For office use only	Team Control Number	For office use only
T1 ________	**2009823**	F1 ________
T2 ________		F2 ________
T3 ________	Problem Chosen	F3 ________
T4 ________	**D**	F4 ________

2020 Interdisciplinary Contest in Modeling (ICM) Summary Sheet
(Attach a copy of this page to each copy of your solution paper.)

Understanding Teaming Strategies with A Double-Layer Network Model

Summary

The desire to live in a collective and interact with other individuals is deeply ingrained in human instinct. Throughout time, we have witnessed the evolution of a complex network in our society, followed by challenging problems. Our goal is to analyze a football team by understanding the modes by which team members cooperate and interact from a network perspective.

First, we process raw datasets and build a passing network with data of all matches. We then employ models like Minimum Spanning Tree and network motifs to unearth the structural and configurational patterns of the passing network. Afterward, we designate four correlated performance indicators like opportunity pass and use a Poisson Regression to confirm their validity. Besides quantitative methods, we also examine the passing networks of the 38 matches geometrically and calculate the graph similarity of each pair of networks. Though the networks implicate the variety of strategies used by the Huskies, certain patterns do exist within the games with the same result.

After that, we establish a double-layer network model to dig into the mechanism of teamwork. Grouping players from the same position like the Midfield and Defense, the bottom-layer network model primarily focuses on the passing tendency of the groups, interactions of players within a group and personal influences on the group. To better analyze the main body of the network, we simplify the original network by using PageRank and degree distribution. The upper-layer network model, on the other hand, adopts a semi-Markov process to explore the interactions and dynamics among the groups. An optimization model and Simulated Annealing are applied to draw useful insights into team performance improvement. Moreover, we explore desired strategies when faced with opponents with different playing styles like strong attack and weak defense.

Based on the elaborate analysis, we prepare a compact summary of current effective strategies of the Huskies and further provide reasonable suggestions according to our modeling results. For example, the centralization of the team the interactions between the Forward and other parts have been effective. Advised strategies include performing relatively more aggressively and straight-forward. Our model accurately depicts the external and internal factors which may influence group dynamics. Finally, we generalize the model and explore possible applications in Hockey games and hospital operations.

Keywords: Double-Layer Complex Network Model; Semi-Markov Process; Simulated Annealing; Performance Indicators; Team Designs

Contents

1 Introduction

1.1 Context and Motivation

Over time, human society has evolved and grown from small networks with limiting connections to the massive community which is formed to tackle perplexing problems. In any sense, one thing has remained constant and that is the inherent practice of teamwork, which is particularly crucial in competitive team sports.

We seek to develop a mathematical model of players networks to analyze team performance to its core. To do this, we need to understand how players interact with teammates and opponents, and the effects these interactions have on the resulting game performance.

1.2 Our Work

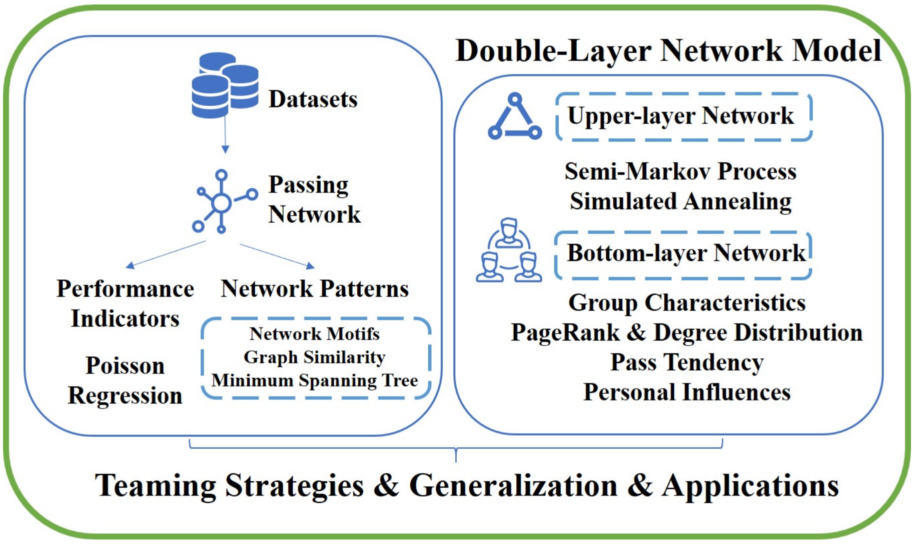


Figure 1: Summary of Our Work

2 Assumptions

A few assumptions are made to perform an effective model. These assumptions are the premise for our subsequent analysis.

1. The data given is authentic and reliable.
2. Players of Huskies can be categorized as Goalkeeper, Defense, Midfield and Forward. There is no player that can either be categorized to one group or categorized to another.
3. Players will act strictly in the team's interest. Players will not neglect their duties and will give full play to one's ability in the football field.

4. Players will follow the instructions of the coach.

5. The substitudes of the team are relatively less competitive than starters. The coach will not hide the star players.

3 Network Analysis

3.1 Network Patterns

Football is the game of teamwork. A single player helps the team by performing one's duty on the field and a team-level network is formed through collective efforts. In order to understand the basic mechanism of the tactical system of Huskies, we create a passing network to visualize the players' interaction. To obtain a more comprehensive characterization of Huskies and remove the interference of random factors, we take all Huskies players' passing data for the whole season into consideration. As previous studies have shown, a general network analysis including information of all members is considered more accurate in characterizing collective feature[1][2].

Figure 2 shows **the passing network of Huskies players**. It's worth noticing that these links are unidirectional and weighted according to the number of passes between two players in the whole season. Each node symbolizes a certain player and its size is proportional to the number of passes the player has made. It is a vivid depiction of the player's status in the tactical system and how they interact with their teammates. Specifically, we treat the goal as a node and involve it in the passing network.

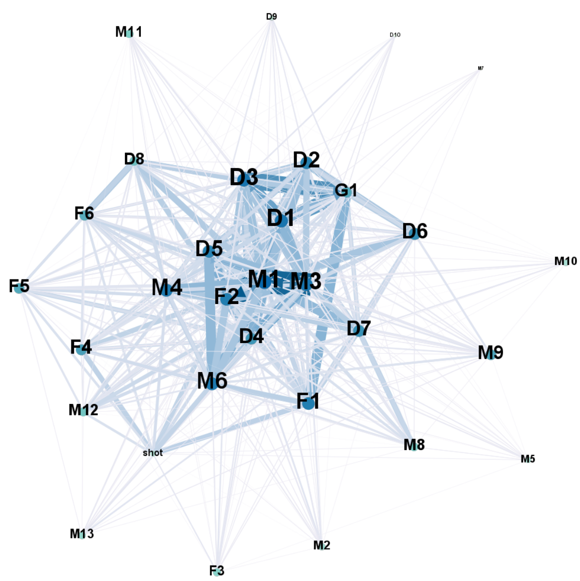


Figure 2: Passing Network of All Huskies players From Last Season

The above network contains 30 nodes and 379 edges, which is indeed a complex network. Therefore, it is necessary to extract the main body of the network to facilitate our analysis of it. We reduce the above network to an undirected weighted graph and apply the Kruskal algorithm to find **the Minimum Spanning Tree**. A minimum spanning tree (MST) or minimum weight spanning tree is a subset of the edges of a connected, edge-weighted undirected graph that connects all the vertices together, without any cycles and with the minimum possible total edge weight. It is widely used to unearth the primary connection and certain properties of a complex network. Note that we use the reciprocal of number of edges between two nodes as the topological distance here. Figure 3 shows **the minimum spanning tree** for the passing network.

