 J. Quant. Anal. Sports 2024; 20(3): 235–250

Research Article

Jonathan Arsenault*, Margaret Cunniff, Eric Tulsky and James Richard Forbes

Spatial roles in hockey special teams

https://doi.org/10.1515/jqas-2023-0019
Received February 28, 2023; accepted March 1, 2024;
published online April 4, 2024

Abstract: Special teams (i.e. power play and penalty kill) situations play an outsized role in determining the outcome of ice hockey games. Yet, quantitative methods for characterizing special teams tactics are limited. This work focuses on team structure and player deployment during in-zone special teams possessions. Leveraging player and puck tracking data from the National Hockey League (NHL), a framework is developed for describing player positioning during 5-on-4 power play and 4-on-5 penalty kill possessions. More specifically, player roles are defined directly from the player tracking data using non-negative matrix factorization, and every player is allocated a unique role at every frame of tracking data by solving a linear assignment problem. Team formations naturally arise through the combination of roles occupied in a frame. Roles that vary on a per-frame basis allow for a fine-grained analysis of team structure. This property of the roles-based representation is used to group together similar power play possessions using latent Dirichlet allocation, a topic modelling technique. The concept of assignments, which remain constant over an entire possession, is also introduced. Assignments provide a more stable measure of player positioning, which may be preferable when assessing deployment over longer periods of time.

Keywords: hockey; player tracking; non-negative matrix factorization; topic modelling

1 Introduction

An ice hockey game is primarily played with five skaters and one goaltender on the ice, which is considered full strength. However, when a team is assessed a penalty, it must temporarily play with fewer than five skaters, causing an imbalance in the number of players on the ice for each team. The team with fewer players on the ice due to being assessed the penalty is said to be on the *penalty kill*, while its opponent is said to be on the *power play*. Collectively, these game states are referred to as *special teams*. During special teams play, man-to-man defense becomes impossible, affording more time and space to the team on the power play. In turn, these teams score at higher rates than when the game is being played at full strength, making special teams situations critical to the game's outcome.

Generally, the objective of the team on the power play is to score a goal while the team on the penalty kill focuses on preventing a goal from being scored. While a team on the penalty kill may seize on offensive opportunities to score a goal of their own, this is a secondary objective. Thus, the team on the power play is referred to as the attacking team and the team on the penalty kill is referred to as the defending team. To generate scoring chances, the attacking team might immediately attempt to score when the opportunity arises due to defensive breakdowns. More often, however, the attacking team will attempt to enter the offensive zone and maintain possession of the puck, leading to what is referred to here as a special teams in-zone *possessions*, or, more succinctly, special teams possessions. During these sequences, both the attacking and defending teams organize themselves in a particular structure, typically predetermined by the coaching staff. Players usually have a general spatial assignment during possessions, meaning they occupy a specific area of the ice.

These spatial assignments have become an integral part of how special teams play is discussed, especially on the power play. Figure 1 shows the instantaneous location of the skaters during a special teams possession labelled with two commonly used spatial assignments: the "point" and the "net front". These can be thought of as spatial *roles* that a player can occupy. The right side of Figure 1 displays heatmaps of these players' locations on the power play over the entire season, demonstrating clear, long-term tendencies in player positioning. However, despite the importance of these spatial roles in hockey special teams and the increasing interest in data analysis in the sport, there is currently

***Corresponding author: Jonathan Arsenault**, Department of Mechanical Engineering, McGill University, Montreal, Canada,
E-mail: jonathan.arsenault@mail.mcgill.ca.
https://orcid.org/0000-0001-5549-8406
Margaret Cunniff and Eric Tulsky, Carolina Hurricanes, Raleigh, USA
James Richard Forbes, Department of Mechanical Engineering, McGill University, Montreal, Canada

