

Research Article

Efficient Strategy Mining for Football Social Network

Taige Zhao,¹ Ningning Cui², Yunliang Chen³, and Man Li⁴

¹School of Computer Science, University of Sydney, Sydney NSW 2017, Australia

²School of Computer Science, Anhui University, Hefei 230601, China

³School of Computer Science, China University of Geosciences, Wuhan 430074, China

⁴School of Information Technology, Deakin University, Geelong VIC 3220, Australia

Correspondence should be addressed to Ningning Cui; willber1988@163.com and Yunliang Chen; 28564395@qq.com

Received 1 September 2020; Revised 9 September 2020; Accepted 13 September 2020; Published 24 September 2020

Academic Editor: Jianxin Li

Copyright © 2020 Taige Zhao et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

With the growing popularity of social network in sport, it expresses the social relationships between individuals and facilitates realistic applications, e.g., social event mining and discovery. Sport network as a specific social network has been widely studied in research and commercial fields. However, most of the existing works utilize a simplex strategy to improve certain indicators in the team and do not consider the effect of strategy adjustment based on the current situation. In this paper, we study the problem of efficient strategy mining in football social network. To address this problem, we propose a quantitative way to combine the aspects of coordination, adaptability, flexibility, and tempo into a passing network, which notably improves the timeliness and information content of the existing network. On this basis, we design a suppression function to express the impact of strategy. Then, we propose a novel passing network and group cooperation scheme based on quantified team performance to obtain the efficient strategies. At last, the experimental results show that, based on the performance of the same team, our optimized passing network has a higher winning rate in practice.

1. Introduction

With the rapid advance of social network in smart city, it expresses the social relationships between individuals and facilitates realistic applications, e.g., social event mining and discovery. Originally, the social network was proposed by Georg Simmel from a sociological point of view, which focuses on analysing the possibilities and limitations of how interpersonal relationships affect their actions [1]. Meanwhile, along with the popular sparked enthusiasm on sport domain, sport social network has proliferating development. Hence, sport social network as a specific social network has been widely studied in research and commercial fields. The sport network is based on social networks to analyse the passing relationship between players. As shown in Figure 1, in the classic passing network, each node represents the entity of each player, and the connection between nodes represents the passing relationship between players [2]. Based on the passing network, the coordination of subteams (special configuration) can be identified through the idea of cohesive subgroups in social networks [2].

In this paper, we study the problem of efficient strategy mining in football social network in smart city. There exist lots of works to improve team performance through strategies from different perspectives. Specifically, Florian Korte and others proposed a passing network based on the player's offensive position, which added the player's real-time position status factor [3]. Duch and Clemente considered the impact of different player configurations on the passing network [4, 5]. However, traditional strategies are fixed and fail to quantify the impact of opponent strategies on real-time team status. The effect after the strategy is formulated cannot be quantified, which results in a great possibility that the theoretical and practical effects are inconsistent. In addition, the effectiveness of the strategy is often affected by the passing habits of the players. Therefore, it is necessary to establish a passing network for making strategies to quantify the current state of players and teams.

To address the above problem, it is necessary to quantify and synthesize a large number of existing team attributes. Therefore, in this paper, based on the cohesive subgroup

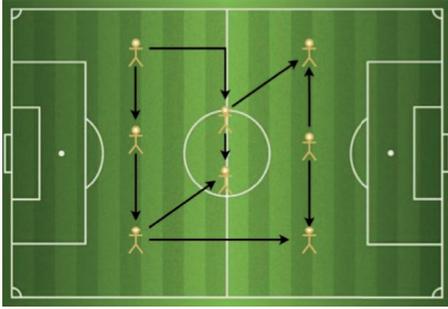


FIGURE 1: Example of the passing network.

analysis, we first establish a special configuration (binary and ternary configuration) recognition model. On the micro level, we use centrality degree and betweenness to quantify the performance of players. On the macrolevel, we define indicators for evaluating team performance from four aspects as follows: *coordination*: analyse the network density and distance; *adaptability*: define and calculate the ball control rate and continuous pass rate; *flexibility*: focus on the variance of delivery coordination time; and *tempo*: summarize the number of offense and defense transitions. After that, based on this idea and considering the opponent's strategy, we propose a suppression function to represent the suppression of players under different strategies and thus generate a new pass network. The new two-tuple and three-tuple structure obtained through the new passing network is used to formulate the optimized strategy. The optimized passing network analyses the real-time status of the game through macro- and microaspects and reflects the performance in the passing network to improve the real-time and comprehensive nature of the passing network.

The main contributions of this paper are as follows:

- (1) To the best of our knowledge, this is the first work that attempts to optimize the passing network to achieve efficient strategies through optimized structures
- (2) We propose a suppression function to quantify the suppression effect on players under different strategies
- (3) We propose a novel passing network based on the quantified team performance for formulating strategies
- (4) We conduct extensive experiments to demonstrate the effectiveness and efficiency of our proposed strategies in practice

The rest of this paper is organized as follows. Section 2 introduces the related work of the passing network. Section 3 provides the architecture of the entire model and gives a brief overview. Section 4 introduces the method proposed to establish an optimized passing network. Section 5 gives a comparison of the win rate under different passing networks and proves that the proposed team performance method is effective. Finally, the work is summarized in Section 6.

2. Related Work

2.1. Passing Network. The traditional passing network has three main forms: (i) player passing network, where the nodes are player entities [6]; (ii) pitch passing networks, where the passing node is a specific area on the football field [7]; and (iii) pitch passing networks, where the passing node is a combination of player and specific position [8].