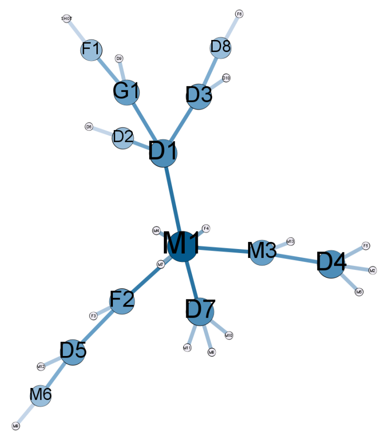
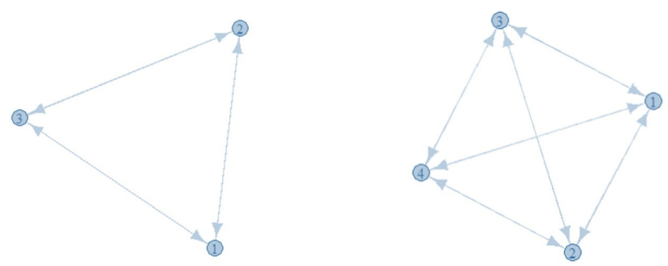


Figure 3: Minimum Spanning Tree

Based on the above figure, player M1 and player D1 both occupy a decisive position in the Huskies. They are the pivot of the whole team and have a strong connection with other players. This result echoes with the structural trait of passing network above. For example, player M1 is distinctly crucial in the passing network. It indicates that the Huskies is a centralized team and several outstanding players have greater influence on the game than others. Notice that player G1 is closely related to player F1. It sounds unreasonable at first, but as shown in the dataset, player G1 passes the ball to player F1 most frequently.

Further, we dive deep into the network patterns to find network motif. Complex networks often display network motifs which indicate dyadic and triadic configurations and team formations. Understanding the pass motifs is vital for the design of counter-strategy [3].

In the passing network of the Huskies, the most common motifs involving 3 and 4 players are as follows.



(a) Most Common Motif with 3 Vertices (b) Most Common Motif with 4 Vertices

Figure 4: Most Common Motif

In the most common motifs shown above, all player pairs are connected, which means the directed passing network is highly connected. As shown in the passing network and the MST, player like M1, M3, D1, D7, F2 are closely connected and forms many dyadic and triadic patterns.

3.2 Performance Indicators

In Section 3.1, we discuss the geometrical properties of the passing network and obtain the network patterns. In this section, we quantify the network by setting up performance indicators and justify their correlation with the team's performance[4]. This would enable us to have a profound understanding of the network.

- *Shot Success Rate* *Shot Success Rate* is a direct measure of a team's offensive capability. We define *Shot Success Rate* as the ratio of goals made to total shots[5].

$$\textit{Shot Success Rate} = \frac{\text{Goals Made By Huskies}}{\text{Total Shots by Huskies}}$$

- *Distribution of Contribution* *Distribution of Contribution* A balanced distribution of contribution is vital to a team. Assume possession of the ball indicates the player's contribution, we count the times of each players' possession and use the variance to quantify the distribution. Let N_i be the number of touches of each Huskies players($i = 1, 2, \dots, 30$). *Distribution of Contribution* is defined as:

$$\textit{Distribution of Contribution} = \frac{\sum_{i=1}^{30}(N_i - \overline{N_i})}{30}$$

- *Opportunity Pass* A basic principle in football games is that a player is more likely to score when he or she is closer to the goal. When the ball is passed to the forecourt, it creates an opportunity and forms a threatening attack. We define *Opportunity Pass* as the numbers of passes directed at the first third of the court(which is closer to the opponent's goal).

- *Mean Communicability* Many dynamical ties of complex networks are closely related to shortest paths. However, there are different scenarios in which non-shortest paths are also used and the passing network is indeed such an example.

  Therefore, to understand the internal mechanism of how players are connected with each other, we introduce the concept of communicability. Communicability indicates the topological property of a complex network, which belongs to the broad generalization of the shortest-path concept. The communicability between nodes u and v based on the graph spectrum is defined as:

$$C(u, v) = \sum_{j=1}^{n} \phi_j(u)\phi_j(v)\lambda_j,$$

  where ϕ_j(u) is the u^{th} element of the j^{th} orthonormal eigenvector of the adjacency matrix associated with the eigenvalue[6].

We calculate the communicability between two nodes by employing an algorithm which is essentially a spectral decomposition of the adjacency matrix. In order to analyze team-level performance, we define *Mean Communicability* as the average of all communicability values between each player in Huskies.

With the performance indicators defined above, we apply **Poisson regression** to explore whether these performance indicators have a statistically significant effect on game performance. Note that we use the data of the Huskies in 38 matches since in such a controlled setting we can ignore the influence of team-related characteristics like playing styles.

Poisson regression is a generalized linear model form of regression analysis that is widely used to model count data. It assumes the response variable has a Poisson distribution and the logarithm of its expected value can be modeled by a linear combination of unknown parameters. In the case of a football game, the event of scoring a goal in a match is generally assumed to follow the Poisson distribution and the outcome of a game is always an integer. Therefore, we could use a Poisson regression model to decide whether these performance indicators are coefficient with the game performance[7].

Here we regard **Goal Difference(GD)** as the response variable to test if the performance indicators defined before are reasonable. We conduct a coefficient analysis including these four performance indicators. Table 1 shows the coefficient result.

Table 1: Coefficient Performance Indicators

Coefficient name	Coefficient Estimate	Standard error	Z.value	P.value
Shot Success Rate	0.8845	0.3618	2.445	0.0145
Distribution of Contribution	3.938e-04	1.992e-04	1.977	0.0481
Opportunity Pass	8.508e-03	5.010e-03	1.698	0.0895
Mean Communicability	-1.341e-05	6.498e-06	-2.064	0.0390

As we can see from the table above, the performance indicators we select are strongly correlated with the performance outcome.

3.3 Network Similarity

In the two sections above, we regard all matches in a season as a whole and calculate its properties. And here comes some intriguing questions:Are there apparent similarities and differences between these matches? Would Huskies apply different strategies against different opponents? In this section, we will discuss the connection between the games and compare each game's graph based on the concept of **similarity**.