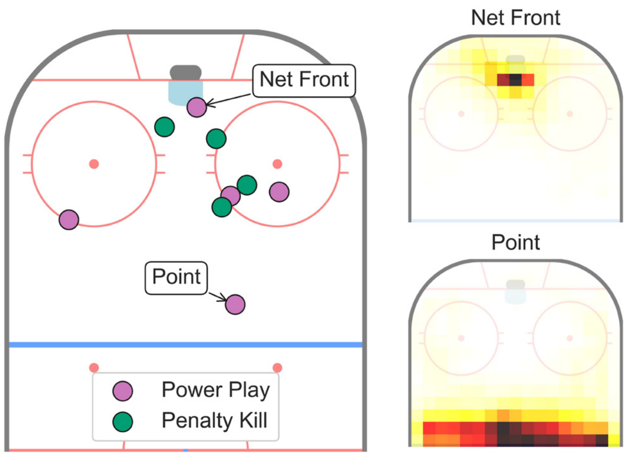


Figure 1: Left: Sample frame of special teams tracking data, annotated with two commonly used roles. Right: Heatmaps of the location of two sample players over the course of a season, demonstrating long-term tendencies in player positioning.

no method for quantifying them, primarily due to the lack of necessary data.

1.1 Data in hockey

A key theme of Nandakumar and Jensen (2019), a historical overview of hockey analytics, is how data availability dictates possible research avenues. The National Hockey League (NHL) publicly releases manually recorded data. The nature of this data, which includes information about shots, player substitutions, giveaways and takeaways, has remained relatively constant since the 2007–2008 season (Kasan 2008). Generally, the available data provides analysts with the identity of each player on the ice throughout the game and information about certain events where a player is performing an action on the puck, referred to here as on-puck events. The quantity of available data improved when private companies began developing computer vision algorithms to automatically detect actions, such as zone entries and passes, from video feeds (Askari et al. 2022; Fani et al. 2017; Tora et al. 2017). While this drastically increased the number of tracked events per game, it still generally only provided information about on-puck events.

To comprehensively quantify player positioning on special teams, it is necessary to know the location of each player at every moment during the game and not just when they are involved in an on-puck event, such as a shot. These on-puck events are relatively rare, and don't provide the location of the players not involved in the event. This desired player location data, referred to as player tracking data, has been collected in professional soccer, basketball, and football leagues for years. In the 2020–2021 season, the NHL began collecting player tracking data of its own, along with puck tracking data. Additionally, work has been done to extract player (Vats et al. 2023) and puck (Vats et al. 2021) tracking data from video feeds. In this work, tracking data is leveraged to quantify player positioning on special teams.

1.2 Prior work

1.2.1 Special teams

Prior research into special teams has primarily focused on measuring team and player performance. Regression-based adjusted plus-minus statistics that isolate the impact a player has on team-level power play and penalty kill performance (Macdonald 2012) represent the state-of-the-art solution. These statistics use a team's performance while a player is on the ice to measure their impact. They are commonly cited in the hockey analytics community as various versions of these statistics are publicly available (McCurdy 2022). While these models can be valuable, they struggle to illuminate why teams and players perform the way they do. Previous work on the tactical choices underpinning successful power plays has focused on the superiority of the four-forward power play (Cane 2017b), the importance of structure on the power play (Cane 2017a; Parnass 2016a), and the importance of generating successful zone entries (Parnass 2016b). In Parnass (2016a,b), the author manually collected data to augment the NHL's data. The richer data sources, such as the event data collected from video feeds and the player and puck tracking data, are proprietary. Thus, the quantity of publicly available research using this data is limited. Event data has been used, for example, to develop novel player evaluation metrics (Liu and Schulte 2018; Schulte et al. 2017). Research using tracking data been mostly focused on passing, in all game situations (Radke et al. 2021; Radke et al. 2022) and specifically on special teams (Ritchie et al. 2022).

1.2.2 Tracking data

In the last decade, tracking data has become available for most major sports. Analysts have found numerous applications for this data, as summarized in Gudmundsson and Horton (2017). Valuing individual player actions, such as soccer passes (Anzer and Bauer 2022), is a particularly well-researched topic that relies on tracking data. Valuing individual actions naturally leads to models that compute the expected value of possessions, like the ones developed for basketball (Cervone et al. 2016), soccer (Fernández et al. 2021), and football (Yurko et al. 2020). A separate area of research is mining tracking data for repeated patterns, including work identifying receiver routes (Chu et al. 2020; Hochstedler and Gagnon 2017) and pass coverage (Dutta et al. 2020) in the National Football League, repeated basketball actions (Miller and Bornn 2017), off-ball runs in soccer (Gregory 2019), and serve returns in tennis (Kovalchik and Albert 2022).