Since then, more and more methods have tried to reflect the quantitative performance of individuals and teams in the passing network. Taking the player passing network as an example, the node size is used to represent the player's performance indicators, and the thickness of the connecting line is used to represent the degree of relationship between entities. The passing network obtained in this way will contain more information, making the entire network more representative of the current state [9–11].

Although the passing network contains a wealth of information, it does not perform well in predicting the future state after the strategy is changed. It only feeds back the current state and cannot effectively predict the future. Therefore, we have changed the order between the two here to obtain the direction of the strategy change from the change of the passing network instead of predicting the new passing network through the changed strategy.

2.2. Performance Analysis Based on Passing Network. Generally, performance analysis based on the passing network is considered from two perspectives: micro and macro. On the micro level, the passing network based on social network analysis (SNA) focuses on the performance of each player and measures individual performance by the number of passes by each player in the passing network [12, 13]. The macrolevel passing network focuses on the overall team performance level, which is particularly reflected in the team tactical analysis [14, 15]. References [16, 17] discussed maximizing the impact on the goal in social networks. Through the opponent's tactics and our own team, we can quantify the team's performance to determine the best tactics for the game [18].

However, the improvement of the correlation between performance and strategy proposed by micro and macro angles is a challenge that needs to be solved. On the micro level, when we consider the player's degree centrality (the number of direct connections between the central node and other nodes) on the basis of player passing and receiving, it often has better results [19]. In addition, sociometric status has also been proved to be a method of quantifying player's performance. It measures the minimum number of steps between the central player and other players as a supplement to the importance of players [18]. Reference [20] discussed the impact of real-time personnel gathering changes. On the macrolevel, with the use of coordination, adaptability, and flexibility as a quantitative indicator of team performance, rhythm is better than the average number of team passes based on SNA ideas [21, 22]. In addition, tactics are also affected by individual abilities, whether they win or lose. Due to the difference in basic tactical thinking between winning

and losing, the past network will also undergo great changes [23, 24]. References [25, 26] pointed out how to process passing network data in a fast way. Although these methods ensure the integrity of the information, they neglect the quantitative combination on the passing network.

Therefore, this paper proposes a passing network that combines macro- and microperformance and considers the influence of opponents with the help of a suppression function to obtain an improved passing network that can significantly improve the winning rate and to formulate strategies accordingly.

3. Scheme Overview

The goal of this work is to mine more efficient strategy such that the team can have a better performance. The overall process is illustrated in Figure 2. Specifically, for a given data set, a personal performance model is used to generate quantitative indicators about nodes for the delivery network. In addition, we propose four new indicators based on the team performance model to enrich the information in the network. Then, the generated rich information passing network model quantifies the impact of the opponent's strategy through the proposed suppression function to generate an optimized pass network model. After that, we generate a special configuration of the dual and triple structure through the optimized passing network. Finally, the special preparation suggestions are sent to the strategy makers to develop dynamic strategies.

4. Establishment of Optimized Passing Network

4.1. Infrastructure of Passing Network

4.1.1. The Elements of Passing Network. The passing network consists of players in the team, where each node represents the entity of each player. Technically, a passing network is essentially topology, and we use transit networks as a tool. In detail, the size of the node reflects the importance of the player in this game; the connection between the nodes represents the passing relationship between the players; and the thickness of the connection indicates the frequency of passing between players.

4.1.2. The Weight of Nodes and Links. As mentioned above, the size of the nodes reflects the importance of the players. We can know that the calculating process (that is, the radius of the vertex in the passing network) is similar to Page-Rank algorithm used for social network. They both map the importance of nodes in the graph to a specific number. The Page-Rank centrality introduced here can also be regarded as a recursive concept of popularity or importance, which follows the principle of a player. Here, we propose the algorithm Player-Rank, which calculates the importance of players in the passing network. Its calculation formula is shown as follows:

$$PR(P) = \beta + \alpha \sum_{j \neq i} \frac{A_{ij}}{L_j^{\text{out}}} X_j, \quad (1)$$

where $PR(P)$ indicates the importance function of the player in the team; $L_j^{\text{out}} = \sum_k A_{jk}$ shows the total number of passes made by player j ; and α is a random weight variable, which indicates the probability that a player decides to give the ball to himself instead of keeping it and continuing to shoot. The value range is $[0, 1]$, which is usually taken by default in Page-Rank, and β is the parameter awarded to "free" popularity by every player. It is worth noting that the player's Page-Rank score also depends on the scores of all teammates. Therefore, all Page-rank scores on the team must be calculated simultaneously.

Page-Rank's centrality roughly assigns each player the possibility of getting the ball after a reasonable number of passes. If this measurement requires higher accuracy, then probability α can be replaced by the player-related probability value, which will make more sense if some players are more inclined to hold the ball than others. In either case, the value of α is not only from the network, because the values between a team may be different usually and should be determined by heuristics. As a proof of concept, in our analysis, we will use uniform values $\alpha = 0.85$ and $\beta = 1$ for all the teams.

4.2. The Establishment of Basic Passing Network

4.2.1. The Basic Process of Modelling. Figure 3 shows a flowchart of the passing model, which clearly illustrates the steps to establish the passing network.

4.2.2. Sample Matches and Data Sets. Here, we adopt the Fullevents as sample data set, which is composed of 38 games played by Huskies. In the initial model building, we first select a match from Fullevents as a data sample to show the process of building the passing network graph. To achieve better performance, we utilize the passing rate rule to choose the sample match. The passing rate is defined as follows:

$$P(\%) = \frac{P(H)}{P(\text{total})}, \quad (2)$$

where $P(H)$ is the total number of passes made by Huskies team, and $P(\text{total})$ is the total number of passes for both sides of the game, which includes head events, simple pass, launch, high pass, and other subevents.