Similarity is a quantitative measure of the likeness of two graphs.We adopt a graph algorithm to evaluate whether nodes and links of the selected two networks resemble each other. Let A_1 and A_2 be the adjacence matrices of graphs G_1 and G_2 respectively. Let also $L_1 = D_1 - A_1$ and $L_2 = D_2 - A_2$ be the laplacians of the graphs, where D_1

and D_2 are the corresponding diagonal matrices of degrees. Based on this method, we calculate the eigenvalues of the laplacians and we define the similarity between two graphs as[8]:

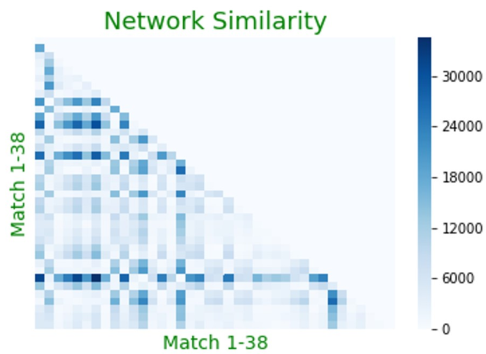
$$\text{sim} = \sum_{i=1}^{k} (\lambda_{1i} - \lambda_{2i})^2$$

where k is chosen s.t.

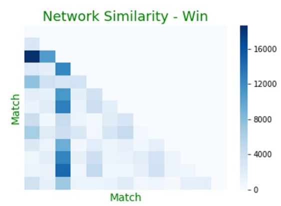
$$\min_j \left\{ \frac{\sum_{i=1}^{k} \lambda_{ji}}{\sum_{i=1}^{n} \lambda_{ji}} \right\} \geq 0.9$$

The constraint condition here indicates that we preserve over 90% of the energy.

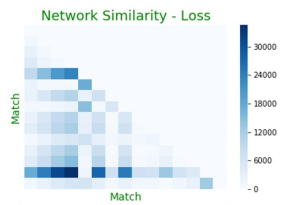
It's worth noticing that if the passing network of one match is similar to that of the other, the similarity between the two matches will be relatively low, vice versa,. Figure 5 is the similarity matrix of Huskies' all 38 games in this season. If two game are more similar in structure than others, their similarity value is lower and their corresponding cube is lighter in colour.



(a) Network Similarity between 38 Huskies Passing Networks



(b) Network Similarity between Passing Networks - Win



(c) Network Similarity between Passing Networks - Loss

Figure 5: Network Similarity Matrix

By comparison between Network Similarity Matrix of wins, losses and overall games, we conclude that this matrix successfully visulizes the difference between different types of the games. For wins and losses, their graph is comparatively lighter than the overall games, indicating that there exists certain pattern of games. In other words, if two games share the same result, their passing networks would look much alike.

4 Modeling

4.1 Introduction to Double-Layer Network

As we can see from the indicators above, there are numerous approaches to measure a team's performance. However, if we want to gain a deep insight into the dynamics of a football team and provides a practical strategy for the coaching team, analyzing a single layer of the network is far from satisfactory. Thus, **a complex network with multi-layers** is more effective in analyzing how a passing network exactly works. Each layer in the complex network runs independently and they also impact each other.

In football games, there are several positions on the court, each of which plays a distinct role in the game. Therefore, we classify the Huskies team into four groups: **Goalkeeper(G), Defense(D), Midfield(M) and Forward(F)**. When Huskies has the ball, one of these four groups must own the ball (since we neglect the time of ball passing). Thus, we could define these four groups as four nodes in the new double-layer network. When the ball is not in control of Huskies, there are two scenarios: the opponents take the ball or Huskies scores a goal, so we add two more nodes: **Shot and Opponent**. With the six nodes combined, we are capable of building **a double-layer network**.

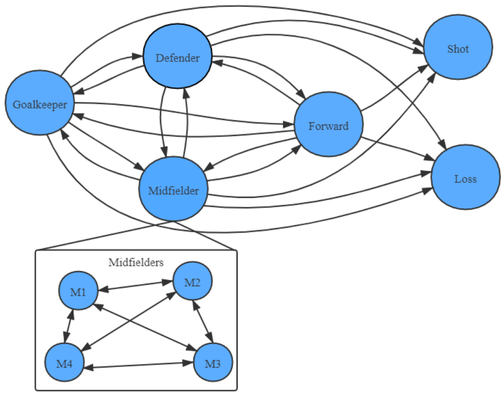


Figure 6: double layer network

Figure 6 depicts the double-layer network. For the upper layer, there are six nodes and the ball flows among them under a certain probability distribution. For the bottom layers, the ball travels between players of the same position, such as midfielders. We can observe team-level cooperation by observing the upper layer and get micro comprehension by looking into a certain node.

4.2 Bottom-layer : Group-level Analysis

4.2.1 Description of the Model

We assume that the players in the same position have a similar pattern so that the players of different roles are divided into groups. According to the assumption, the interaction between each player can be respectively analyzed as passes in the group and among the groups. As Figure 7 shows, the passes among the groups are considered as the external behavior of a group instead of a player's action, and passes in the group are considered as the internal moves which have nothing to do with other groups. The analysis method doesn't mean that the internal passes are neglected, but the passes' influence is contained in the external attributes of the group. A certain player's impact can be evaluated similarly, according to the changes in the external properties of the groups.

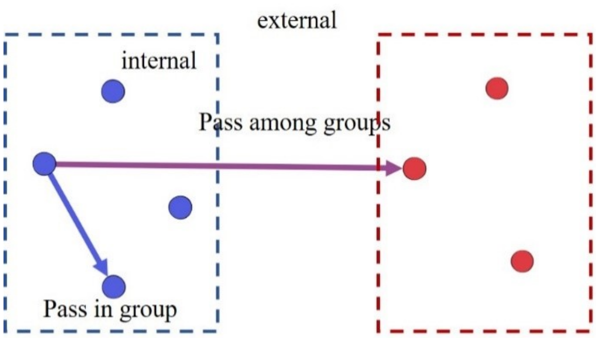


Figure 7: Internal and External Passes of a Group

Considering that the groups are similar in structure but different in functions, the same analysis can be imposed on the groups and different evaluations can be made according to the upper layer's requirements.

4.2.2 Indicators

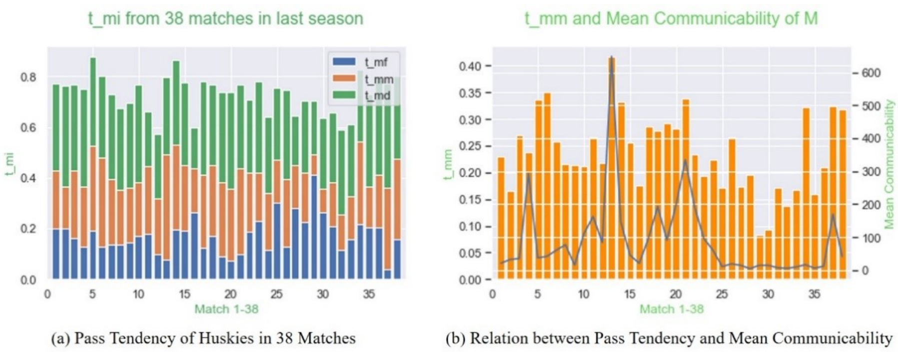
To evaluate the external property of a specific group, we use the passing tendency of the group to other groups as the indicators, including both the tendency that the ball staying in the original group and the ball passed to other groups.