More closely related to the problem at hand is the literature surrounding role-based representations of player positioning. Roles were first used to align field hockey player trajectories in Lucey et al. (2013) using a manually-labelled dataset of player roles. This approach was then applied in basketball (Lucey et al. 2014) to uncover patterns in the lead-up to three-point shots and in soccer to find patterns in how teams score goals (Wei et al. 2013) and to investigate home-field advantage (Bialkowski et al. 2014). In Bialkowski et al. (2016), an alternate method is proposed where the role definitions are not known *a priori*. Instead, formations, role definitions and role assignments are simultaneously computed using minimum entropy data partitioning. Shaw and Glickman (2019) propose a different framework for analyzing team structure in soccer, where roles are bypassed altogether. Formations are characterized solely using the relative location of teammates. Distinct types of formations are then found by clustering the observed formations.

1.3 Contribution

Developing similar methods for hockey special teams presents a unique set of challenges that require new approaches. In soccer, for example, the concept of a formation is very well-defined and teams are expected to maintain these formations for extended periods of time. The developed techniques have leveraged this, aggregating player positioning over several minutes (Bialkowski et al. 2016; Shaw and Glickman 2019), assuming the team structure remains relatively constant over that time span. This is not directly applicable to hockey, as possessions are very short, and team formations may constantly be in flux due to the fast-paced and unpredictable nature of the game. The prior work in basketball (Lucey et al. 2014) may hold greater relevance, as player roles are assigned at each frame. However, no attempt is made to characterize formations, as they hold no importance within basketball. The roles were also defined manually, while a more data-driven approach is desired. Lastly, in all previous work, the number of roles is equal to the number of players in the situation of interest. A detailed analysis of hockey special teams using player roles may require more roles than the players on the ice, meaning more than five roles on the power play and more than four roles on the penalty kill. This would provide a more flexible and detailed framework for characterizing player positioning.

This paper introduces a novel method for describing player positioning during special teams possessions. This approach employs non-negative matrix factorization to characterize player roles directly from data, resulting in interpretable roles that are then uniquely allocated at each frame to yield a low-dimensional encoding of player positioning. To achieve smoother role allocations over the course of a possession, the cost matrix used for role allocation is computed in a way that allocates players to roles they have more frequently been observed. As a by-product of this role-based representation, formations arise, which offer insight into the overall team structure during special teams. Two applications are then presented that demonstrate how this dataset can be leveraged to reveal insights into special teams play. The first application groups similar power play possessions, while the second analyzes large-scale player deployment by aggregating a player's role occupation over a possession to define player assignments.

2 Data preparation

2.1 Tracking dataset

The dataset contains 956 games of tracking data from the 2021–2022 NHL season, shared with the authors by the NHL. All 32 NHL teams are represented, with the number of games per team ranging from 43 to 68. Simultaneous measurements of the locations of players on the ice are available at 10 Hz, or 10 frames per second. The identity of the team in possession of the puck, available as part of the tracking dataset, is also used. The puck may be in possession of either team, but it may also be in a third state, where the team in possession is ambiguous. In this state, the possession is deemed to be *contested*.

2.2 Extracting special teams in-zone possessions

This analysis is limited to 5-on-4 power plays and 4-on-5 penalty kills, as they are, by a wide margin, the most frequent special teams situations. This excludes 5-on-3 and 4-on-3 power plays and the corresponding penalty kills, relatively rare situations with different geometries that would require a separate analysis. Thus, 5-on-4 and 4-on-5 are referred to when using the abbreviations PP for the power play and PK for the penalty kill, respectively. This work is concerned with game segments in which the attacking team has gained the offensive zone with control of the puck and the players have arranged themselves in their desired structure. These segments are special teams in-zone possessions, more succinctly referred to as *possessions* here. A possession is defined as an uninterrupted segment of the game during which

1. five offensive skaters are in their offensive zone,
2. four defensive skaters are in their defensive zone,
3. the segment lasts more than 5 s, and
4. the possession of the puck is maintained by the attacking team.