It is worthy to note that this passing rate is not a pass success rate, which may include incidents of failed passes. The pass rate indicator is used to screen out matches that are performing actively (or differently). After calculation, Match 6 has the highest passing rate, and hereafter we use it to illustrate. The detailed information of Match 6 is shown in Table 1.

In addition, based on the statistics of the pass frequency, we construct an $n \times n$ adjacency matrix, and it represents the adjacency relationship between vertices. In this case, n represents the number of players of the team in the game, and the columns of row j and matrix k represent player j and player passing through k . For example, for Match 6, there exist 14 players in this game totally; hence, its adjacency matrix is 14×14 . Note that since the passing is a one-way

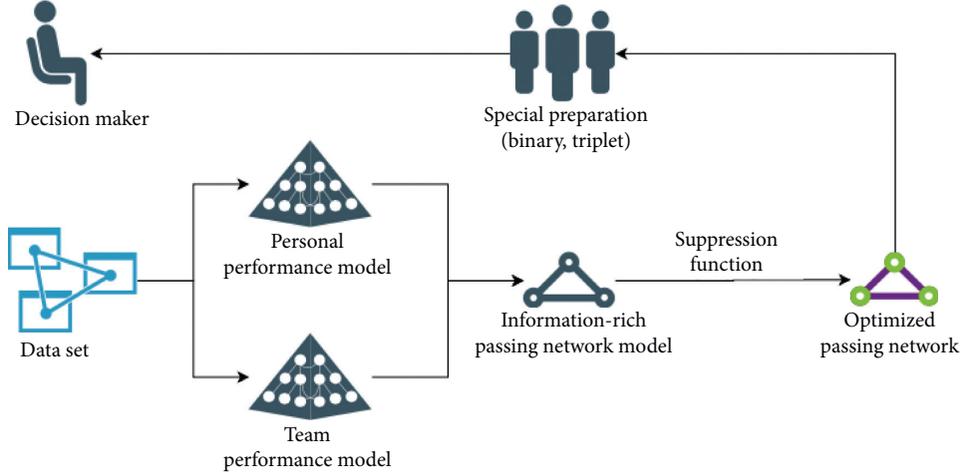


FIGURE 2: Overview of mining strategy for football social network.

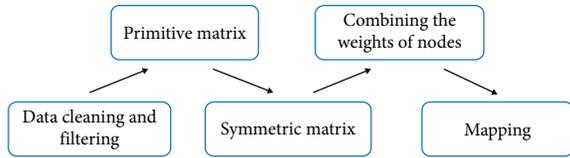


FIGURE 3: Overview of mining strategy for football social network.

TABLE 1: The information of Match 6.

Match ID	Outcome	Own score	Opponent score	Side
6	Win	2	1	Home

relationship, the adjacency matrix is a directed multivalued matrix. To reflect the two-way passing relationship among players, we transform the original matrix into a symmetric matrix based on the combination of out-degree and in-degree.

4.3. Player Performance Based on Social Network Analysis (SNA). After establishing the basic pass network, we need to quantify the player's importance indicators, i.e., centrality degree and betweenness, to enrich the information of player entities in the passing network.

4.3.1. Centrality Degree. For the centrality, we have the following definition.

Definition 1. Centrality is one of the evaluation indexes, which refers to the number of direct connections between nodes and other nodes in the network.

Note that if a node has the highest degree, it is at the centre of the local area network and is capable. In the passing network model, the output degree of the node represents the number of times the player passes the ball, and the input degree of the node represents the number of times the player obtains the ball. The total degree is the sum of the number of passes and catches.

4.3.2. Betweenness. For the betweenness, it is another indicator and can be defined as follows.

Definition 2. Betweenness refers to the degree to which the node is located "in the middle" of other nodes.

Specifically, a node is located on the shortcut of many other node pairs; that is, the node has a high intermediate centrality. The centre of a node measures how much that node controls the interaction between other nodes. If a player is between multiple pairs of players, even if his/her degree is low, he/she may play an important intermediary role. Therefore, this player is often located in the centre of the passing network. Here, the larger number of degrees means that the player represented by the node is at the centre of the passing network. This shows that players with a larger intermediary centre are regarded as "hubs" in the passing network, which means "connecting" players on the field, "metronomes" of team passing, and controlling the rhythm of the game.

4.4. Teamwork Performance Based on Multifaceted Analysis. Based on the individual performance model, we propose a team performance model to enrich the information of the passing network from a macro perspective. The team performance model is more inclined to the overall structural change of the passing network. Here, we introduce four indicators: coordination, adaptability, flexibility, and rhythm, to quantify the overall performance of the team.

4.4.1. Coordination. For ease of exhibition, we utilize the network density and distance to represent the coordination of the team. In the passing network, the network density and distance exhibit the closeness of the connection between players. For the network density, it refers to the ratio of the actual number of connections to the maximum number of possible connections in an informal network, which can be used to measure the closeness of the connections between network members. For the network distance, it refers to the length of the geodesic line between two points in the

network, which can be used to measure the minimum number of people who need to go through to get in touch with any two members of the network.

4.4.2. Adaptability. To measure the overall adaptability of the team, we consider two aspects of adaptability: ball possession and continuous pass ratio.

- (1) Ball possession can also be called ball control ratio.

The rate of ball possession is the ratio of time that one party controls the football during the game. The sum of the ball possession rates of two teams is 100%. The possession rate is used to detect who controls the initiative and the rhythm of the match. Generally, the higher the possession rate of a team is, the more mastery the team has in the match. But the ball control rate is only one of the monitored factors, which needs to be analysed along with other factors.

The ball control time is the length of time the team obtaining the ball during the game. Passing ball between any players belonging to the same team is regarded as the valid time. In addition, before being blocked by the opponent, the time of flying in the air is also regarded as the valid time. After being intercepted by the opponent, the valid time is belonging to opponent's possession time.