The tendency t_{ij} can be calculated as the ratio of passes to different groups.

$$t_{ij} = \frac{\text{passed made between i and j}}{\text{all attempts of i}}$$

Attempts include successful passes, fail passes(lose possession of the ball) and shots. t_{ii} reflects the tendency that the players of i pass the ball between each other while t_{ij}, where $i \neq j$, reflects the tendency that the group passes the ball to other groups.

4.2.3 Group Effect



(a) Pass Tendency of Huskies in 38 Matches (b) Relation between Pass Tendency and Mean Communicability

Figure 8: the Pass Tendencies of the Midfield Group

Figure 8 shows the pass tendencies of the Midfield group (figures of other groups are shown in the Appendix) across the entire season. Tendency differences reflect the variance of team strategy in different games. To depict the tendency of players to pass inside the group, Mean Communicability, which is proved related to the consequences of the matches in Section 3.2, is used to evaluate the internal relationships. The Mean Communicability is positively correlated to the tendency of a group to control the ball. Hence, this index can be a reminder of team strategy. When the team requires a specific group to keep the ball for longer time, both adding a member to the group and increasing the passes among the group are possible solutions. The question of how to adjust the tendency will be discussed in the model of the upper layer.

4.2.4 Individual Effect

To analyze a specific player's influence on the bottom layer, we simplify the overall network based on PageRank and degree distribution.

- **PageRank** PageRank (PR) is an algorithm used by Google Search to measure the importance of website pages. It works by counting the number and quality of links to a page and it assumes more important websites are likely to receive more links. As for football players, we can use PR to obtain the 'rank' of players similarly.

Figure 9: 'Page Rank' of Players

As shown in the above figure, we can tell that player M1, F2 and D1 play important roles in the team, which is in accordance with the analysis in Section 3.1. Meanwhile, some of the players like D10 and M10 are relatively less important. Since they are substitutes who spend little time in the football court, the 'rank'is reasonable.

- **Degree Distribution** Let us consider the following problem. Suppose there are s people. Randomly draw n pairs from the group with replacement. The probability of drawing any person in the group is the same. Denote the number of times that person A and B are drawn as a pair to be X, which is a random variable. Then what is the probability that $X = x$? The probability of drawing A and B as a pair in one trial is $p_{AB} = \frac{2}{s(s-1)}$. Then $P(X = x) = \binom{n}{m} \cdot p_{AB}^x \cdot (1 - p_{AB})^{(n-x)}$.

  The above scenario can be an analogy to the degree distribution of a random unweighted graph with s nodes and n edges. Based on the passing network above, $s = 30$, $n = 10697$. Therefore, when $x = 25$, the probability is approximately the largest. Further, when x is too small or too big, the probability is approximately 0.

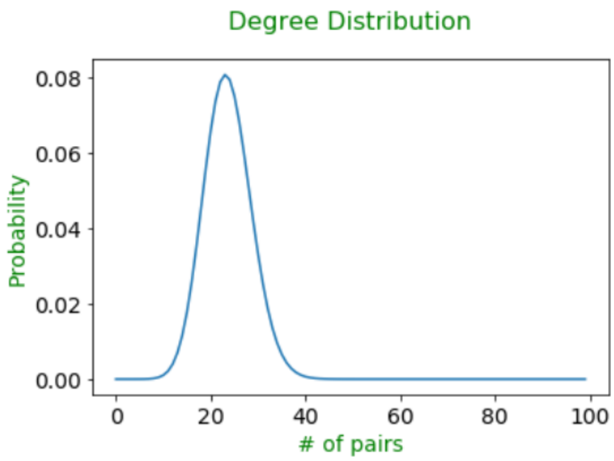


Figure 10: Degree Distribution

Combing the results we obtain in PageRank and degree distribution, the last 10 players in PageRank (Player F3, M2, M5, M7, M9, M10, M11, M13 D9, D10) are excluded from the entire network and edges that represent pass times less than 20(0.5 per game) are removed. Figure 11 shows the simplified network.

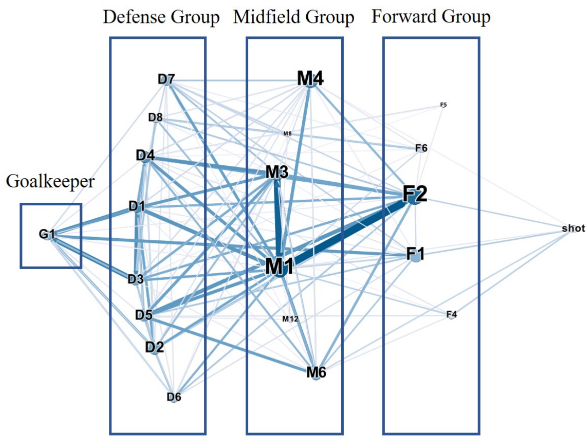


Figure 11: Simplified Passing Network of Huskies Players From Last Season

The remained players attend much more matches than the excluded players, so they can be seen as key players or regular substitutions. As the larger number of attended matches, the differences in the pass tendencies of each group with and without these players are considered significant.

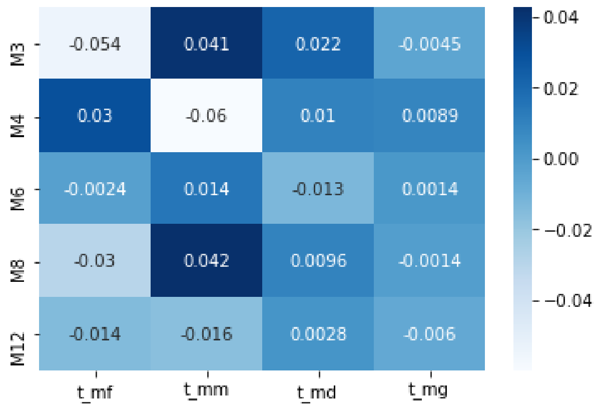


Figure 12: Tendency Difference of Middlefield Group with/without a Middlefield player

Player M1 attended all the matches so his influence is not calculated. The differences with/without Forward players and Defense players are shown in the Appendix. From the figure above, different player's traits can be drawn. For example, M4 is more likely to pass the ball to the Forward instead of other Middlefield players, so M4 plays more aggressively than others. The differences can be used to instruct the lineup and substitution plan against different opponents and form a more reasonable structure of the team, which will be discussed later.