To allow time for the teams to organize, the first 2 s of each possession are discarded, meaning no possession lasts less than 3 s. The attacking team is said to have maintained possession of the puck if the defending team never possesses the puck and puck possession is contested for less than 30 % of the duration of the possession. These parameters were chosen to ensure a balance between maximizing the number of possessions in the dataset while capturing only possessions with observable structure. From the 956 games in the tracking dataset, 9347 possessions were extracted for an average of 9.78 possessions per game. A breakdown of the dataset on a per-team basis is provided in Table 1. The entire possessions dataset consists of 869,534 frames of tracking data.

3 Defining and allocating roles

This section details the procedure used to define roles from the player tracking data and to allocate each player a unique role at each frame.

3.1 Defining roles using non-negative matrix factorization

Non-negative matrix factorization (NMF) is a dimensionality reduction technique useful in uncovering patterns in tracking data, as extensively detailed in Kovalchik (2023). Of note, NMF has previously been used in basketball to obtain low-dimensional representation of shooting (Miller et al. 2014) and player positioning (Cervone et al. 2016) tendencies. The latter application closely aligns with the application presented in this work.

Table 1: Per-team summary statistics of the possession dataset, extracted from 956 games of data from the 2021–2022 NHL season.

	Max.	Min.	Mean	Std. dev.
Total games	68.0	43.0	59.2	5.5
Total poss. (PP)	424.0	206.0	292.1	45.4
Poss. per game (PP)	6.2	3.9	4.9	0.5
Mean poss. duration (PP) [s]	10.7	7.6	9.1	0.8
Total poss. (PK)	387.0	216.0	292.1	47.2
Poss. per game (PK)	6.7	3.6	4.9	0.7
Mean poss. duration (PK) [s]	10.7	7.6	9.2	0.7

3.1.1 Methodology

The playing area of interest is divided into $L = 249$ bins. Each bin is 5 ft by 5 ft. This excludes bins that are fully outside the ice surface. For each player $n = 1, \ldots, N$ in the dataset, a *player count matrix* is created by counting the number of times the player is observed in bin ℓ, $\ell = 1, \ldots, L$. These player count matrices are normalized to unit sum to remove any effect caused by variations in playing time, referred to as time-on-ice (TOI). These new matrices are called *player occupancy matrices*. The *occupancy matrix*, created by flattening and stacking the player occupancy matrices, is denoted $\mathbf{X} \in \mathbb{R}_{+}^{N\times L}$, where $X_{n,\ell}$ is the estimated proportion of time player n spent in bin ℓ. This value is only an estimate as the raw counts have been smoothed using a Gaussian blur to reduce the effect of noise caused by both the small sample size and the arbitrary binning of the playing area. This process is described in more detail in Appendix A.

Using NMF, the occupancy matrix is approximated by the product of two matrices, $\mathbf{X} \approx \mathbf{WB}$, where $\mathbf{W} \in \mathbb{R}_{+}^{N\times K}$ is the *weight matrix* and $\mathbf{B} \in \mathbb{R}_{+}^{K\times L}$ is the *basis matrix*. The rows of $\mathbf{B}$ provide non-negative bases for the occupancy matrix, while the rows of $\mathbf{W}$ provide the basis loadings for each player. In other words, $\mathbf{B}$ defines the roles and $\mathbf{W}$ contains the contribution of each role to each player's occupancy matrix. The values of $\mathbf{B}$ and $\mathbf{W}$ are found by minimizing a Kullback-Leibler-type divergence,

$$e_{\mathrm{KL}}(\mathbf{X}, \tilde{\mathbf{X}}) = \sum_{ij} X_{ij} \log \frac{X_{ij}}{\tilde{X}_{ij}} - X_{ij} + \tilde{X}_{ij} \tag{1}$$

that is minimized using the multiplicative update method (Lee and Seung 2000). The reconstructed matrix is denoted $\tilde{\mathbf{X}} = \mathbf{WB}$.