Here, we use the existing data to compute the ball control rate. Assuming that there is no significant difference about the time of touching the ball, the ball control rate can be regarded as the number of touching balls. Therefore, the ball control rate is the quotient of the number of touching balls and the total number of touching balls. Intuitively, the touching ball includes the number of passes, shots, and free kicks. Note that, the total ball possession ratio of the two sides is 100%. The calculation formula is as follows:

$$\gamma = \frac{\sum_i (P_{1i} + S_{1i} + F_{1i})}{\sum_{j=1}^2 \sum_i (P_{ji} + S_{ji} + F_{ji})}, \quad (3)$$

where γ is the rate of possession, $j=1$ is Huskies' player, $j=2$ is the opponent, P is the number of passes, S is the number of shots, and F is the number of free kicks.

- (2) For continuous pass ratio, we count the number of all single-pass ratio in a game to get the total number of pass ratios for that team in this match. If the total number of passing ratios is highly related to the possession rate, then the continuous pass rate can better reflect the team's overall passing control level. We have counted the number of successful passing ratios of two teams to get the total number of consecutive passing ratios. The ratio of the total number of consecutive passing ratios is the

continuous pass ratio, which is used to reflect the team's ability to maintain continuous pass ratio. The calculation formula is as follows:

$$\gamma = \frac{\text{consecutive passes}}{\text{passing times}}, \quad (4)$$

where γ is the continuous passing ratio.

4.4.3. Flexibility. The cooperation of the team is inseparable from the participation of each player. A player's pass participation rate can reflect the importance of his team cooperation. However, for the game, the opposing players tend to focus on defending players who have a high participation rate in passing coordination. Therefore, to evaluate the flexibility of the tactical coordination of the entire team, we pay more attention to the variance of passing coordination time, where passing coordination time represents the time of passing the ball from one player to the other. Note that the larger the variance is, the greater the difference in the participation rate between players is. Since the key players are more likely to focus on the offense and defense, hence, the smaller the variance is, the better the team's flexibility is. Its definition is as follows:

$$\sigma^2 = \frac{\sum (X - \mu)^2}{N}, \quad (5)$$

where σ^2 is the variance of the number of passing ratios, X is the number of passing fits per player, μ is the average number of passing fits, and N is the total number of players.

4.4.4. Tempo. The number of offensive and defensive conversions can reflect the rhythm of the entire game to a certain extent, which can also reflect the tempo well. The characteristic of modern football is confrontational, and the conversion is the immediate change of attack and defense after the confrontation. It is the physical and tactics of offense and defense reflected in the game. To defend well, sometimes the player only needs to destroy the ball to complete the task, but if he/she can get the ball right and motivate the offense; often the opponent can be caught off guard and exhausted, so that the team has more offensive initiative. By analysing the conversion of the ball, we can obtain the data of the players grabbing the ball and starting the offense, that is, the number of defenses. In the same way, the number of offenses and defenses reflects the rhythm of defense. Therefore, we define the number of offense and defense conversions as the sum of the offenses and defenses.

4.5. Suppression Function of Passing Network. After measuring individual performance and team performance, we can get a rich information passing network. At this moment, it is nontrivial to take the impact of the opponent's strategy on the passing network into consideration. Intuitively, when a team has star players, the rival team will take actions for the targeted defense. Generally, the rival team will arrange the defensive player to follow the star player closely, prevent our

star player from taking the ball, or cooperate with other teammates to cut off the receiving line of our player. Even if the individual ability of the star player is excellent, he/she does not have the ability to address all the problems. It is inevitable to cooperate with other teammates and enable themselves to avoid the rival threats. Therefore, when the opposing team has a better background knowledge of our players, the opposing team is more likely to cut off the connection between our players and the star players by blocking our players' ball-to-star routes.

To solve this problem, we have introduced a passing network suppression function to enhance the flexibility of our team's passing network. The function is as follows:

$$C(\text{PR}) = h^{\ln(\text{PR})}, \quad (6)$$

where h is the defensive ability coefficient of the opposing team. The stronger the opposing player's defensive ability is, the smaller h ($0 < h < 1$) is. PR is the importance of the player in the team. The higher the importance is, the smaller the suppression function value is, and the stronger the connection suppression effect on the side of the passing network is.

5. Experiment Evaluation

Data Set. In experiment, we used `fullMatches.csv` to get the result data of all matches in the league, `Matches.csv` to get the result of the target team Huskies, `fullEvents.csv` to get the time and location of each event in each game, and `Events.csv` to obtain the time and location of each game event of the target team.

Comparison between Methods. In this experiment, we compare the visualized passing network structure with our proposed passing network to illustrate the influence of specific factors on the passing network and analyse the strategy accordingly.

Setup. In experiment, we use `Ucient6.2` to calculate and count the basic attributes in the passing network. `MATLAB 2016b` was used to calculate the required optimization indicators and generate the required data format for the passing network. Finally, we use the `networkD3` package in `R Studio` to generate a visual passing network structure.

5.1. Comparison with Existing Passing Network. Firstly, we use only individual performance, only team performance, and both of these performances to generate the information-rich passing network. And we compare these passing network structures through the data of existing matches, calculate their winning rate, and compare with the basic passing network structure. Here we have made statistics on the winning rate of the team before the optimization and the winning rate of the game with the optimized network attributes in the data set for more than 300 games. The result is shown in Table 2. The result shows that, compared with the winning rate of the basic passing network structure, the win rate of our optimized

passing network considering only individual performance and only considering team performance has increased by 8% and 9%, respectively, and the passing network winning rate, which takes two performances together, increases by 12%.

