solitons and Fractals 174 (2023) 113778

Contents lists available at ScienceDirect

Chaos, Solitons and Fractals

journal homepage: www.elsevier.com/locate/chaos

CMPN: Modeling and analysis of soccer teams using Complex Multiplex Passing Network

Arash Beheshtian-Ardakani [a,1], Mostafa Salehi [a,b,*,2], Rajesh Sharma [c,d,2]

[a] Faculty of New Sciences and Technologies, University of Tehran, Tehran, Iran
[b] School of Computer Science, Institute for Research in Fundamental Science (IPM), P.O. Box 19395-5746, Tehran, Iran
[c] Institute of Computer Science, University of Tartu, Tartu, Estonia
[d] Indraprastha Institute of Information Technology, Delhi, India

ARTICLE INFO

Keywords:
Multilayer network
Performance analysis
Core player
Long short-term memory
Outcome prediction
Soccer

ABSTRACT

Nowadays, coaches exploit data analysis in soccer (football) matches to plan their strategies against opponents. Network science, a subdomain of data analytics, is widely used to analyze soccer matches by treating players as nodes and passes between them as edges. However, single-layer methods for analyzing games overlook critical information by aggregating different types of passes into one layer. This paper introduces a new model called the Complex Multiplex Passing Network (CMPN) for analyzing team sports performance, with a focus on soccer matches. We utilized a real-world dataset to construct the multilayer structure of the CMPN. Each layer represents a specific type of pass between players. Using the CMPN, we conducted various analysis tasks at different topological scales. Firstly, we identified the core players of teams by calculating the PageRank versatility of each player. Next, we discovered the types of passes between trios of players based on multilayer motifs. Additionally, we measured similarities between passing tactics using the Pearson inter-layer assortativity measure. Finally, we employed a long short-term memory network to predict the outcomes of attacking plays using the CMPN model. The predictions achieved over 90% accuracy and approximately 70% F-measure. These findings offer practical value to coaches and performance analysts, as they enable appropriate planning by predicting playing styles in different competitions and neutralizing the strategies of opposing teams.

1. Introduction

During the FIFA World Cup Qatar 2022™, FIFA shared insights, metrics, and performance data of various games which happened during the tournament [1]. Data scientists of soccer clubs can use analytical methods on this data to prepare critical information like patterns of plays, tactics, and key players of opponent teams for coaches to improve team performance strategies [2]. In soccer, one of the most important methods for analyzing a team's performance is by considering players and passes between them through passing networks. Sports scientists use passing networks to (i) characterize a player's role in a team, (ii) find players with similar passing skills, and (iii) determine the importance of players [3]. The following topological scales are used in the literature for analyzing a team's performance [4]: (1) The microscale level investigates the importance and role of players in their teams' networks. (2) The mesoscale level discovers motifs and communities of players in their teams. (3) The macroscale level considers the whole network of a team to find out the team's tactics and strategies.

In this work, we propose a *complex multiplex passing network (CMPN)* by considering each type of pass between players as a separate layer. We create a multiplex network for each team separately. Players connect with intra-layer edges in every layer, and replicas between layers connect with inter-layer edges. So far, the methods presented for analyzing passing networks of teams only involved single-layer (monoplex) networks. These methods lose lots of information due to their simplification by neglecting the importance of distinct types of passes by aggregating them into just one layer. This approach changes topological and dynamical properties throughout the passing networks. Consequently, they are prone to achieve misleading results in determining opponent key players, tactics, and strategies. The main contributions of this study are as follows:

1. A novel CMPN model is proposed by considering different types of passes made by team players in a game. The proposed model can help

* Corresponding author at: Faculty of New Sciences and Technologies, University of Tehran, Tehran, Iran.
E-mail addresses: arash.beheshtian@ut.ac.ir (A. Beheshtian-Ardakani), mostafa_salehi@ut.ac.ir (M. Salehi), rajesh.sharma@ut.ee (R. Sharma).

[1] PhD Candidate.
[2] Associate Professor.

https://doi.org/10.1016/j.chaos.2023.113778
Received 17 December 2022; Received in revised form 28 June 2023; Accepted 30 June 2023
Available online 17 July 2023

data scientists to analyze team sports performance at the microscale, mesoscale, and macroscale levels.

2. At the microscale level, we identify *core players* using the proposed CMPN. We define the core players of teams as ones having an essential effect on implementing different tactics. We determine these core players by preserving the information in the multilayer nature structure of team sports. We use the PageRank versatility measure on the proposed CMPN to explore the players with the most impact in playing teams' tactics.

3. At the mesoscale level, we find multilayer motifs using the proposed CMPN to discover patterns leveraging different tactical passes by trios of teams. We detect motif building blocks of teams to find frequent passing patterns that consist of different layers in the proposed CMPN.

4. At the macroscale level, we present a method for calculating similarities between different passing tactics of a team using the proposed CMPN. We determine similarities between distinct passing tactics of teams that compose layers in the CMPN. We compare the Pearson inter-layer assortativity measure on the out-degree of every node, i.e., the passes made by each player, to determine similar tactics in the CMPN. Finally, We present a novel method for *Outcome Prediction of Attacking Plays (OPAP)* in soccer competitions. Our approach utilizes the Attacking plays Detector (APD) algorithm to extract the trajectory of passes from the moment of ball recovery by a team until the ball is either shot or successfully scored at the opponent's goal. By leveraging the CMPN model to capture various types of passes and incorporating the LSTM model, we achieve a prediction accuracy of over 90% and an approximate F-measure of 70% when analyzing attacking plays in real-world leagues, teams, and the World Cup.

The results prove the advantages of the proposed CMPN over an aggregated single-layer network in all analytical scales by exploiting the multilayer structure of passing networks.

The remainder of this paper is structured as follows: Section 2 provides a comprehensive review of the related works in the literature pertaining to the topic. Section 3 presents the novel methods proposed in this study, including the CMPN model and the OPAP method. Section 4 presents the results of applying these methods, showcasing their effectiveness and performance. Section 5 offers a detailed discussion highlighting the superiority of the CMPN model and its advantages in analyzing team sports performance. Finally, Section 6 concludes this article by summarizing the key findings and contributions.

2. Related works

In this section, we review the literature of passing network studies for team sports performance analysis at all topological scales.

2.1. Microscale analysis

Most works at the microscale level are dedicated to determining the importance level of players in team passing networks. Researchers usually have used centrality measures defined in complex networks for finding teams' key players and their positions in pass sequences [5–7]. One such measure in past works is the degree centrality of a node, which is the number of edges connected to it. In sports analysis, the degree centrality of a player shows the strength of this player in a team [8]. In directed graphs, two types of degree centrality exist, namely out-degree and in-degree centralities. Degree centrality or out-degree centrality is the number of passes made by a player to his or her teammates, and degree prestige or in-degree centrality is the number of passes received by the player from their teammates during a game.

The PageRank algorithm is a recursive idea of importance for assessing the relative importance of nodes in a network [9]. In the context of soccer analysis, the PageRank centrality takes into account the passes received by a player from other influential players. The PageRank centrality of player i in a soccer match is calculated using Eq. (1). In this equation, A_{ji} is the number of passes made from player j to player i, and $L_j^{out} = \sum_k A_{jk}$ is the total number of passes made by player j, i.e., the out-degree of that node. The heuristic parameter p shows the probability that a player passes a ball instead of keeping or shooting it to the goal. In addition, the parameter q gives free popularity to each player in the team. The values of p and q are usually set to 0.85 and 1, respectively [10].

$$x_i = p \sum_{j \neq i} \frac{A_{ji}}{L_j^{out}} x_j + q \tag{1}$$

A player's PageRank score recursively depends on the PageRank scores of his or her teammates. Therefore, all the PageRank scores of a team's players must be calculated simultaneously. The PageRank centrally approximately assigns a player the probability that this player owns the ball after some appropriate number of passes are made and received between team players [2].

2.2. Mesoscale analysis

At the mesoscale level, motifs with three/four players are studied in the literature [4]. Motifs are specific patterns (subgraphs) that are statistically very common in a complex network. The importance of these patterns is measured using their presence in the target network compared to their occurrences in a set of randomly generated networks.