3.1.2 Results

Players with less than 180 s of TOI in the possession dataset are omitted. As a result, 417 players for the PP and 419 players for the PK are considered. This threshold roughly corresponds to 13 players per team, which is approximately two 5-player units on the PP and three 4-player units on the PK, including supplemental players accounting for changing units over the course of the season. Thus, most players who have spent some time as either a PP or PK regular are included. The most critical parameter to be selected is the number of roles K. Selecting a large number of roles would minimize the reconstruction error (1), but would result in highly-granular roles that are difficult to interpret and of limited value for analysis. Instead, the number of roles are constrained to be in the range of 6–12, to be consistent with the expectations of domain experts and how they communicate special teams tactics. The precise value is selected by visual inspection of the results. This lead to the selection of $K = 9$ for the PP and $K = 6$ for the PK.

The role definitions for both the PP and PK are shown in Figures 2 and 3, respectively. Note that each individual role definition has its own color scale. Each role has been given a meaningful label, reinforcing their interpretability. The PP role definitions roughly divide the offensive zone into nine regions. Three roles are reserved for players playing the point, corresponding to regions near the blue line. Two role definitions align with the half-wall, an area of the ice close to the boards on both the left and right sides of the offensive zone. Three roles are near the goal line, and the remaining role is centrally located in the offensive zone, an area known as the bumper spot. These definitions are sparse and well-defined, which is a desired feature. In contrast, the PK role definitions are less well-defined, particularly the High role. This role is a residual role, and its occupancy map includes a significant portion of the playing area. This role captures the fluidity of forward positioning on the penalty kill, as forwards are constantly adjusting their positioning based on the location of the puck. The role definitions obtained from this procedure are purely spatial, but there are functional differences in the roles, as the location a player occupies in space affects the actions they can take.

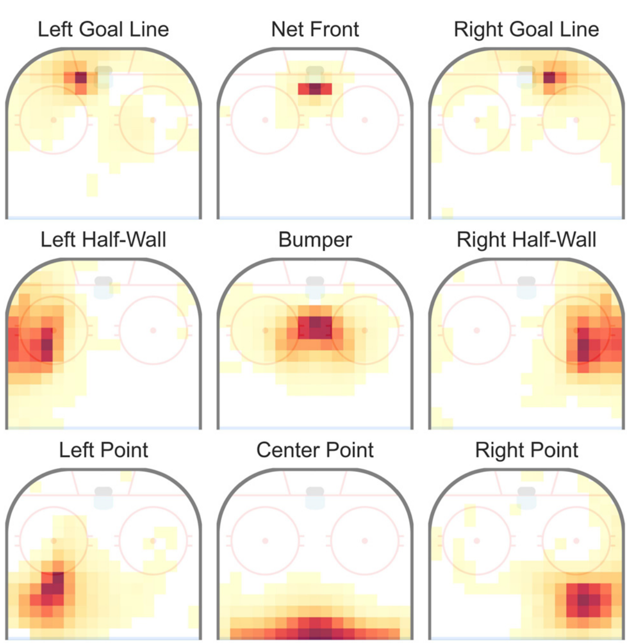


Figure 2: Power play role definitions. The color scale for each role is unique. The role definitions are sparse and well-defined. They symmetrically divide the offensive zone into nine regions.

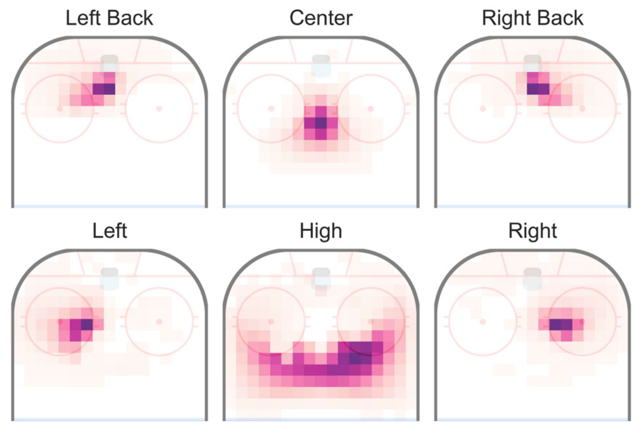


Figure 3: Penalty kill role definitions. The color scale for each role is unique. The role definitions are less well-defined than the PP role definitions. The high role is a residual role, capturing the fluidity of forward positioning on the PK.