5.2. The Results of Passing Networks. As shown in Figure 4, the nodes in the figure represent the players. The position and number of the player in the team are marked next to the node. The stronger the player's personal ability, the larger the circle displayed by the node. The connection between the nodes represents the number of direct passes by the player. The more passes, the thicker the connection.

We do not fix the position of the players because we were not able to know the formation of the team at first, so in our passing network diagram, the distance between the nodes has no meaning.

The dyadic configurations selected by the model include M1 & M3, M3 & D2, and M1 & F2. It can be seen that the attack and defense are centred on the two centre players of M1 and M3; the triadic configuration is D6 & D2 & M3 configuration, and the rear defensive coordination is good, ensuring that the backcourt ball quickly passes into the midfield and launches the attack.

However, we know that one side of a football game only allows 11 players on the field at the same time, and the reason this passing network has 14 players is that we also take into account the substitutes. Although this also shows the cooperation and interaction between players, this is different from the passing network during the same period of the game.

5.3. Effect Evaluation of the Suppression Function. The weighted influence factors are programmed into the model of the passing network. Take Match 6 as an example, and the first half and the second half of the first half of the network pass map as an example. Data shows that the opponent has suppressed the main members M1 and M4. We can see if a new formation can be formed to reasonably respond through the comparison, shown in Figure 5.

It can be seen that the game formation has changed, the core players have also shifted, and a new two-ternary structure has appeared.

5.4. Evaluation of the Four Quantitative Parameters. In this part, we utilize a mathematical model for multifaceted analysis to evaluate the overall performance of a team's game from four aspects: coordination, adaptability, flexibility, and tempo.

5.4.1. Coordination. The result is shown in Table 3; from the perspective of network density and network distance, the team's network distance is equal to 1. It means all nodes are reachable; there are no players without passing ball. Most of the nodes can be connected to other nodes, which shows that most players have a direct connection. By comparing the opponent's network density and distance data, we can see the difference between the two teams.

TABLE 2: Comparison of the win rate of different optimizations in the passing network.

Basic (%)	Individual performance only (%)	Team performance only (%)	Consider all (%)
34.21	42.67	43.13	46.14

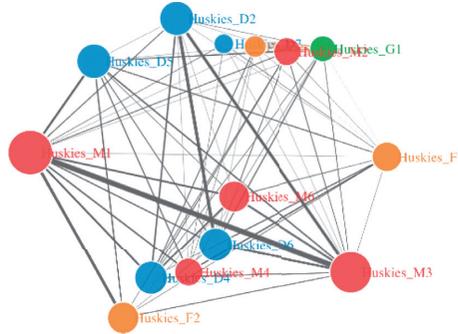


FIGURE 4: Match 6 huskies' passing network includes all players.

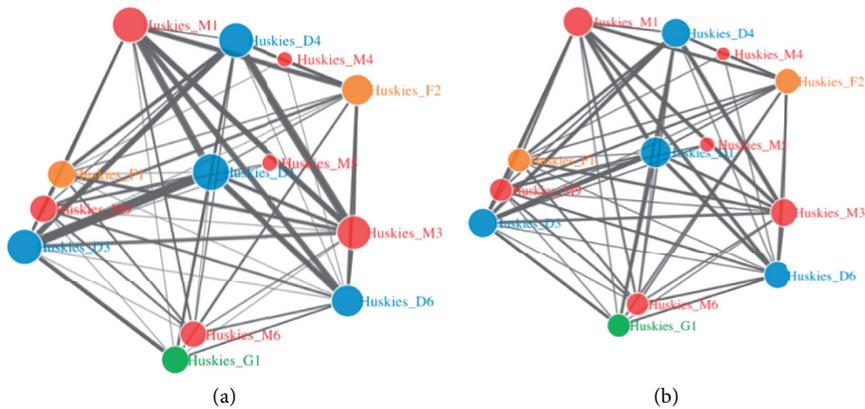


FIGURE 5: The comparison of original (a) and opponent's interference (b) in the first stage of Match 6.

TABLE 3: Index of passing track network of Huskies in Match 6.

Density	Distance
2.9906	1.0000

5.4.2. Adaptability. We calculated the possession rate and continuous passing rate of all the team in 368 games (one match against each opponent), both of which can reflect the team's adaptability. If the team adapts well, the possession rate and continuous passing rate are high. The result is shown in Figure 6.

5.4.3. Flexibility. As shown in Figure 7, we illustrate the flexibility indicators in 368 games.

5.4.4. Tempo. In 368 games (one match against each opponent), the number of offensive and defensive conversions (that is, team flexibility indicators) of all the team is shown in Figure 8.

5.5. Evaluating the Impact of Changing Player on the Binary and Ternary Relationship. Players are often changed in football matches. Therefore, we need to evaluate whether the optimized passing network is sensitive to the player change event. In other words, we need to evaluate whether the optimized passing network retains the previous information after the personnel changes. If the binary and ternary structures can still be identified, it turns out that the optimized passing network is insensitive to further player events and there is no need to rebuild the network every time a player change occurs. Instead, we need to rebuild the network every time you change players.

Let us take Match 6 as example. There are two substitutions in this game, so we can divide the total time of Match 6 into three subperiods. Time division and personnel changes are shown in Table 4.