Motif analysis of passing networks shows that frequent specific types of passes between players can be related to unique styles of teams. These motifs also help data scientists to distinguish leaders in passing networks [4]. Flows in the motifs of a passing network explain how players are involved in ball-passing sequences of their team [11].

2.3. Macroscale analysis

At the macroscale level of analysis, various metrics provide a comprehensive understanding of a team's network and its outcomes in matches. These metrics include edge count, edge density, and average path length, which offer insights into the characteristics of teamwork [12]. Metrics such as the clustering coefficient, network average shortest path, and the number of nodes occupied by a team further explain the organization and structure of the team's network [13]. Additionally, the importance of positions within soccer passing networks can be measured by comparing various network attributes across different playing positions [14]. Furthermore, the weighted adjacency matrix of soccer teams is studied to recognize the main players in offensive processes [15].

Predicting soccer match outcomes is crucial for both soccer clubs and betting organizations. To address this, machine learning algorithms are employed to forecast the results of soccer games using various match and player attributes as inputs [16]. Furthermore, integrating domain knowledge specific to soccer into the modeling process enhances the accuracy and reliability of outcome predictions [17]. Ball-passing distribution data and social network analysis help predict the win-lose outcome of soccer matches and provide information such as the need to change passes between specific players for winning the subsequent games [18]. The match outcome and goal scoring are related to the passing network centrality of various playing positions during passing sequences [7]. The main trends of a soccer match based on probabilities of winning, losing, and drawing as home or visiting teams can be determined by incorporating measures of weighted and directed passing networks [19]. Mainly, Long short-term memory (LSTM) networks are used in the literature for soccer match predictions [20,21].

Researchers have also studied trajectories in soccer matches via deep learning algorithms. Trajectory data encodes information about the dynamics of the process that is implicit in the original event data [22]. LSTM networks are beneficial for planning the optimal shooting trajectory of players to score a goal [23]. Deep bidirectional long short-term memory and mixture density networks are utilized on trajectory data for shot prediction in basketball games which could help coaches and players decide when and where to shoot [24].

To our knowledge, no study utilized multilayer networks in team sports performance analysis [4,25], except an idea given by Ramos et al. for analyzing attacking plays in soccer matches [26]. In their work, the concept of an attacking play in soccer was defined as the period from the moment of ball recovery by a team until the ball is either lost or shot at the opponent's goal. Consequently, they drew a temporal bipartite network for each attacking play and projected each network into two separate single-layer networks, i.e., events projection and players projection networks. However, their analysis was limited to projecting the bipartite networks into single-layer networks, neglecting the valuable multilayer structure.

3. Methods

In a single-layer network, a passing network can be represented using a graph $G = (V, E)$ in which nodes V demonstrate the players of a team and edges E show the passes between them. In contrast, multilayer networks encode more information than single-layer networks [27]. A multilayer network can be represented using a quadruple $G = (A, L, V, E)$ in which A and L show the sets of actors and layers, respectively. Nodes and edges of the multilayer network are shown as $V \subseteq A \times L$ and $E \subseteq V \times V$ [28]. Each actor (player in the CMPN) must be present in at least one layer. It is to be noted that each layer is not required to contain all actors. Nodes exist in layers if at least there is an edge connection to them.

Multiplex networks are a special type of multilayer networks with a one-to-one mapping of the nodes in different layers, also called replica nodes. Inter-layer edges can exclusively connect to corresponding replica nodes in a multiplex network. These networks often describe the interactions between the same set of nodes, with each layer characterizing a different type of interaction [29].

We use the Wyscout company dataset for building our CMPN model. This dataset includes soccer logs of events during men's soccer matches for some first-division leagues of European countries in 2017–2018, World Cup 2018, and European Cup 2016 in JSON format [30]. The Wyscout dataset includes the following types of passes [30,31]:

1- Cross: A player passes the ball from a side of the pitch into the penalty area, trying to create some dangerous situation.

2- Head: A player tries to pass the ball with the head.

3- High: These are long-range passes in the air with a specific target or idea.

4- Launch: These are long-range passes in the air without a specific target or idea.

5- Simple: These are short-range or long-range passes on the ground, which are not difficult to be sent.

6- Smart:A player makes a smart pass between two or three opposing players, leading his or her teammates to a suitable attacking position.

We build the layers of the CMPN model based on these six types of passes. Hence, the set of layers in the CMPN is L = {Cross, Head, High, Launch, Simple, Smart}.

Algorithm 1 shows the pseudocode for building the proposed CMPN model. First, we build a Python Pandas DataFrame from spatio-temporal log events named Events (Line 1). Then, we distinguish types of passes from the Events which form the layers of the proposed CMPN and save them in an array named Layers (Line 2). Next, we choose the players of teams and store their events in Actors DataFrame (Line 3). For every layer of CMPN, we build a distinct DataFrame and label Actors in each layer as sender or receiver of each pass in events (Line 5). We calculate players' locations by averaging the length and width positions of the players on the field when they make passes. We use these positional metadata to draw players on a soccer pitch (Line 8).

To accurately calculate multilayer network measures in the CMPN model, we assign edges with a weight of 1 between replica nodes. These edges represent the interconnections between all pairs of layers in the network (Line 13). In each layer of the CMPN, we create directed edges from a sender player to a receiver player for each pass. The weight assigned to these edges represents the number of passes between the corresponding players(Line 15).

Algorithm 1 Building Complex Multiplex Passing Network (CMPN)

Require: Spatio-temporal log events
Ensure: CMPN
1: Events $\Leftarrow$ Spatio-temporal log events
2: Layers $\Leftarrow$ PASS_TYPES(Events)
3: Actors $\Leftarrow$ SELECT_PLAYERS(Events)
4: **for each** pass $\in$ Layers **do**
5: Senders[pass] & Receivers[pass] $\Leftarrow$ SEND_RECEIVE[Actors]
6: **end for**
7: **for each** $P1 \in Senders$ **do**
8: Position[p1] = MEAN(p1[pass][length], p1[pass][width])
9: **end for**
10: **for each** $pass \in Layers$ **do**
11: **for each** P1 $\in$ Senders & P2 $\in$ Receivers **do**
12: **if** P1 = P2 **then**
13: $Weight[p1][p2] \Leftarrow 1$
14: **end if**
15: $Weight[p1[pass]][p2[pass]] \Leftarrow PASS\_NUM(P1, P2, pass)$
16: **end for**
17: **end for**

In the past, researchers used a weighted directed adjacency matrix in single-layer network analysis of team performance [8]. Formally, a weighted directed adjacency matrix W_j^i is used to display the edge's weight from each node i to each node j. However, in multilayer network analysis we use a weighted directed adjacency tensor. A tensor is a geometric element defined by indexes that indicate its rank. Its superscript index shows a canonical contravariant vector, and its subscript index shows a canonical covariant vector [32]. Formally, we use a weighted directed adjacency tensor $M_{j\beta}^{i\alpha}$ to display each node's edge weight i in each layer α to each node j in each layer β.

In the following steps, we use the CMPN model for microscale, mesoscale, and macroscale analyses of soccer teams. We use Python and R programming languages to implement these steps.

3.1. Microscale analysis

In soccer matches, identifying key players who are important for implementing team tactics and strategies is vital. In our study, we term these players as "core players" within their respective teams. To identify these core players in the teams' CMPN, we utilize the Pagerank versatility measure, an extension of single-layer PageRank centrality known for effectively capturing the most central node in a multilayer network (De Domenico et al. 2015).

The Pagerank versatility measure is particularly suitable for modeling ball-passing dynamics in soccer. As players have the ability to make different types of passes to their teammates, the Pagerank versatility measure allows us to account for this versatility within the proposed CMPN. To calculate the Pagerank versatility, we incorporate a transition tensor that enables a random walker to either jump to a neighboring node or teleport to any other node in the network. Teleportation is possible across multiple layers, with the rate varying depending on each layer. For the sake of simplicity, we assume a uniform teleportation rate across all layers in our CMPN.