3.2 Allocating roles

The role definitions provide an aggregate view of player positioning. Using the role definitions, a fine-grained analysis is possible by determining the role each player occupies at each frame of tracking data.

3.2.1 Methodology

Let P be the number of players on the ice in a frame of tracking data. For the PP $P = 5$ and for the PK $P = 4$. Let $\mathbf{C}$ be a $K \times P$ cost matrix, where each element $C_{k,p}$ represents the cost of allocating player p to role k for $p = 1, \ldots, P$ and $k = 1, \ldots, K$. Let $\mathbf{A}$ be a $K \times P$ allocation matrix, where

$$A_{k,p} = \begin{cases} 1 & \text{if player } p \text{ occupies role } k, \\ 0 & \text{otherwise.} \end{cases} \tag{2}$$

To allocate a role to each player, the matrix $\mathbf{A}$ that maximizes the total cost

$$C_{\text{total}} = \sum_{k=1}^{K} \sum_{p=1}^{P} C_{k,p} A_{k,p} \tag{3}$$

is found. Solving for $\mathbf{A}$ corresponds to solving a linear assignment problem, which can be done using the Hungarian algorithm (Kuhn 1955), ensuring that each role is allocated to at most one player at each frame. In practice, both $\mathbf{C}$ and $\mathbf{A}$ must be zero-padded such that they are of dimension $K \times K$. This simply assigns fictional players to the unallocated roles.

The elements of the cost matrix $C_{k,p}$ are determined by computing $\Pr(k_p = k|\ell_p = \ell)$, the probability that the role occupied by player p, k_p, is role k conditioned on the player's observed location, ℓ_p, being location ℓ. Using Bayes' rule,

$$\Pr(k_p = k|\ell_p = \ell) = \frac{\Pr(\ell_p = \ell|k_p = k)\Pr(k_p = k)}{\Pr(\ell_p = \ell)}, \quad (4)$$

$$= \frac{\Pr(\ell_p = \ell|k_p = k)\Pr(k_p = k)}{\sum_{k'=1}^{K} \Pr\left(\ell_p = \ell|k_p = k'\right)\Pr\left(k_p = k'\right)}. \quad (5)$$

The information in $\mathbf{W}$ and $\mathbf{B}$ can be used to compute the prior and likelihood terms of (5).

Recall that the nth row of $\mathbf{W}$, denoted $\mathbf{w}_n$, corresponds to the nth player's basis loadings, meaning this row contains information about the amount of time this player spends in each role. Normalizing each row to unit sum,

$$\bar{\mathbf{w}}_n = \frac{\mathbf{w}_n}{\sum_{k'=1}^{K} W_{n,k'}}, \quad (6)$$

yields $\bar{\mathbf{W}}$, the normalized player weights. Let $\bar{\mathbf{w}}_{n_p}$ be the normalized weights for player p, where n_p is the row index of player p in the weight matrix. The role occupied by player p at frame i, $k_{p,i}$ can be modelled as $k_{p,i} \sim \text{Multinomial}(\bar{\mathbf{w}}_{n_p})$. Thus, $\Pr(k_{p,i} = k) = \bar{W}_{n_p,k}$.

Similarly, define the normalized basis matrix $\bar{\mathbf{B}}$, where each row is normalized to unit sum,

$$\bar{\mathbf{b}}_k = \frac{\mathbf{b}_k}{\sum_{\ell'=1}^{L} B_{k,\ell'}}. \quad (7)$$

The location of player p, ℓ_p, given k_p is then modelled as $\ell_p|k_p \sim \text{Multinomial}(\bar{\mathbf{b}}_{k_p})$. Thus, $\Pr(\ell_p = \ell|k_p = k) = \bar{b}_{k,\ell}$.