4.3 Upper-layer : Team-level Analysis

4.3.1 Description of the Model

In this part, we use a bottom-up approach to model the dynamics of ball possession as a Semi-Markov Process. Semi-Markov process is a generalization of Markov process. Assume the transition process of ball possession has Markov property. Markov property refers to the memoryless property of a stochastic process. More specifically, the conditional probability distribution of future states depends only upon the present state, not on history. It is a dynamic process that can be in any one of N states $1, 2, \dots, N$, and that each time it enters state i it remains there for a random amount of time having mean w_i and then makes a transition into state j with probability p_{ij} [9]. The semi-Markov process is a generalization of the Markov process. It is a dynamic process that can be in any one of N state

We follow the idea of aggregating players of the same position as one node and define 4 corresponding states $\{F, M, D, G\}$. For example, state D refers to the possession stage where the Defense holds the ball. State S means the team's own shooting. Further, we treat the whole opponent team as one node and define state O to be the possession stage where the opponent team gets the ball. Here we assume that the transition probabilities between two states remain constant through the entire sequence of events. Note that the probability transition matrix defined here is equivalent to the idea of passing tendency defined in Section 4.2.2.

In a stochastic system, first passage time is the time taken for a state variable to reach a certain value. Based on the model and data drawn from the given dataset, we can calculate the mean first passage time from any state to S and the mean inter-visit time of S. As our Poisson Regression model in Section 3.3 indicates, shot success rate is highly related to the team's performance. Therefore, we focus on the mean inter-visit time of S and build an optimization problem to minimize it.

4.3.2 Notations

In order to clearly describe the flow of ball specifically,we set up several index to analyze upper layer.

Table 2: Notations

Symbol	Definition
$\mathbf{T} = \{F, M, D, G, O, S\}$	Set of Semi-Markov Process states
F	State: Possession by Forward
M	State: Possession by Midfield
D	State: Possession by Defense
G	State: Possession by Goalkeeper
O	State: Possession by Opponent team
S	State: Shooting
$FOW(G) = \{F, M, D\}$	Set of states closer to the goal than G
$FOW(D) = \{F, M\}$	Set of states closer to the goal than D
$FOW(M) = \{F\}$	Set of states closer to the goal than M
$FOW(F) = \{S\}$	Set of states closer to the goal than F
k	Relative Competitiveness of the opponent, $k \in \mathbb{R}$
p_{ij}	The probability of transition from state i to j
w_i	Expected sojourn time of state i
m_{ij}	Expected first passage time from state i to j

4.3.3 Optimization Model

$$
\begin{aligned}
\min \quad & m_{OS} \\
\text{s.t.} \quad & \sum_{j \in \mathbf{T}} p_{ij} = 1, && (1) \\
& p_{iO} = f(Q_i, k), \forall i \in \mathbf{T} \backslash \{O, S\}, Q_i = \{p_{ij} \mid j \in FOW(i)\} && (2) \\
& w_i = g(p_{ii}, i), \forall i \in \mathbf{T} \backslash \{O, S\} && (3) \\
& w_O = h(w_F, w_M, w_D, k) && (4) \\
& m_{ij} = w_i + \sum_{k \in \mathbf{T}, k \neq j} p_{ik} m_{kj}, \forall i, j \in \mathbf{T} && (5) \\
& w_S = 1 && (6) \\
& p_{GG} = 0 && (7) \\
& p_{SO} = 1 && (8) \\
& p_{OO} = 0 && (9)
\end{aligned}
$$

As we have discussed before, the objective is m_{OS}. Constraint (1) ascertains that we will obtain a probability transition matrix of a Markov chain. Constraint (2) depicts the positive correlation between the probability of losing possession and the probability of passing the ball towards the goal. The function f is a user-defined increasing function. Constraint (3) states that w_i is related to p_{ii} and state i itself where function g can also be user-defined. Specifically, for state O, w_O is also related to k. The more competitive the opponent is, the longer w_O will be. Constraint (5) shows how to calculate w_{ij} based on w_i and p_{ij} which is the property of a Semi-Markov Process. The remaining constraints are about the assumptions, for example, we assume once we reach state S, the next state is L with probability 1. Therefore, the inter-visit time m_{SS} is equivalent to m_{OS}. We set the optimization objective to be m_{OS}.

Here we will use Simulated Annealing to perform an approximate optimization. Detailed optimization results are as follows.

4.3.4 Optimization Results

Denote P to be the probability transition matrix with original parameters extracted from the given dataset. Denote optP to be the optimized probability transition matrix. The optimized w_{OS} is 114.28s while the original w_{OS} is 223.26s, reducing 48.79% while the difference between optP and P is smaller than 5%. Therefore, by changing the probability transition matrix slightly, the Huskies may obtain better performance.

$$P = \begin{pmatrix} 0.120 & 0.282 & 0.264 & 0.002 & 0.269 & 0.062 \\ 0.165 & 0.253 & 0.325 & 0.008 & 0.225 & 0.023 \\ 0.168 & 0.255 & 0.248 & 0.074 & 0.248 & 0.008 \\ 0.172 & 0.105 & 0.332 & 0 & 0.390 & 0 \\ 0.186 & 0.321 & 0.410 & 0.084 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \end{pmatrix} \quad optP = \begin{pmatrix} 0.114 & 0.285 & 0.264 & 0.002 & 0.269 & 0.062 \\ 0.157 & 0.240 & 0.325 & 0.008 & 0.225 & 0.023 \\ 0.160 & 0.242 & 0.236 & 0.074 & 0.248 & 0.008 \\ 0.163 & 0.099 & 0.315 & 0 & 0.390 & 0 \\ 0.186 & 0.321 & 0.410 & 0.084 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \end{pmatrix}$$

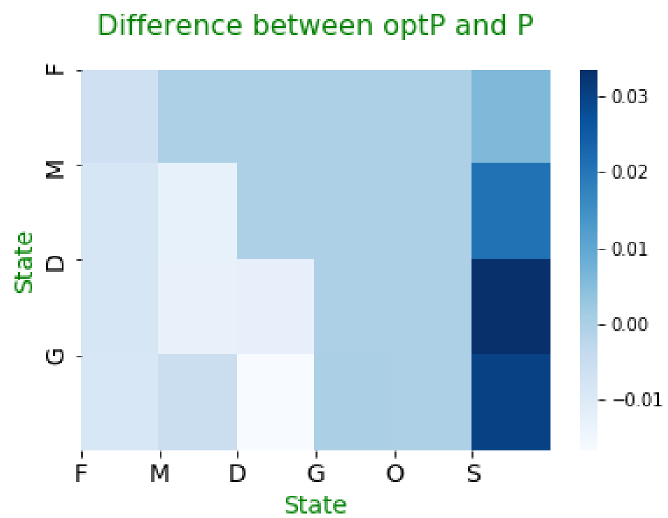


Figure 13: Optimization Results of General Case

Since we aggregate data from all matches, we presume the aggregated opponent can be a general representation. Further, based on the dataset about matches information, we presume the aggregated opponent is as competitive as the Huskies, indicating $k = 1$. Therefore, we can draw useful insight into universally effective strategies the

Huskies can adopt. From the heatmap above, we can see clearly that in order to perform better, the Huskies may need to perform more aggressively and straight-forward.

4.3.5 Specific Case Study-Sensitivity Analysis

Now we will consider the influence of different opponent's playing styles on the counter-strategy.