The rank-4 transition tensor $R_{j\beta}^{i\alpha}$ for a random walker that jumps to its neighbors at rate r and teleports to other nodes in the multilayer network at rate $1 - r$ is calculated using Eq. (2). In this equation, $T_{j\beta}^{i\alpha}$

denotes the tensor of transition probabilities for jumping between pairs of nodes and switching between pairs of layers, N is the number of nodes, L is the number of layers, and $U^{i\alpha}_{j\beta}$ is a rank-4 tensor, in which all of its components have a value equal to 1. We set the r value to 0.85 as in the classical PageRank algorithm to match the literature [10].

$$R^{i\alpha}_{j\beta} = rT^{i\alpha}_{j\beta} + \frac{(1-r)}{NL}U^{i\alpha}_{j\beta} \tag{2}$$

The steady-state solution of the master equation for transition tensor $R^{i\alpha}_{j\beta}$ calculates the PageRank centrality of nodes in the multilayer network [33]. The eigentensor of the transition tensor $R^{i\alpha}_{j\beta}$, denoted by $\Omega_{i\alpha}$, indicates the probability of the steady-state solution for finding a random walker in node i and layer α [34]. As a result, the PageRank centrality for node i and layer α is calculated according to Eq. (3). In this equation, λ_1 is the largest eigenvalue.

$$R^{i\alpha}_{j\beta}\Omega_{i\alpha} = \lambda_1 \Omega_{j\beta} \tag{3}$$

The determination of the importance of pass types involves subjective judgments that can be effectively addressed through pair-wise comparisons. In this study, we utilize the analytic hierarchy process (AHP) to determine the weights of CMPN layers. AHP is a measurement theory that allows the evaluation of intangibles in relative terms, particularly useful when decision-making involves personal involvement and preferences [35]. In our study, we implement AHP using Saaty's fundamental scale, which assigns an absolute scale from 1 (equal importance) to 9 (extreme importance) to assess the intensity of importance [36]. We then employ tensor contraction to calculate player versatility based on the obtained weights [32]. The player with the highest versatility is referred to as the core player within the team.

3.2. Mesoscale analysis

Multilayer motif analysis can reveal new insights which could not be obtained by studying single-layer motifs. We investigate motifs with three nodes in the proposed CMPN to discover patterns leveraging various tactical passes by trios of teams, which are missing in the aggregated single-layer network [37].

We use MuxViz, an open-source software containing a collection of algorithms for analyzing multilayer networks [38], to detect multilayer motifs of the CMPN. MuxViz incorporates FANMOD, a tool that implements the RAND-ESU algorithm for network motif detection [39]. We set the FANMOD parameters as follows: motif size = 3, number of samples for approximated calculation = 1000, null model = local const, number of random networks to use for assessing significance = 100, Exchanges per edge = 3, and exchange attempts = 3.

3.3. Macroscale analysis

At this level, our investigation focuses on two main objectives: analyzing the similarity between passing tactics employed by soccer teams and predicting the outcomes of attacking plays in soccer competitions.

3.3.1. Similar tactics

To analyze team plays effectively, we employ a method that evaluates the similarity between distinct tactics utilized by each team during a game. By identifying the opposing team's players involved in different tactics, coaches can take appropriate actions to neutralize several strategies employed by the opposition. In the proposed CMPN model of a team, each layer represents a tactical relationship between the team players. We determine similar tactics by calculating the correlation of the inter-layer assortativity for each pair of the CMPN layers. We define assortativity type based on the out-degree of every node, in other words, the passes made by each player.

There are various similarity measures for comparing individual structures of layers in multilayer networks. We use the Pearson inter-layer assortativity measure to calculate the similarity between the types of passes in the CMPN model. For this purpose, we calculate the Person correlation coefficient between the number of passes made by players and their replicas in other layers for the whole types of passes of the proposed CMPN. Eq. (4) presents the Person correlation coefficient in which $cov(x, y)$ is the covariance of degree value for two nodes x and y. s_x and s_y show the standard deviation of these nodes [29].

$$r = \frac{cov(x, y)}{s_x s_y} \tag{4}$$

3.3.2. Outcome prediction

The outcome prediction of attacking plays holds significant importance for coaches and soccer analysts as it allows them to enhance the organization of their teams. We propose the Attacking plays Detector (APD) algorithm that leverages the CMPN model to distinguish the distinct passing trajectories that begin from a ball recovery and end either with a shot or goal. These trajectories incorporate the inherent multilayer nature of pass sequences in soccer matches. In the APD algorithm, we required spatio-temporal log events and the distinguished types of passes which construct the layers of the CMPN model. The APD pseudocode is presented in Algorithm 2

Attacking plays begin with a ball recovery or kick-off, which are set to "B" parameter events (Line 2), and end with a shot or goal, which are set to "E" parameter events (Line 3). These parameters also can be adjusted to include additional types of events. We store the number of events in "z" (Line 4). The parameter "t" store the time steps of events as the events are temporal (Line 5). We store the maximum length of attacking plays in MaxLen parameter (Line 6).

During match events (Line 7), when an event from the "B" set occurs (Line 8), we start a new trajectory array until we reach an "E" event (Line 11) and append events from the types of passes of the CMPN (Line 12) to the trajectory (Line 13). If we reach a "B" event during these steps, we put away the previous trajectory and start a new trajectory because we want trajectories from "B" to "E" sets (Line 14). If the last event of the trajectory is a goal, we label it as "1", and if the last event is shot, we label it as "0" (Lines 19–22). We update the MaxLen variable if the length of the detected trajectory is greater than its value (Lines 24–25). Finally, As the trajectory from "B" to "E" is completed, we append that to the Attacking plays array (Lines 27).

Now, we present a novel method for outcome prediction of attacking plays in soccer matches using multilayer networks and deep learning concepts. LSTM networks comprehend data sequences and can deal with the vanishing gradient problem in recurrent neural networks. Therefore, we use the LSTM network to discover which attacking plays are more likely to end up with goals at different competition levels, such as the World Cup, European Leagues, and teams.

The pseudocode of the OPAP method for predictability analysis of competitions is presented in Algorithm 3. First, we split the derived attacking plays dataset from Algorithm 2 into train and test data sets. We use stratified sampling to randomly selects samples from each class in proportion to their representation in the dataset and dedicate 20 percent of the data for testing and 80 percent for training the LSTM model (Line 1). The dataset exhibits an imbalance, with a significantly lower number of goals compared to shots in soccer matches. To address this issue, we adopt a standard approach from the existing literature, specifically designed to handle imbalanced datasets present in our study [40]. We apply undersampling exclusively to the training set, while the test set remains unaltered. In other words, no undersampling is applied to the test set. Consequently, our model is trained solely on the undersampled training set.

Pass-type features are encoded using the One-Hot encoder for preparing data to input the model (Line 2). After that, we used these encoded layers to transform the dataset of attacking plays into One-Hot labeled sequences (Line 3). The LSTM model needs an equal size of sequences for training. Therefore, the sequences must be padded to the maximum length of attacking plays to have the same size to input into the model. We select post-padding at this step to use the CUDA Deep

Algorithm 2 Attacking Plays Detector (APD)

Require: Spatio-temporal log events & Layers from CMPN (Algorithm 1)
Ensure: Attacking Plays
1: Events ⇐ Spatio-temporal log events
2: B ⇐ Set beginning events of trajectories
3: E ⇐ Set goals and shots events as ends of trajectories
4: $z \Leftarrow LENGHT(Events)$
5: $t \Leftarrow 0$
6: $MaxLen \Leftarrow 0$
7: **while** $t \neq z$ **do**
8:   **if** $Events[t] \in B$ **then**
9:     new $Trajectory$
10:     $t \Leftarrow t + 1$
11:     **while** $Events[t] \notin E$ **do**
12:       **if** $Events[t] \in Layers$ **then**
13:         $Trajectory.append(Events[t])$
14:       **else if** $Events[t] \in B$ **then**
15:         new $Trajectory$
16:       **end if**
17:       $t \Leftarrow t + 1$
18:     **end while**
19:     **if** $Events[t] = Goal$ **then**
20:       $Trajectory.Label \Leftarrow 1$
21:     **else if** $Events[t] = Shot$ **then**
22:       $Trajectory.Label \Leftarrow 0$
23:     **end if**
24:     **if** $LENGTH(Trajectory) > MaxLen$ **then**
25:       $MaxLen \Leftarrow LENGTH(Trajectory)$
26:     **end if**
27:     $AttackingPlays.append(Trajectory)$
28:   **end if**
29:   $t \Leftarrow t + 1$
30: **end while**

Neural Network (cuDNN) implementation in the Keras deep learning framework [41] (Line 4).