As shown in Table 4, origin player ID represents the player being replaced. Destination player ID represents the replacement player. Match period represents matches, where 1H represents the first half of football match and 2H represents the second half of football match. Event time

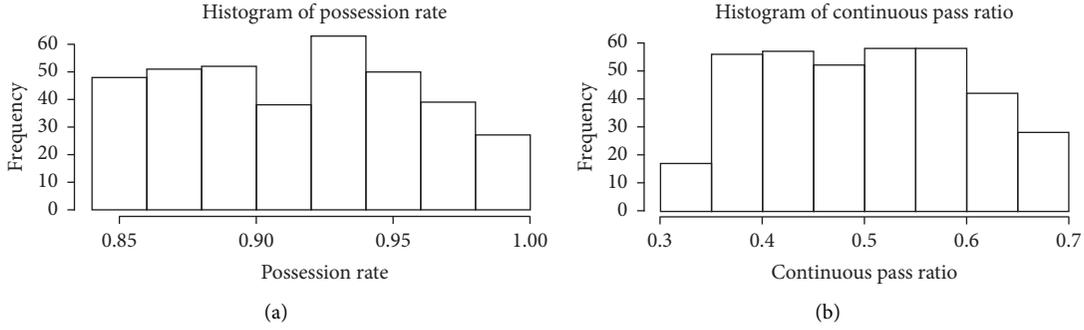


FIGURE 6: Distribution histogram of possession rate (a) and continuous pass rate (b) in 368 games.

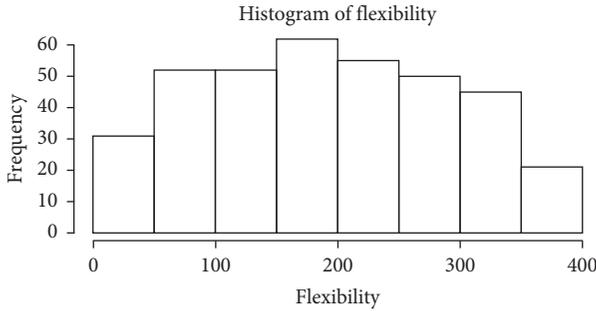


FIGURE 7: Distribution histogram of flexibility in 368 games.

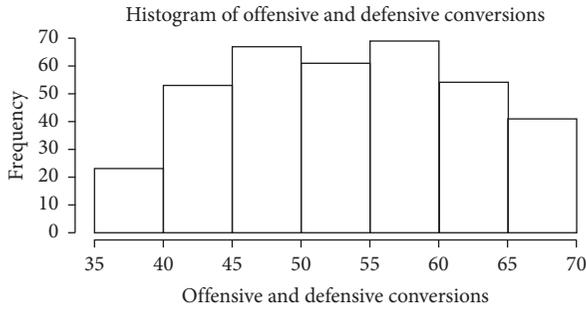


FIGURE 8: Distribution histogram of offensive and defensive conversions in 368 games.

TABLE 4: Huskies substitution in Match 6.

Original player ID	Destination player ID	Match period (H)	Event time
Huskies_M2	Huskies_M4	2	600
Huskies_D2	Huskies_F4	2	600
Huskies_D5	Huskies_D7	2	1860

represents how much time the game played when the substitution occurred in half of match.

Next, in each subperiod, we establish their optimized passing network. The structure of the optimized passing network in three time periods is shown in Figure 9.

Through Figure 9, we can get their binary and triplets. Finally, compare them with binary and triplets of optimized passing network identified in total time period. The results are shown in Table 5.

As shown in Table 5, the binary structure in the total time period is the same as the binary structure in the first period, and the ternary structure is the same as the ternary structure in the second period. From this, we can get that the information-rich passing network needs to be rebuilt when players are changed.

5.6. Evaluating the Impact of Team Performance on Final Score. Due to many factors, we need to determine whether the information contained within the team's performance is reasonable. In other words, consider whether the current state of the team can be correctly reflected based on the complex information. Under normal circumstances, the larger the certain index, the higher the chance of the team scoring. In this case, the indicator is correct. Therefore, we need to find matches with similar strategies and configurations but different scores. Figures 10 and 11 contain the passing network structure of the game under two similar strategies.

Although they have differences in the shape of the network caused by different point locations, the similarity of their data reaches more than 80%. At the same time, from Figures 10 and 11, it can be seen that most of the binary and triplets are the same; they all have the main binary configuration M1M3. Their results are shown in Table 6.

Through Figures 6–8, we get the indicators of coordination, adaptability, flexibility, and tempo in these four games and display them in Table 7.

As shown in Table 7, for Match 8 and Match 17, the main change is the flexibility parameters. The other three parameters are considered to be the same because they do not change much. We can get from Table 7 that Match 8 has a flexibility of 195.73 and Match 17 has a flexibility of 311.75. From Table 6, Match 17 has a better score than Match 8, which is in line with the facts.

For Match 7 and Match 13, their differences are mainly reflected in continuous pass rate and flexibility, so the other two parameters can be regarded as the same. We can get from Table 7 that Match 7's continuous pass rate is 0.657 and flexibility is 321.00; Match 13's continuous pass rate is 0.373 and flexibility is 105.63. From Table 6, Match 7 has a better score than Match 13, which is in line with the facts.

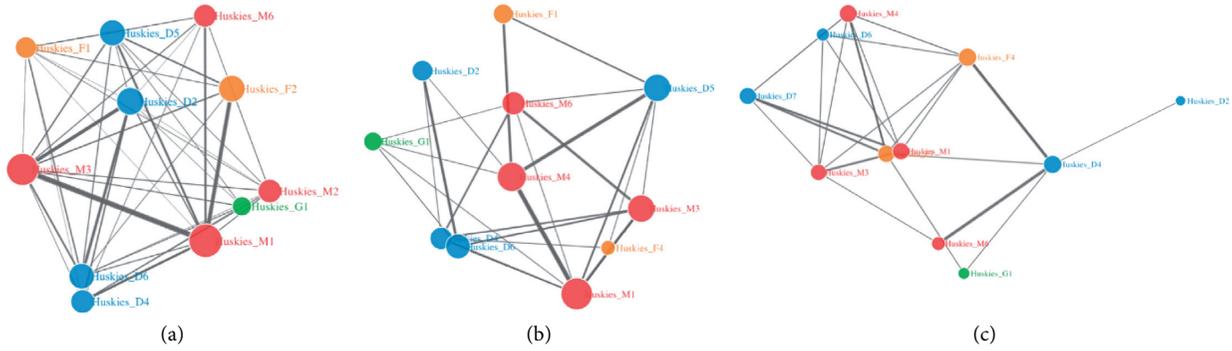


FIGURE 9: Three subperiod passing networks.