1. **Strong attack & Weak defense** To depict this playing style, we set $p_{OF} = 1$ and get the optimized probability transition matrix optP_d. The optimized m_{OS} is 108.19s under this setting. Then we use a heatmap to depict the difference between optP and optP_d.

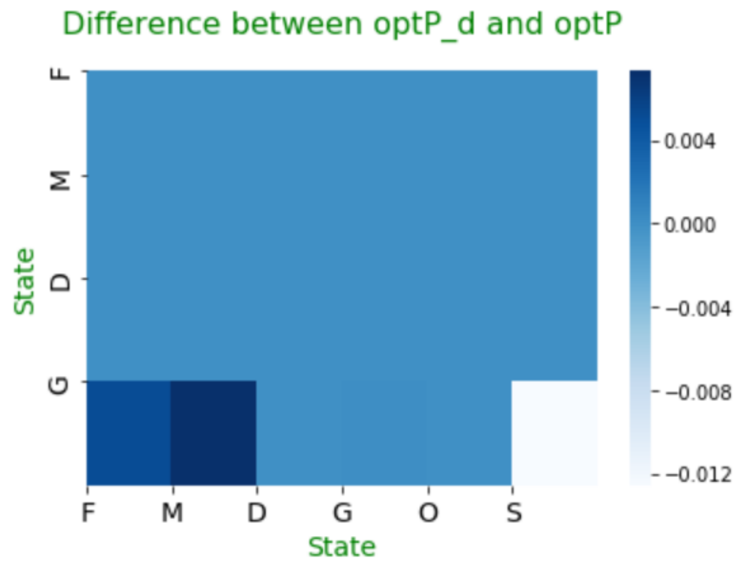
   

   Figure 14: Strong attack & Weak Defense Case

   From the heatmap above, we can see clearly that when faced with a team featured with strong attack and strong defense, the Goalkeeper of the Huskies may need to pass the ball to its Forward and Midfield more frequently in order to obtain shorter inter-visit times, which is in accordance to our intuition.

2. **Weak attack & Strong Defense Case** To depict this playing style, we set $p_{OD} = 1$ and get the optimized probability transition matrix optP_f. The optimized m_{OS} is 120.26s under this setting. Then we use a heatmap to depict the difference between optP and optP_f.

   From the heatmap above, we can see clearly that when faced with a team featured with weak attack and strong defense, to perform better, the Forward and Midfield of the Huskies may need to pass the ball back and avoid confronting the strong opponent's Defense head-on, which is also in accordance to our intuition.

3. **Competitive Opponent** As we defined before, k can depict the competitiveness of the opponent. The larger k is, the more competitive the opponent is. Therefore, we keep the probability transition matrix unchanged and set $k = 2$ to understand

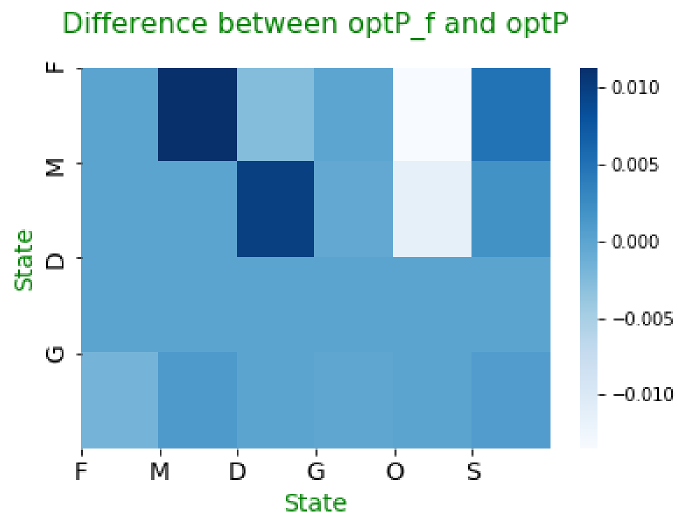


Figure 15: Weak attack & Strong Defense Case

how a competitive opponent affects the Huskies counter-strategy. Denote the optimized probability transition matrix to be optP_s. The optimized m_{OS} is 161.50s under this setting since a more competitive opponent will hold the ball longer.

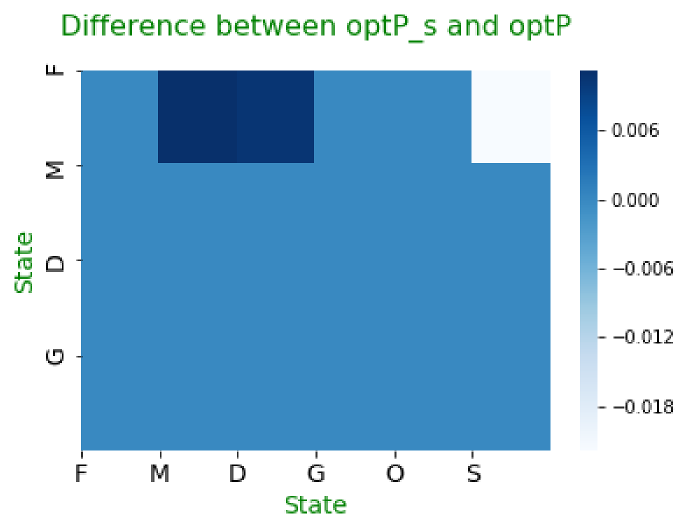


Figure 16: Competitive Opponent Case

From the heatmap above, we can see clearly that when faced with a more competitive team, the Forward of the Huskies may need to pass the ball back more frequently, no matter how competent the opponent's Defense is.

5 Strategy

In this section, we would focus on task 3 in the problem set, summarizing our model and refining it to insights which are conducive to Huskies's long-term development.

1. **We use the results and insights obtained in the above sections to inform**

the coach about what kinds of structural strategies have been effective for the Huskies.

- The overall passing network has been adaptive and flexible while stable.
- The dyadic and triadic configurations have been greatly contributing to the formations of the passing network and formed network motifs to support attacks and defenses.
- The centralization of the team has facilitated the ordered coordination and organization of attacks and defense.
- The interactions between the Forward and other parts have been effective.

2. **We advise the coach on what changes they should make next season to improve team success.**

    - **Adjusting the player number of different groups.**When the team requires a specific group to keep the ball for a longer time, adding a member to the group is a direct solution.
    - **Choosing players according to opponents.** There are different player styles in Huskies according to Figure 12. Among Forward players, Player F5 and F6 are more likely to interact with other Forward players, while Player F2 is a more conservative player, for he makes the Forward group pass more to Middlefield and Defense Group. For Midfield Players, M4 is more willing to pass forward, while M3 and M8 increase the control ability of the Midfield Group. In the Defense Group, Player D1 and D4 are defensive players while D8 is more aggressive for his increment in passes to the Forward group.
    - **Perform relatively more aggressively and straight-forward.**
    - **Adapt strategies according to the playing style of the opponent.** When faced with a team featured with strong attack and strong defense, the Goalkeeper of the Huskies may need to pass the ball to its Forward and Midfield more frequently. When faced with a team featured with weak attack and strong defense, to perform better, the Forward and Midfield of the Huskies may need to pass the ball back and avoid confronting the strong opponent's Defense head-on. When faced with a more competitive team, the Forward of the Huskies may need to pass the ball back more frequently.