We split the TrainData into ImbalancedData and ValdataionData sets. We again use stratified sampling to randomly selects samples from each class in proportion to their representation in the dataset and dedicate 20 percent of the train data for validating and 80 percent for training the LSTM model (Line 5). The ValidationData has the same distribution as TestData, and we use it to tune the hyperparameters of the OPAP model. The training procedure does not used the ValidationData and evaluation metrics are resulted on ValidationData at the end of each epoch. After that, we use the random undersampling method on the ImbalancedData to select all attacking plays with goal outcomes and a subset of attacking plays with shot outcomes for training the proposed deep learning model (Line 6). We select random undersampling due to its simplicity and relatively low execution time [42].

Next, we build the deep learning outcome prediction model. We set the number of steps to MaxLen and the number of features to distinct types of passes in the CMPN model (Lines 7–8). We use a sequential model consisting of an LSTM layer with units as same as the number of features followed by a dense layer with one unit and the sigmoid activation function to guarantee that the output of this unit will always be between 0 and 1 (Line 9).

To compile the model, we use the Adam optimizer and set its parameters as learning_rate = 0.001, beta_1 = 0.9, beta_2 = 0.999, and define binary cross-entropy as the loss measure (Line 10). The model is trained using batch sizes of 8 samples over 100 epochs. Throughout the training process, we monitored and recorded accuracy and loss values

Table 1
Actors in the Juventus team CMPN for the benchmark game.

Actor number	Wyscout-ID	Label
1	3318	S.Khedira
2	3323	G.Higuain
3	20395	M.De-Sciglio
4	20443	M.Pjanic
5	20455	G.Buffon
6	20461	G.Chiellini
7	20579	M.Benatia
8	20588	K.Asamoah
9	25437	B.Matuidi
10	89186	P.Dybala
11	105334	Douglas-Costa

for both the training and validation datasets (Lines 11–12). Finally, we present the evaluation metrics for the test dataset (Line 13).

Algorithm 3 Outcome Prediction of Attacking Plays (OPAP)

Require: Layers (Algorithm 1) & Attacking Plays, MaxLen (Algorithm 2)
Ensure: The LSTM accuracy in outcome prediction
1: $TrainData$ & $TestData \Leftarrow SPLIT(AttackingPlays)$
2: $EndocdedLayers \Leftarrow OneHotEncoder(Layers)$
3: $EndocdedDataset \Leftarrow Transform(TrainData, EndocdedLayers)$
4: $PaddedData \Leftarrow Padding(EndocdedDataset, MaxLen)$
5: $ImbalancedData$ & $ValidationData \Leftarrow SPLIT(PaddedData)$
6: $Dataset \Leftarrow Undersampling(ImbalancedData)$
7: $Steps \Leftarrow MaxLen$
8: $Features \Leftarrow LENGTH(Layers)$
9: $OPAP\_Model \Leftarrow ADD(LSTM(Steps, Features), Dense(sigmoid))$
10: $OPAP\_Model \Leftarrow COMPILE(Adam, binary\_crossentropy)$
11: $Accuracy$ & $Loss \Leftarrow OPAP\_Model(Dataset)$
12: $Val\_accuracy$ & $Val\_Loss \Leftarrow OPAP\_Model(ValidationData)$
13: $Evaluation\_Metrics \Leftarrow OPAP\_Eval(TestData)$

4. Results

We tested and evaluated the proposed CMPN using multiple soccer matches from the Wyscout dataset. Fig. 1 displays the CMPN model for the match between Juventus and Napoli, which took place on December 1, 2017. This match is considered a benchmark in the reference dataset paper by Pappalardo et al. [30]. Hence, we selected this game to demonstrate the results of microscale, mesoscale, and macroscale analyses conducted on the CMPN.

In the CMPN, players are considered as the actors, while types of passes represent the different layers. In Fig. 1, layer numbers are below the types of passes. Within each passing layer, the intra-layer edges represent the passes made and received between players. Inter-layer edges with a weight equal to 1 exist between replica players, but we do not draw them in this figure to have a clear view of the CMPN. Table 1 shows the actor number, Wyscout-ID, and label for each player in the starting lineup of the Juventus team in the benchmark game.

Fig. 2 illustrates all passes between players as layers are ignored. In other words, it shows when the layers of the CMPN (see Fig. 1) are aggregated into a single-layer network. Due to the loss of information when aggregating different types of passes between players in the aggregated single-layer network, it is non-trivial to determine core players, discover multilayer ball-passing motifs, distinguish similarities between different team ball-passing tactics, and predict the outcome of attacking plays via types of passes. In the next steps, we demonstrate the benefits of the CMPN over an aggregated single-layer network in detail.

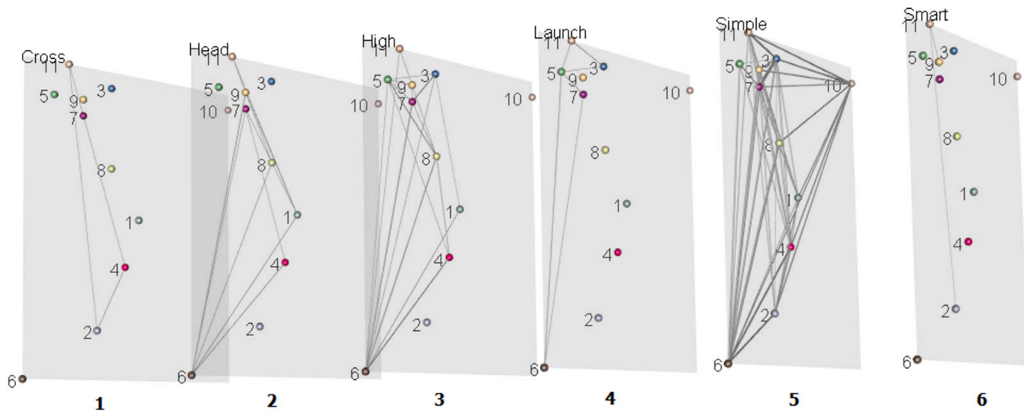


Fig. 1. CMPN of the Juventus team in the benchmark game. Each layer in the CMPN corresponds to a specific type of pass, with the layer numbers indicating the different pass types.

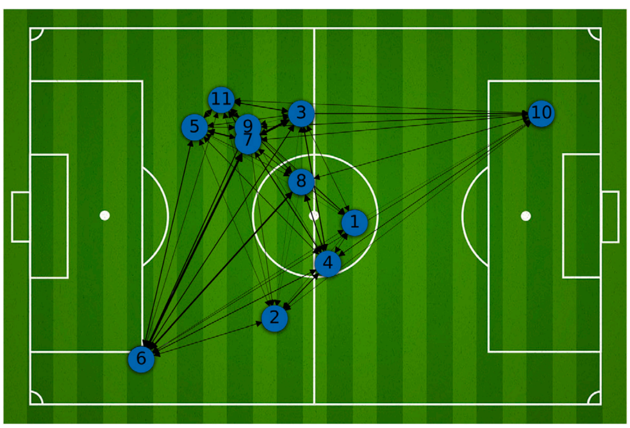


Fig. 2. Aggregated single-layer passing network of the Juventus team in the benchmark game.

Table 2
CMPN Statistics of the Juventus team in the benchmark game.