TABLE 5: Dyadic and triadic configuration during different period.

Configuration	Entire game	First subperiod	Second subperiod	Third subperiod
Dyadic	M1M3	M1M3	M1M4	
	M3D2	M3D2	D6D2	D4F4
	M1F2	M1F2	M1F3	D4M6
Triadic	D2D6M3	M1M4D5	M3M6	
			D2D6M3	None

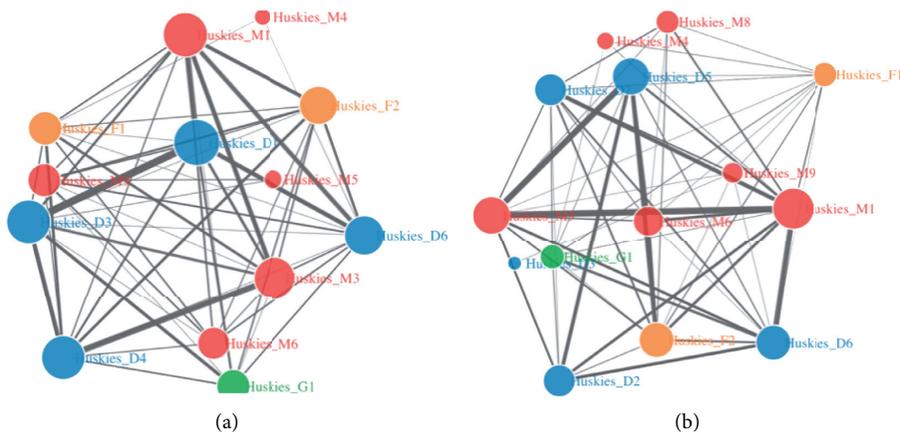


FIGURE 10: Match 8 (a) and Match 17 (b) passing network.

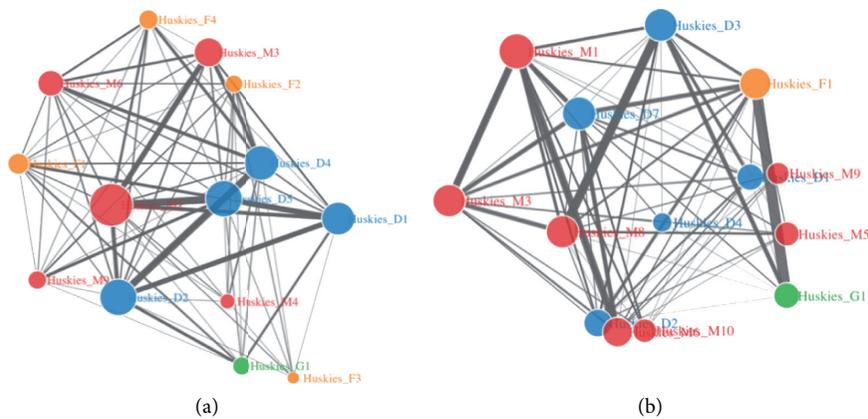


FIGURE 11: Match 7 (a) and Match 13 (b) passing network.

TABLE 6: Results of four games.

Opponent ID	Outcome	Own score	Opponent score
Opponent 8	Tie	1	1
Opponent 17	Win	1	0
Opponent 7	Lose	0	1
Opponent 13	Lose	1	4

TABLE 7: The performance indicators of the four games.

Index	Match 8	Match 17	Match 7	Match 13
Continuous pass rate	0.5198	0.4720	0.6570	0.373
Ball possession	0.9728	0.9440	0.9250	0.943
Flexibility	195.73	311.75	321.00	105.63
Tempo	57.00	56.00	49.00	52.00

In summary, it can be judged that our team's quantitative indicators have no internal errors, which is basically in line with the facts.

6. Conclusions

In this paper, we investigated and studied the issue of efficient strategy mining for social network. To the best of our knowledge, this is the first work that aims to optimize the passing network to address this problem. Compared with the traditional passing network, we first introduce the four aspects: coordination, adaptability, flexibility, and tempo, and integrate them into a passing network. Next, we propose a suppression function to express the impact of strategy. Based on the above schemes, we optimize the passing network to obtain the efficient strategy. Finally, through the comprehensive experiments, we demonstrate the feasibility of the proposed methods.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this article.

Acknowledgments

This work was partially supported by the National Natural Science Foundation of China (nos. U1711266, 41925007, and 62076224).