6 Application

Although football is only one specific team sport, our modeling ideas can be generalized into other fields that require teamwork.

1. Divide the team into parts that have different functions but similar structure. Each part contains several individuals.
2. Consider the individual's influence on his/her group and neglect the individual's direct influence on the whole team. The individual's impact will be contained in the group's external properties and eventually affect the team.

3. Model the interaction between groups. By step 2, the complexity has been decreased so that it becomes easier to optimize cooperation among groups and provide general suggestions for improvement.

4. Reconstruct suggestions on the group level. Usually, we will get two aspects of improvement. For the whole group, it can improve by better structural strategy. For the individuals, they can adjust themselves according to their influence on the group, making the group conforms to the requirement of the optimization of the team.

- **Scenario 1: Ice Hockey**

  Ice Hockey is a popular team sport which has much in common with football. One team contains 3 Forwards, 2 Guards and 1 Goalkeeper. As the players at the same position have similar missions, we can construct a network, in which Forwards, Guards, and the Goalkeeper are abstracted as 3 nodes. Then we can analyze the interaction of the network, similar to what we have done in the paper. To construct a reliable model, more data are needed and special rules in ice hockey should take into consideration.

- **Scenario 2: Hospital surgery**

  Hospital is a typical scenario that needs excellent teamwork to function properly. For example, to cure a surgical patient, we need diagnosis, surgery and recovery procedure, which involves the cooperation of the corresponding groups. Follow our modeling idea, we can study the interaction of these involved groups. It has to be noted that the interaction mechanism is greatly different from a football game, for one is a sequence situation and another is a simultaneous case. Hence, a different upper-layer model should be established and more specific data need to be collected.

7 Strength and Weakness

7.1 Our Strength

- Our passing network involves the goal as a node, which can better depict the pass tendencies and attacking patterns.
- By using PageRank and degree distribution, we successfully simplify the passing network, thus making the main body of the network prominent.
- Instead of linear regression, we use a Poisson Regression to test the statistic significance of the performance indicators defined, which better fits the games setting. Our indicators are also validified.
- Our double-layer network model follows the idea of divide-and-conquer. Upper-layer model takes the interaction between two competing teams into consideration while bottom-layer model explores the dynamics within groups of players.
- Our optimization model is not sensitive to the different playing styles and the relative competitiveness of the opponent, which guarantees its robustness.

- We have splendid presentations of our results through a variety of figures and charts with well-designed colors combinations, especially using heatmaps to show the values of matrices.

7.2 Our Weakness

- The determination of the form of f, g, h functions defined in the optimization model may be subjective.
- By simulated annealing, we may get an approximate solution.
- We'd better conduct more quantitative analysis on how specific strategies are connected with the direct changes on the performance. This requires us to collect more research data on a wider range.

References

[1] J. Peña and H. Touchette, "A network theory analysis of football strategies," 06 2012.

[2] P. Cintia, S. Rinzivillo, and L. Pappalardo, "A network-based approach to evaluate the performance of football teams," 09 2015.

[3] N. GÜRSAKAL, F. Yilmaz, H. O. Çobanoğlu, and S. Cagliyor, "Network motifs in football," vol. 20, pp. 263–272, 12 2018.

[4] J. Ramos, R. J. Lopes, and D. Araujo, "What's next in complex networks? capturing the concept of attacking play in invasive team sports," *Sports Medicine*, vol. 48, 09 2017.

[5] J. Castellano, D. Casamichana, and C. Peñas, "The use of match statistics that discriminate between successful and unsuccessful soccer teams," *Journal of human kinetics*, vol. 31, pp. 139–47, 03 2012.

[6] E. Estrada, D. J. Higham, and N. Hatano, "Communicability betweenness in complex networks," *Physica A: Statistical Mechanics and its Applications*, vol. 388, no. 5, pp. 764 – 774, 2009.

[7] T. U. Grund, "Network structure and team performance: The case of english premier league soccer teams," *Social Networks*, vol. 34, no. 4, pp. 682 – 690, 2012.

[8] D. Koutra, A. Ramdas, A. Parikh, and J. Xiang, "Algorithms for graph similarity and subgraph matching," 2011.

[9] J. López Peña, "A Markovian model for association football possession and its outcomes," *arXiv e-prints*, p. arXiv:1403.7993, Mar 2014.