Layer	Nodes	Edges
Cross	3	3
Head	7	12
High	9	23
Launch	5	8
Simple	11	81
Smart	3	3

Table 3
Pairwise comparisons of CMPN layers.

Layer	Cross	Head	High	Launch	Simple	Smart
Cross	1	7	5	9	7	0.11
Head	0.14	1	0.33	9	3	0.11
High	0.2	3	1	9	5	0.11
Launch	0.11	0.11	0.11	1	0.11	0.11
Simple	0.14	0.33	0.2	9	1	0.11
Smart	9	9	9	9	9	1

Table 4
Importance weight of CMPN layers.

Layer	Weight
Cross	0.21
Head	0.08
High	0.11
Launch	0.02
Simple	0.08
Smart	0.5

4.1. Microscale analysis

Fig. 1 demonstrates that certain players, like Matudi, play key roles in implementing diverse tactics due to their ability to make various types of passes. However, the versatility of these players in making passes within a single-layer network cannot be captured because the passing information is lost when aggregating different passing tactics together.

Table 2 shows some of the statistics for the CMPN of the Juventus team in the benchmark game. The first column is the number of non-isolated nodes for every layer. All the players are present in the Simple pass layer. However, their presence in other layers is less. The second column is the number of directed edges disregarding their weights in each layer. Smart and Cross passes play a crucial role in creating significant scoring opportunities for teams. However, their lower occurrence in soccer matches can be attributed to the inherent difficulty involved in making these passes.

To validate the CMPN model from a sports science perspective, we consulted with the subject matter experts of a famous Iranian soccer analysis television program focused on soccer matches played in the Persian Gulf Pro League. The TV program had a soccer analysis department with eight experts to analyze soccer matches played in the Persian Gulf Pro League every week. They were familiar with the AHP method and provided us with the final pairwise comparisons of the CMPN layers in Table 3. After that, we calculate the importance weight of CMPN layers via the AHP method in Table 4.

Fig. 3 shows the PageRank results for the Juventus players in the benchmark match. Ranking players based on multilayer PageRank values is essential to distinguish the core players of teams. Therefore, we calculated the PageRank versatility of Juventus players in the proposed CMPN versus the PageRank centrality in the aggregated single-layer network for some Juventus matches. The PageRank values indicate that Matuidi has higher rankings than Chiellini in the proposed CMPN for Juventus matches. However, this ranking contradicts the aggregated network due to the information loss that is caused by projecting it onto a single-layer network. This discrepancy arises from the fact that player Chiellini primarily performs Simple passes. In contrast, player Matuidi is actively involved in crucial layers, such as the Smart pass layer. As a result, player Matuidi emerges as the core player within his team.

The calculation of centrality measures in the proposed CMPN is more reliable than calculating over the aggregated network or calculating centrality for each layer and then adding them with a selective heuristic. So far, researchers for determining key players of teams via

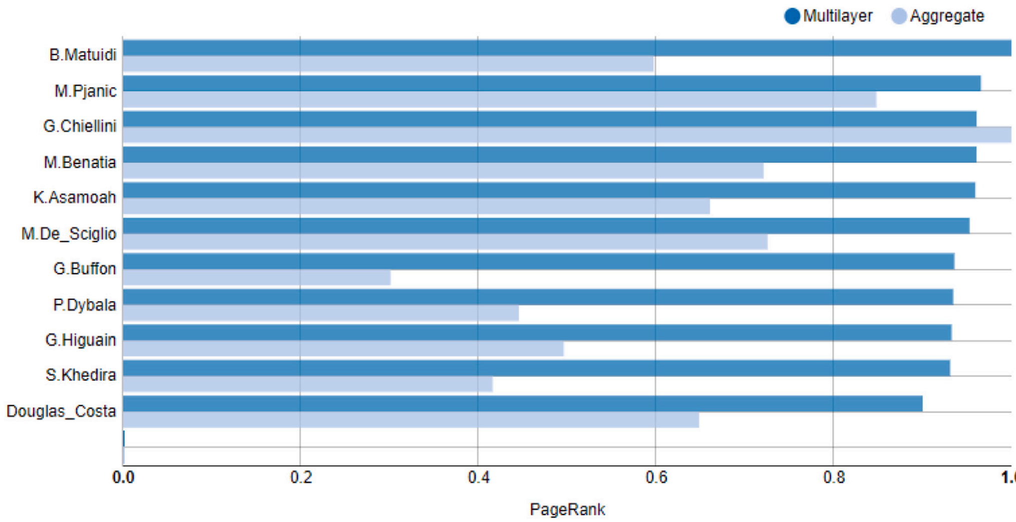


Fig. 3. Multilayer vs Aggregated PageRank comparison of Juventus players.

Table 5
Motifs inside the Juventus team CMPN.

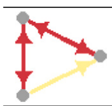
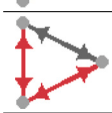
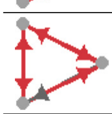
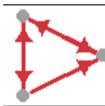
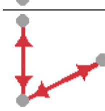
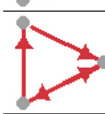
Motif	Adj.Matrix	Frequency
	$\begin{pmatrix} 0 & 5 & 0 \\ 5 & 0 & 5 \\ 3 & 5 & 0 \end{pmatrix}$	0.038
	$\begin{pmatrix} 0 & 6 & 5 \\ 6 & 0 & 5 \\ 5 & 5 & 0 \end{pmatrix}$	0.025
	$\begin{pmatrix} 0 & 5 & 6 \\ 5 & 0 & 5 \\ 5 & 5 & 0 \end{pmatrix}$	0.025

Table 6
Motifs inside the Barcelona team CMPN.

Motif	Adj.Matrix	Frequency
	$\begin{pmatrix} 0 & 5 & 0 \\ 5 & 0 & 5 \\ 5 & 5 & 0 \end{pmatrix}$	0.203
	$\begin{pmatrix} 0 & 0 & 5 \\ 0 & 0 & 5 \\ 5 & 5 & 0 \end{pmatrix}$	0.127
	$\begin{pmatrix} 0 & 0 & 5 \\ 5 & 0 & 0 \\ 5 & 5 & 0 \end{pmatrix}$	0.044

their passing networks did not consider the ability of a player to make different types of passes. For simplification, they assumed the same importance level for all types of passes in calculating the centrality of players. However, players who can make different types of passes play a critical role in assisting the team in accomplishing their tactical objectives.

4.2. Mesocale analysis

Coaches can discover the building blocks of opposing teams by analyzing their multilayer motifs. In Table 5, we present the most frequently occurring multilayer motifs of the Juventus team in the benchmark game, sorted by their frequency. The layer numbers in the adjacency matrix correspond to the CMPN model layers. Additionally, in Table 6, we showcase the most frequent motifs discovered in the CMPN of the Barcelona team during their match against Real Madrid on December 23, 2017. These multilayer motifs exhibit p-values lower than 0.001, indicating their statistical significance [43].

Based on the findings presented in these tables, we conclude that high passes play a central role in the primary tactical patterns executed by the trios of the Juventus team. In contrast, Barcelona follows a distinctive style known as Tiki-taka, renowned in Spanish soccer for its emphasis on short and simple passing.

4.3. Macroscale analysis

In this section, we begin by identifying the similar passing tactics utilized by the Juventus team in the benchmark game. Following that, we present a comprehensive analysis of various evaluation metrics of the OPAP method in selected competitions.

Table 7
Pearson inter-layer assortativity for Juventus CMPN.

Layer	Cross	Head	High	Launch	Simple	Smart
Cross	1	−0.261	−0.463	−0.251	0.099	**0.722**
Head		1	0.247	−0.259	0.324	−0.188
High			1	0.607	0.19	−0.437
Launch				1	−0.399	−0.028
Simple					1	−0.165
Smart						1

4.3.1. Similar tactics

In Table 7, we present the Pearson inter-layer assortativity measure values for the CMPN layers of the Juventus team. The results highlight that the Cross and Smart pass layers demonstrate the highest similarity, with a Pearson criterion value of 0.722. Consequently, we can conclude that there are shared players involved in these two distinct passing tactics. Opposing teams could think of a way to control these players, which can help them in increasing their chance of winning against the Juventus team.