References

- [1] G. O. C. Bongomin, J. C. Munene, J. M. Ntayi, and C. A. Malinga, "Social network," *African Journal of Economic and Management Studies*, vol. 9, no. 3, pp. 388–406, 2018.
- [2] A. Chin, M. Chignell, and H. Wang, "Tracking cohesive subgroups over time in inferred social networks," *New Review of Hypermedia and Multimedia: Web Science*, vol. 16, no. 1-2, pp. 113–139, 2010.
- [3] S. Mclean, P. M. Salmon, A. D. Gorman, J. Wickham, E. Berber, and C. Solomon, "The effect of playing formation on the passing network characteristics of a professional football team," *Human Movement*, vol. 2018, no. 5, pp. 14–22, 2018.
- [4] J. Duch, J. S. Waitzman, and L. A. N. Amaral, "Quantifying the performance of individual players in a team activity," *PLoS One*, vol. 5, no. 6, Article ID e10937, 2010.
- [5] F. Clemente, F. Martins, D. Kalamaras, and R. Mendes, "Network Analysis in Basketball: Inspecting the Prominent Players Using Centrality Metrics," *Journal of Physical Education and Sport*, vol. 15, no. 2, pp. 212–217, 2015.
- [6] Y. Yamamoto and K. Yokoyama, "Common and unique network dynamics in football games," *PLoS One*, vol. 6, no. 12, Article ID e29638, 2011.
- [7] P. Passos, K. Davids, D. Araújo, N. Paz, J. Minguéns, and J. Mendes, "Networks as a novel tool for studying team ball sports as complex social systems," *Journal of Science and Medicine in Sport*, vol. 14, no. 2, pp. 170–176, 2011.
- [8] C. Cotta, A. M. Mora, J. J. Merelo, and C. Merelo-Molina, "A network analysis of the 2010 FIFA world cup champion team play," *Journal of Systems Science and Complexity*, vol. 26, no. 1, pp. 21–42, 2013.
- [9] F. M. Clemente, F. M. L. Martins, D. Kalamaras, P. D. Wong, and R. S. Mendes, "General network analysis of national soccer teams in FIFA World Cup 2014," *International Journal of Performance Analysis in Sport*, vol. 15, no. 1, pp. 80–96, 2015.
- [10] J. Ribeiro, P. Silva, R. Duarte, K. Davids, and J. Garganta, "Teamsports performance analysed through the lens of social network theory: implications for research and practice," *Sports Medicine*, vol. 47, no. 9, pp. 1689–1696, 2017.
- [11] K. Florian and L. Martin, "Passing Network Analysis of Positional Attack Formations in handball," *Journal of Human Kinetics*, vol. 70, no. 1, pp. 209–221, 2019.
- [12] F. M. Clemente, F. M. L. Martins, P. D. Wong, D. Kalamaras, and R. S. Mendes, "Midfielder as the prominent participant in the building attack: a network analysis of national teams in fifa world cup 2014," *International Journal of Performance Analysis in Sport*, vol. 15, no. 2, pp. 704–722, 2015.
- [13] S. Mclean, P. M. Salmon, A. D. Gorman, M. Naughton, and C. Solomon, "Do inter-continental playing styles exist? Using social network analysis to compare goals from the 2016 EURO and COPA football tournaments knock-out stages," *Theoretical Issues in Ergonomics Science*, vol. 18, no. 4, pp. 370–383, 2017.
- [14] J. Buldu, J. Busquets, I. Echegoyen, and F. Seirullo, "Defining a historic foot-ball team: using network science to analyse guardiola's F.C. barcelona," vol. 9, no. 1, pp. 1–14, 2019.
- [15] P. Cintia, F. Giannotti, L. Pappalardo, D. Pedreschi, and M. Malvaldi, "The harsh rule of the goals: data-driven performance indicators for football teams," in *Proceedings of the IEEE International Conference on Data Science and Advanced Analytics (DSAA)*, pp. 1–10, Paris, France, October 2015.
- [16] J. Li, T. Cai, K. Deng, X. Wang, T. Sellis, and F. Xia, "Community-diversified influence maximization in social networks," *Information Systems*, vol. 92, 2020.
- [17] T. Cai, J. Li, A. S. Mian, R. Li, T. Sellis, and J. X. Yu, "Target-aware holistic influence maximization in Spatial social networks," *IEEE Transactions on Knowledge and Data Engineering*, 2020.
- [18] H. Touchette, "A network theory analysis of football strategies," 2012, <https://arxiv.org/abs/1206.6904>.

- [19] C. M. Clemente, M. S. Santos, F. M. L. Martins, and R. S. Mendes, "Using network metrics in soccer: a macro-analysis," *Journal of Human Kinetics*, vol. 45, no. 1, pp. 123–134, 2015.
- [20] S. Sun, N. Akhtar, H. Song, C. Zhang, J. Li, and A. Mian, "Benchmark data and method for real-time people counting in cluttered Scenes using depth Sensors," *IEEE Transactions on Intelligent Transportation Systems*, vol. 20, no. 10, pp. 3599–3612, 2019.
- [21] F. A. Moura, L. E. B. Martins, R. D. O. Anido, R. M. L. De Barros, and S. A. Cunha, "Quantitative analysis of Brazilian football players' organisation on the pitch," *Sports Biomechanics*, vol. 11, no. 1, pp. 85–96, 2012.
- [22] R. Trequattrini, R. Lombardi, and M. Battista, "Network analysis and football team performance: a first application," *Team Performance Management*, vol. 21, no. 1-2, pp. 85–110, 2015.
- [23] C. Winter, C. Rasche, and M. Pfeiffer, "Linear vs. non-linear classification of winners, drawers and losers at FIFA World Cup 2014," *Science and Medicine in Football*, vol. 1, no. 2, pp. 164–170, 2017.
- [24] H. Srivastava and R. Sankar, "Information dissemination from social network for extreme weather Scenario," *IEEE Transactions on Computational Social Systems*, vol. 7, no. 2, pp. 319–328, 2020.
- [25] R. Zhou, J. C. Li, and J. X. Yu, "ELCA evaluation for keyword search on probabilistic XML data," *World Wide Web*, vol. 16, no. 2, pp. 171–193, 2013.
- [26] N. Tang, J. X. Yu, K. F. Wong, and J. Li, "Fast XML structural join algorithms by partitioning," *Journal of research and practice in information technology*, vol. 40, no. 1, pp. 33–54, 2008.