8 Appendix

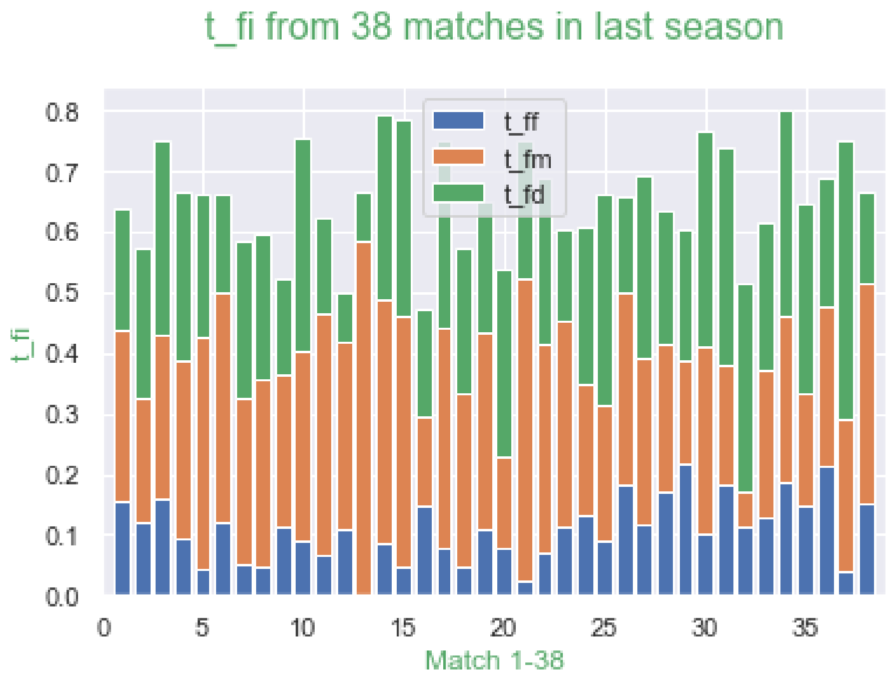


Figure 17: Pass Tendency of Huskies in 38 Matches-Forward Group

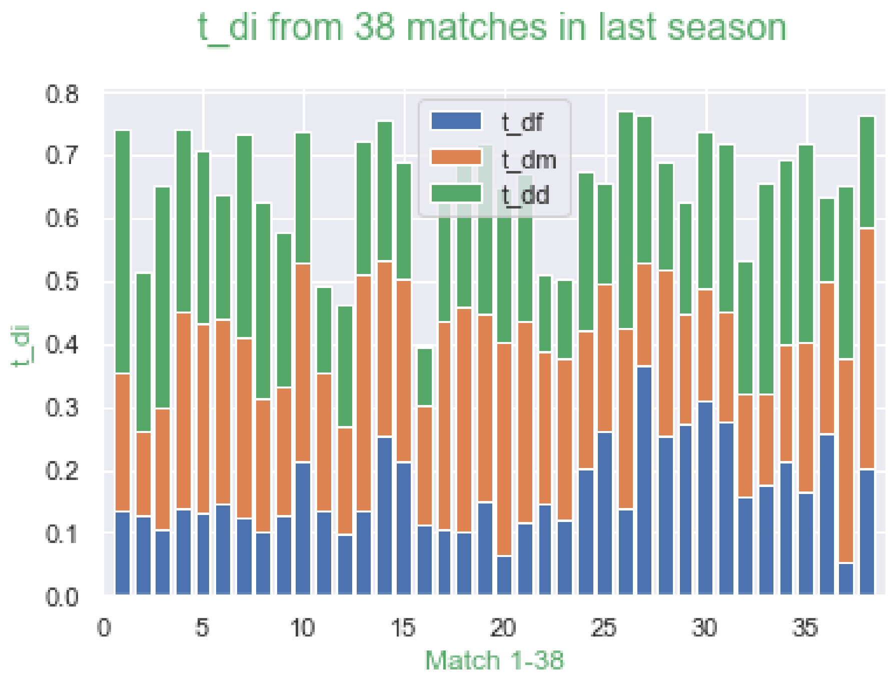


Figure 18: Pass Tendency of Huskies in 38 Matches-Defense Group

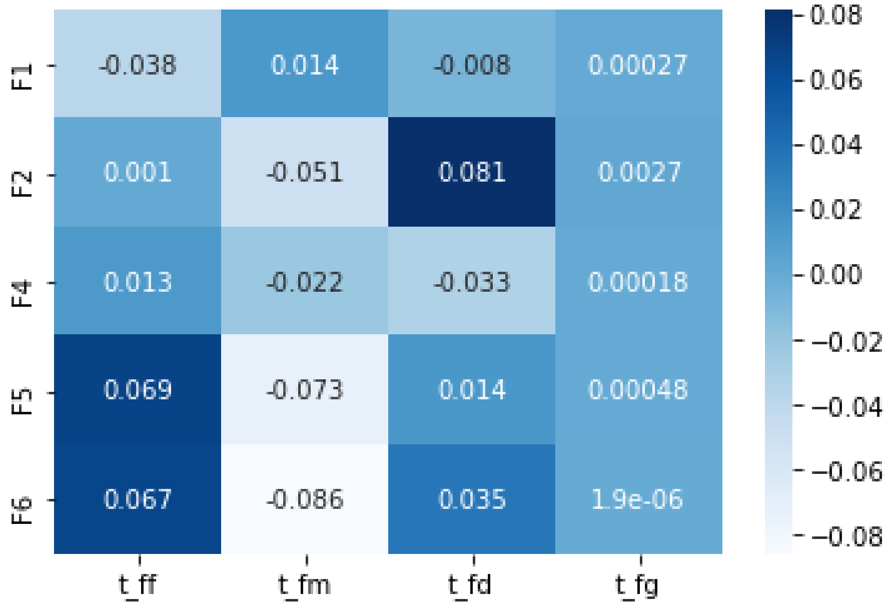


Figure 19: Tendency Difference of Forward Group with/without a Forward Player

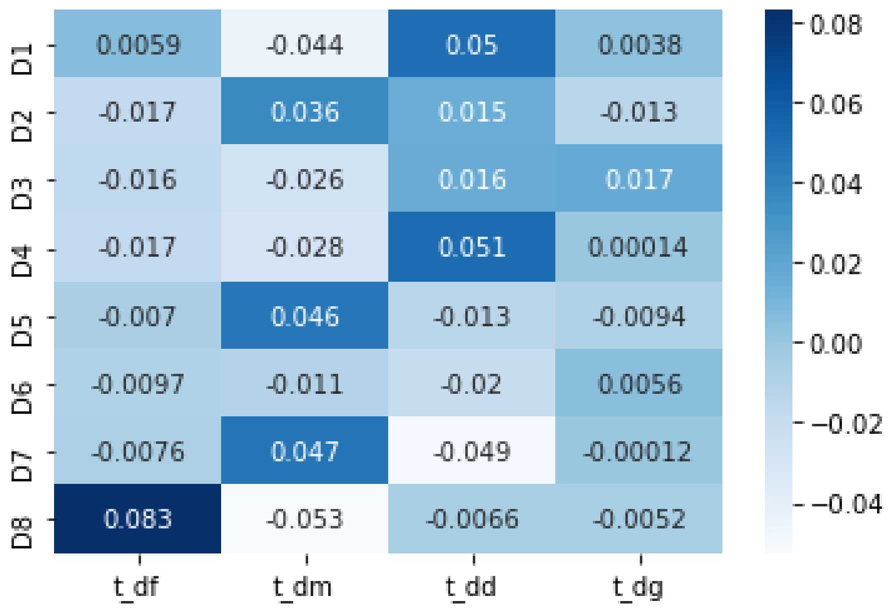


Figure 20: Tendency Difference of Defense Group with/without a Defense Player

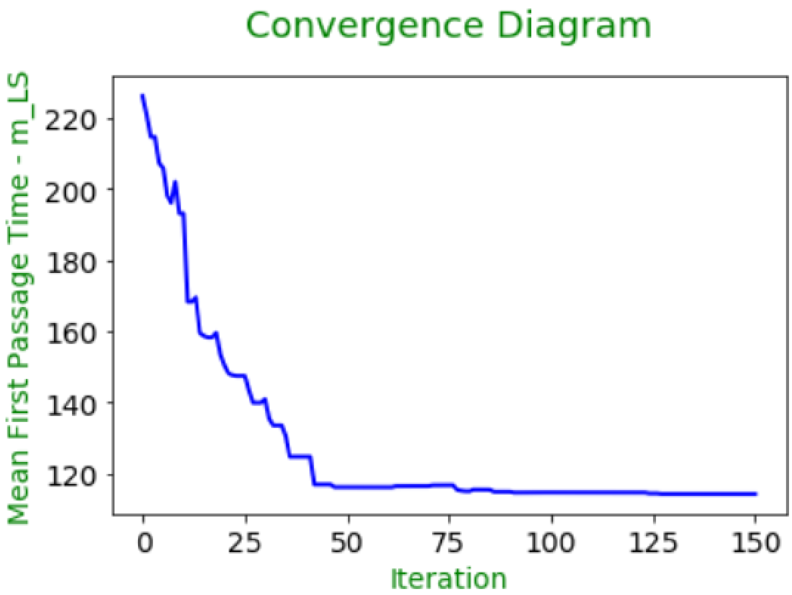


Figure 21: The Convergence Diagram of the Simulated Annealing Algorithm when Trying to Obtain $optP$.