4.3.2. Outcome prediction

Due to the imbalanced nature of our datasets, we conducted a comprehensive analysis of the OPAP method using diverse evaluation metrics. This analysis aimed to provide deeper insights into the performance of the method. A loss (cost) function accounts for the probabilities or uncertainty of a prediction based on how much its predictions

Table 8
Competition levels.

Competition	World Cup	Italy league	Juventus
Matches	64	380	38
Passes	58081	346771	21861
Attacking plays	903	5973	335
Goals	80	626	59

Table 9
Evaluation metrics for test data.

Competition	World Cup	Italy league	Juventus
Accuracy	0.95	0.95	0.96
Precision	0.85	0.81	0.83
Recall	0.6	0.6	0.71
F-measure	0.7	0.69	0.76

vary from the true value. Accuracy is the count of predictions where the predicted values are equal to the actual values. Accuracy, Precision, Recall and F-measure values are calculated using the number of true positives (TP), true negatives (TN), false positives (FP), and false negatives (FN) according to Eqs. (5), (6), (7), and (8), respectively [44].

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{5}$$

$$Precision = \frac{TP}{TP + FP} \tag{6}$$

$$Recall = \frac{TP}{TP + FN} \tag{7}$$

$$F - measure = \frac{2 \times Precision \times Recall}{Precision + Recall} \tag{8}$$

To demonstrate the results of the OPAP method at various competition levels, we have chosen to analyze matches from the World Cup, the Italy league, and specifically, all the matches played by the Juventus team in the Italy league. By considering these different competition levels, we aim to showcase the effectiveness and applicability of the OPAP method across various soccer tournaments and teams. Table 8 enumerates the number of matches, passes, attacking plays, and goals for these competitions. According to this table, the ratio of goals to total attacking plays is 9%, 10%, and 17% for the World Cup, the Italy league, and the Juventus team, respectively. The Juventus team was the champions of the 2017–18 Serie A. As a result, it has a higher ratio value than other general competition levels that include all teams.

The accuracy and loss values during the training process of the OPAP method for predicting the outcome of attacking plays in the competition levels are shown in Fig. 4. This figure proves that the loss value is decreasing and the accuracy value is increasing during training epochs of the OPAP method. The OPAP method achieves an accuracy rate of over 90% and a loss rate below 30%, providing compelling evidence of its ability to accurately predict the outcome of attacking plays based on the distinct playing styles observed in various competitions.[3] The plots in leagues such as the Italy league appear smoother due to the larger number of games in these competitions. With more data available for training the OPAP method, the results tend to be more consistent. Conversely, the plots for the World Cup and individual teams exhibit some variation due to the lower number of matches in these scenarios.

Table 9 shows the Accuracy, Precision, Recall, and F-measure values of the test data for each competition level. The Accuracy values indicate that the OPAP method predicts the outcomes well. As a result, the OPAP method can perfectly distinguish goal outcome structures of attacking plays in each competition level by incorporating the CMPN model.

Our results indicate that coaches and soccer analysts can benefit from the OPAP method to discover the best possible attacking plays for scoring goals and plan to defend against the attacking strategies of opposing teams.

5. Discussion

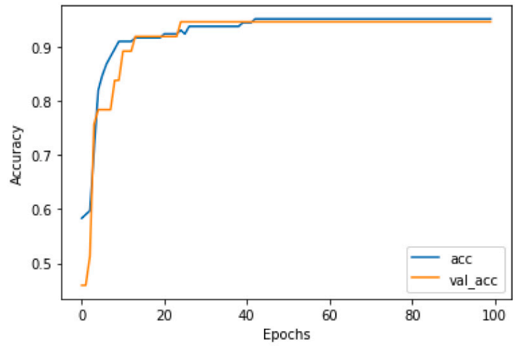
To compare the OPAP method with other existing approaches, we implemented a random forests (RF) model for predicting the outcomes of attacking plays in the World Cup competition, achieving a mean accuracy of 80%. The RF algorithm is an ensemble method known for its black-box nature, combining multiple independent decision trees for classification and regression tasks. Additionally, it proves useful for feature selection tasks [45].

Fig. 5 illustrates the SHAP (SHapley Additive exPlanations) values for both the OPAP method and the RF outcome prediction model. These values help identify important features for predicting the outcomes of attacking plays. SHAP employs a game theoretic approach to explain the outputs of powerful, albeit black-box, artificial intelligence models [45]. Fig. 5(a) demonstrates the sequence evaluation of the CMPN model in the OPAP method, in which large consecutive sequences of the Simple pass do not necessarily end up with goals. The RF algorithm is useful for understanding the importance of the number of occurrences for each type of pass in trajectories ending with goals. Fig. 5(b) illustrates the occurrence evaluation of the CMPN, in which large values of the Simple pass in an attacking play have a low impact on the outcome prediction, which means that although the most frequent pass in soccer matches is the Simple pass, it has a low value for achieving goals. On the other hand, the large values of the Cross pass have a high effect on attacking plays ending with goals in this competition. In addition, the Launch pass has no impact on the model output.

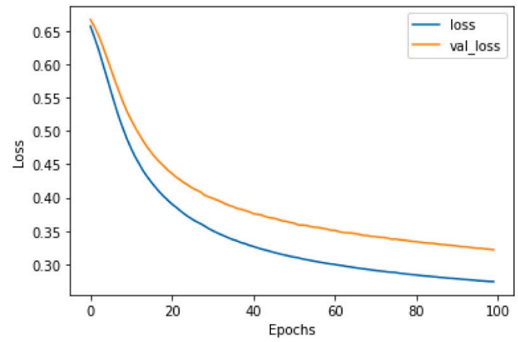
We calculate the mean percentage value for each CMPN layer to find the structures of attacking plays ending with goals for each competition. Our findings show that the lowest and highest values belong to the Launch and Simple passes. According to the SHAP value of these passes, they have little impact on determining the outcome of attacking plays. Fig. 6 shows the values for the most important CMPN layers in the World Cup, Italy league, and Juventus team. The Cross pass has the highest percentage value in the World Cup, which means the main plan for scoring a goal in that competition was to make this type of pass. One reason for this decision could be that team players in the World Cup have lower time compared to league competitions to match their tactics, and coaches decide to use these long passes from the sides of the pitch in their attacking play strategies as the easiest way to score a goal. On the other hand, the Smart pass has the highest value in the Italy league, which justifies coaches in Serie A having more time for planning tactical strategies to design attacking plays.

The implementation of the OPAP method on the Wysout dataset yields promising results, demonstrating the effectiveness of the CMPN model in predicting the outcomes of attacking plays across various competition levels. These findings prove valuable for coaches in adjusting strategies, training teams, and analyzing the tactics of opponent teams. Our model demonstrates a high level of trustworthiness, achieving an accuracy rate of over 90% and an F-measure of approximately 70% in predicting the outcome of attacking plays across all investigated competition levels. This information is very beneficial for coaches and soccer analysts to discover the best pass sequences based on the characteristics of each competition to train their teams and plan for increasing the chance of scoring goals in attacking plays.

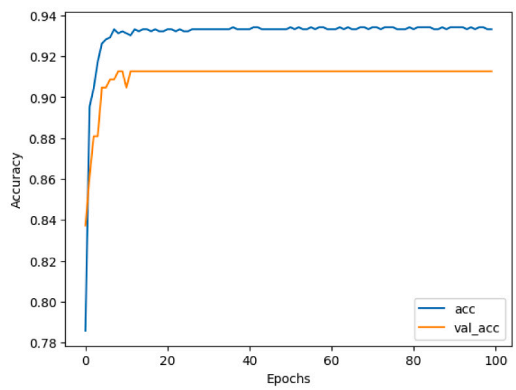
[3] The attacking plays datasets are available on “https://github.com/Beheshtian/CMPN”.



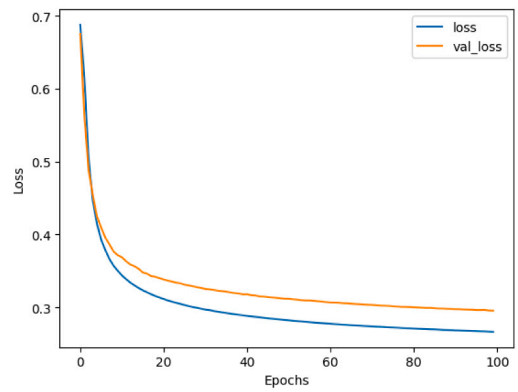
(a) World Cup Accuracy



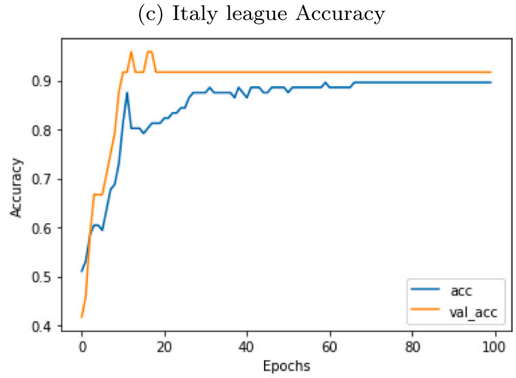
(b) World Cup Loss



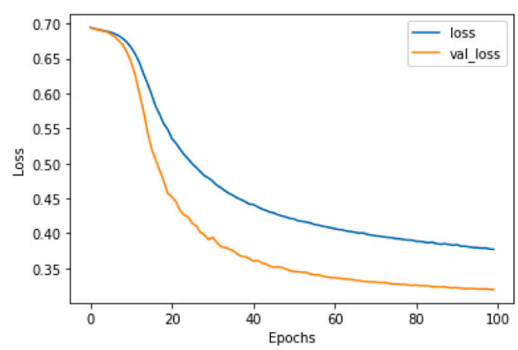
(c) Italy league Accuracy



(d) Italy league Loss

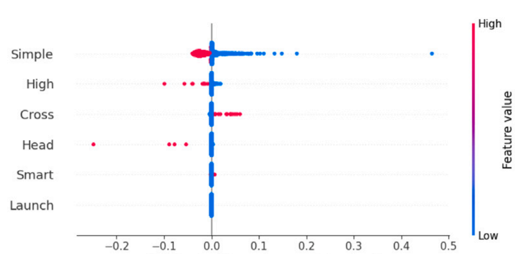


(e) Juventus team Accuracy

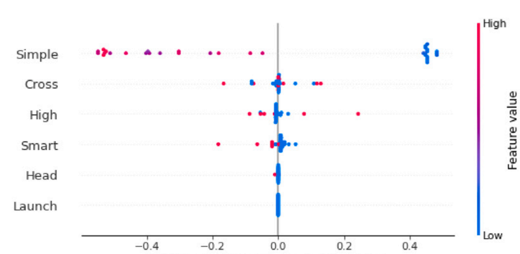


(f) Juventus team Loss

Fig. 4. The Accuracy and Loss values of the OPAP method for selected competitions. The blue lines represent the values for the training dataset, while the orange lines depict the values for the validation dataset. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)



(a) SHAP for OPAP



(b) SHAP for RF

Fig. 5. SHAP values of CMPN layers for goal trajectories in World Cup. (a) shows sequence evaluation via the OPAP method. (b) shows occurrence evaluation via the RF algorithm.

Fig. 6. Mean percentage of CMPN layers in goal trajectories for competition levels.

6. Conclusion

This paper proposed a novel model for performance analysis of soccer teams using the Complex Multiplex Passing Network (CMPN), focusing on soccer matches. In contrast with the traditional aggregated single-layer passing network, the proposed CMPN model leverage distinct types of passes between teammates in its performance analysis.

We demonstrated the advantages of the proposed CMPN in analyzing team sports performance across various topological scales, utilizing a real-world dataset. Our study comprehensively investigated the microscale, mesoscale, and macroscale levels, providing valuable insights at each level. At the microscale level, we effectively identified core players within the teams. Moving to the mesoscale level, we successfully discovered patterns of multilayer passing motifs that contribute to overall team performance. Lastly, at the macroscale level, we identified similar passing tactics and proposed a method for predicting the outcome of attacking plays. Remarkably, our predictive model achieved an accuracy rate of over 90% for outcome prediction, with an approximate F-measure of 70%. These findings highlight the effectiveness and practicality of our approach in enhancing team sports analysis and performance evaluation. Coaches can leverage this valuable information to analyze opposing teams, determine effective attacking and defending strategies for their own teams, and maximize goal-scoring opportunities across multiple matches.

As a future direction of this research, we aim to study relations between the core players, motifs, and tactics of teams found in this paper. For example, we will investigate the cooperation between motifs in the proposed CMPN to understand how these building blocks are related. In addition, we plan to utilize machine learning algorithms to calculate the weights of the CMPN layers. Moreover, we aim to employ Graph Neural Networks to predict the roles of core players in attacking plays. The requirement of comprehensive datasets containing events of soccer matches is a potential limitation of implementing the proposed CMPN.

CRediT authorship contribution statement

Arash Beheshtian-Ardakani: Conceptualization, Methodology, Software, Writing – original draft, Visualization, Investigation. **Mostafa Salehi:** Methodology, Validation, Supervision. **Rajesh Sharma:** Conceptualization, Writing – reviewing.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

The datasets supporting the findings of this study are available in [30].

Acknowledgments

Mostafa Salehi is supported by a grant from IPM, Iran (No. CS1402-4-162). Rajesh Sharma is supported by EU H2020 program under the SoBigData++ project (No. 871042) and CHIST-ERA (No. CHIST-ERA-19-XAI-010).

References

[1] FIFA. FIFA to introduce enhanced football intelligence at FIFA world cup 2022™. 2022, URL https://www.fifa.com/technical/football-technology/media-releases/fifa-to-introduce-enhanced-football-intelligence-at-fifa-world-cup-2022-tm.
[2] 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(23):2631–8. http://dx.doi.org/10.1080/02640414.2019.1589919.
[3] Gong B, Zhou C, Gómez M-A, Buldú J. Identifiability of Chinese football teams: A complex networks approach. Chaos Solitons Fractals 2023;166:112922. http://dx.doi.org/10.1016/j.chaos.2022.112922.
[4] Buldú JM, Busquets J, Martínez JH, Herrera-Diestra JL, Echegoyen I, Galeano J, Luque J. Using network science to analyse football passing networks: Dynamics, space, time, and the multilayer nature of the game. Front Psychol 2018;9. http://dx.doi.org/10.3389/fpsyg.2018.01900.
[5] Ievoli R, Palazzo L, Ragozini G. On the use of passing network indicators to predict football outcomes. Knowl-Based Syst 2021;222:106997. http://dx.doi.org/10.1016/j.knosys.2021.106997.
[6] Ichinose G, Tsuchiya T, Watanabe S. Robustness of football passing networks against continuous node and link removals. Chaos Solitons Fractals 2021;147:110973. http://dx.doi.org/10.1016/j.chaos.2021.110973.
[7] Clemente FM, Sarmento H, Aquino R. Player position relationships with centrality in the passing network of world cup soccer teams: Win/loss match comparisons. Chaos Solitons Fractals 2020;133:109625. http://dx.doi.org/10.1016/j.chaos.2020.109625.
[8] Gudmundsson J, Horton M. Spatio-temporal analysis of team sports. ACM Comput Surv 2017;50(2). http://dx.doi.org/10.1145/3054132.
[9] Zhang P, Wang T, Yan J. PageRank centrality and algorithms for weighted, directed networks. Physica A 2022;586:126438. http://dx.doi.org/10.1016/j.physa.2021.126438.
[10] Peña JL, Touchette H. A network theory analysis of football strategies. In: Clanet C, editor. Sports Physics: proc. 2012 euromech physics of sports conference. Palaiseau: Éditions de l'École Polytechnique; 2012, p. 517–28, arXiv:1206.6904.
[11] Håland EM, Wiig AS, Hvattum LM, Stålhane M. Evaluating the effectiveness of different network flow motifs in association football. J Quant Anal Sports 2020;16(4):311–23. http://dx.doi.org/10.1515/jqas-2019-0097.
[12] Young CM, Luo W, Gastin P, Lai J, Dwyer DB. Understanding effective tactics in Australian football using network analysis. Int J Perform Anal Sport 2019;19(3):331–41. http://dx.doi.org/10.1080/24748668.2019.1605562.
[13] Herrera-Diestra JL, Echegoyen I, Martínez JH, Garrido D, Busquets J, Io FS, et al. Pitch networks reveal organizational and spatial patterns of Guardiola's F.C. Barcelona. Chaos Solitons Fractals 2020;138:109934. http://dx.doi.org/10.1016/j.chaos.2020.109934.
[14] Wu Y, Xia Z, Wu T, Yi Q, Yu R, Wang J. Characteristics and optimization of core local network: Big data analysis of football matches. Chaos Solitons Fractals 2020;138:110136. http://dx.doi.org/10.1016/j.chaos.2020.110136.
[15] Sarmento H, Clemente FM, Gonçalves E, Harper LD, Dias D, Figueiredo A. Analysis of the offensive process of AS monaco professional soccer team: A mixed-method approach. Chaos Solitons Fractals 2020;133:109676. http://dx.doi.org/10.1016/j.chaos.2020.109676.
[16] Stübinger J, Mangold B, Knoll J. Machine learning in football betting: Prediction of match results based on player characteristics. Appl Sci 2020;10(1). http://dx.doi.org/10.3390/app10010046.
[17] Berrar D, Lopes P, Dubitzky W. Incorporating domain knowledge in machine learning for soccer outcome prediction. Mach Learn 2019;108(1):97–126. http://dx.doi.org/10.1007/s10994-018-5747-8.
[18] Cho Y, Yoon J, Lee S. Using social network analysis and gradient boosting to develop a soccer win–lose prediction model. Eng Appl Artif Intell 2018;72:228–40. http://dx.doi.org/10.1016/j.engappai.2018.04.010.
[19] Medina P, Carrasco S, Rogan J, Montes F, Meisel JD, Lemoine P, et al. Is a social network approach relevant to football results? Chaos Solitons Fractals 2021;142:110369. http://dx.doi.org/10.1016/j.chaos.2020.110369.
[20] Rahman MA. A deep learning framework for football match prediction. SN Appl Sci 2020;2(2):165. http://dx.doi.org/10.1007/s42452-019-1821-5.

[21] Zhang Q, Zhang X, Hu H, Li C, Lin Y, Ma R. Sports match prediction model for training and exercise using attention-based LSTM network. Digit Commun Netw 2021. http://dx.doi.org/10.1016/j.dcan.2021.08.008.
[22] Mattsson CES, Takes FW. Trajectories through temporal networks. Appl Netw Sci 2021;6(1):35. http://dx.doi.org/10.1007/s41109-021-00374-7.
[23] Guan Y, Qiu Y, Tian C. Trajectory planning in college football training using deep learning and the internet of things. J Supercomput 2022. http://dx.doi.org/10.1007/s11227-022-04619-9.
[24] Zhao Y, Yang R, Chevalier G, Shah RC, Romijnders R. Applying deep bidirectional LSTM and mixture density network for basketball trajectory prediction. Optik 2018;158:266–72. http://dx.doi.org/10.1016/j.ijleo.2017.12.038.
[25] Caicedo-Parada S, Lago-Peñas C, Ortega-Toro E. Passing networks and tactical action in football: A systematic review. Int J Environ Res Public Health 2020;17(18). http://dx.doi.org/10.3390/ijerph17186649.
[26] Ramos J, Lopes RJ, Araújo D. What's next in complex networks? Capturing the concept of attacking play in invasive team sports. Sports Med 2018;48(1):17–28. http://dx.doi.org/10.1007/s40279-017-0786-z.
[27] Salehi M, Sharma R, Marzolla M, Magnani M, Siyari P, Montesi D. Spreading processes in multilayer networks. IEEE Trans Netw Sci Eng 2015;2(2):65–83. http://dx.doi.org/10.1109/TNSE.2015.2425961.
[28] Interdonato R, Magnani M, Perna D, Tagarelli A, Vega D. Multilayer network simplification: Approaches, models and methods. Comp Sci Rev 2020;36:100246. http://dx.doi.org/10.1016/j.cosrev.2020.100246.
[29] Bianconi G. Multilayer networks: structure and function. New York, NY, USA: Oxford University Press; 2018.
[30] Pappalardo L, Cintia P, Rossi A, Massucco E, Ferragina P, Pedreschi D, Giannotti F. A public data set of spatio-temporal match events in soccer competitions. Sci Data 2019;6(1):236. http://dx.doi.org/10.1038/s41597-019-0247-7.
[31] Wyscout. Wyscout main events description. 2018, URL https://footballdata.wyscout.com/wp-content/uploads/2018/03/Wyscout-Events-Manual.pdf.
[32] De Domenico M, Solé-Ribalta A, Cozzo E, Kivelä M, Moreno Y, Porter MA, et al. Mathematical formulation of multilayer networks. Phys Rev X 2013;3:041022. http://dx.doi.org/10.1103/PhysRevX.3.041022.
[33] Solé-Ribalta A, De Domenico M, Gómez S, Arenas A. Random walk centrality in interconnected multilayer networks. Physica D 2016;323–324:73–9. http://dx.doi.org/10.1016/j.physd.2016.01.002.
[34] De Domenico M, Solé-Ribalta A, Omodei E, Gómez S, Arenas A. Ranking in interconnected multilayer networks reveals versatile nodes. Nature Commun 2015;6(1):1–6. http://dx.doi.org/10.1038/ncomms7868.
[35] Munier N, Hontoria E. Uses and limitations of the AHP method. Springer International Publishing; 2021.
[36] Saaty RW. The analytic hierarchy process—what it is and how it is used. Math Model 1987;9(3):161–76. http://dx.doi.org/10.1016/0270-0255(87)90473-8.
[37] Sharma R, Magnani M, Montesi D. Missing data in multiplex networks: A preliminary study. In: 2014 Tenth international conference on signal-image technology and internet-based systems. 2014, p. 401–7. http://dx.doi.org/10.1109/SITIS.2014.65.
[38] De Domenico M, Porter MA, Arenas A. MuxViz: a tool for multilayer analysis and visualization of networks. J Complex Netw 2014;3(2):159–76. http://dx.doi.org/10.1093/comnet/cnu038.
[39] Wernicke S, Rasche F. FANMOD: a tool for fast network motif detection. Bioinformatics 2006;22(9):1152–3. http://dx.doi.org/10.1093/bioinformatics/btl038.
[40] Fernández A, García S, Galar M, Prati RC, Krawczyk B, Herrera F. Learning from imbalanced data sets, vol. 10. Springer; 2018.
[41] Chollet F, et al. Keras. 2015, URL https://keras.io/.
[42] Yen S-J, Lee Y-S. Under-sampling approaches for improving prediction of the minority class in an imbalanced dataset. In: Intelligent control and automation: International conference on intelligent computing. Springer; 2006, p. 731–40. http://dx.doi.org/10.1007/978-3-540-37256-1_89.
[43] Zhu W. $P< 0.05, < 0.01, < 0.001, < 0.0001, < 0.00001, < 0.000001$, or $< 0.0000001\ldots$ J. Sport Health Sci 2016;5(1):77–9. http://dx.doi.org/10.1016/j.jshs.2016.01.019.
[44] Ballı S, Özdemir E. A novel method for prediction of EuroLeague game results using hybrid feature extraction and machine learning techniques. Chaos Solitons Fractals 2021;150:111119. http://dx.doi.org/10.1016/j.chaos.2021.111119.
[45] Sun J, Yu H, Zhong G, Dong J, Zhang S, Yu H. Random Shapley forests: Cooperative game-based random forests with consistency. IEEE Trans Cybern 2022;52(1):205–14. http://dx.doi.org/10.1109/TCYB.2020.2972956.