Combining these results, each element of the cost matrix is

$$C_{k,p} = \frac{\bar{B}_{k,\ell}\bar{W}_{n_p,k}}{\sum_{k'}\bar{B}_{k',\ell}\bar{W}_{n_p,k'}}. \quad (8)$$

It would be possible to compute the cost using solely the role definitions and each player's location. However, incorporating information about a player's historical role occupation tendencies through the weight matrix stabilizes the results. Players are constantly in motion and will undoubtedly move across role boundaries without a meaningful change in role occurring. By allocating players roles that they have more often been previously observed in, these trivial and often short-lived role changes are avoided. It does happen, however, that players do not have weights, either because they were not present in the original dataset or they did not meet the TOI requirement. In this case, a uniform prior is used, meaning $W_{n_p,k} = 1/K$.

3.2.2 Results

As there are more roles than players on the ice, not all roles are allocated at each frame. Figure 4 shows the percentage of frames that each role is allocated in. On the PP, five roles (Center Point, Right Half-Wall, Bumper, Left Half-Wall and Net Front) are allocated in over 60 % of frames, while the remaining four roles are allocated in fewer than 40 % of frames. On the PK, three roles are allocated in over 80 % frames.

Figure 5 shows the percentage of frames a role was allocated to a left-shooting player, along with the general location of each role. A role's general location is a qualitative description of the area of the ice that the role definition occupies. For example, the Left Point, Left Half-Wall, and Left Goal Line roles are generally on the left side of the ice, and are differentiated by their distance from the goal. On the PP, players tend to play their off-side, meaning a left-shooting player is more likely to play on the right side of the PP, which aligns with a well-understood concept in hockey. By playing on their off-side, players have their stick pointing towards the middle of the ice, making it easier

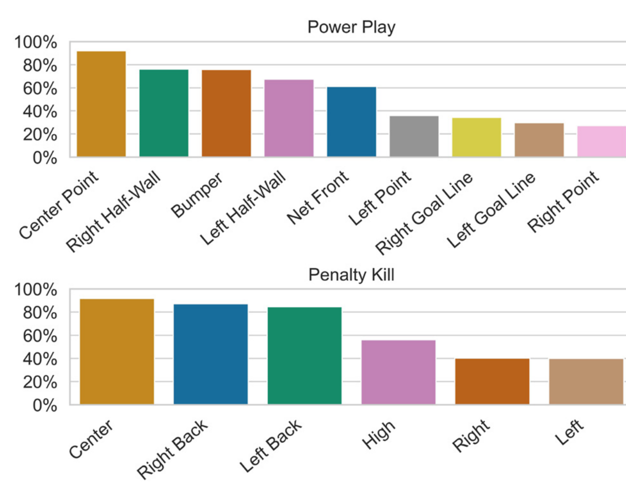


Figure 4: Percentage of frames each role is allocated in, for both the PP and PK, illustrating the tendency for certain roles to be allocated more often than others.

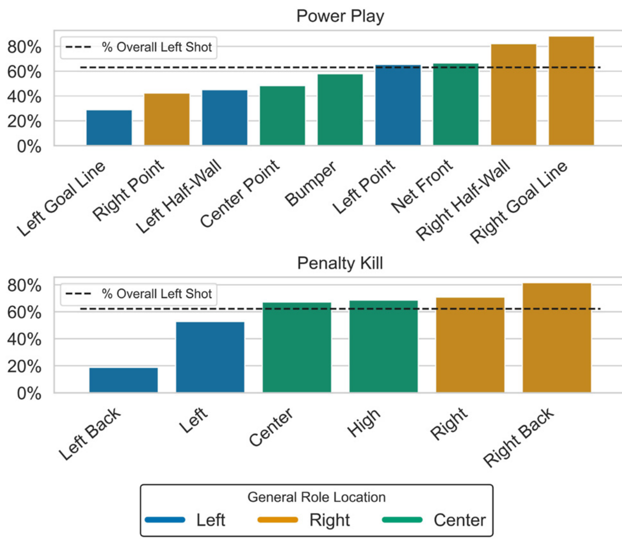


Figure 5: Percentage of frames a role is allocated to a left-shooting player. On the PP, players tend to play their off-side more frequently, while on the PK, players tend to play their strong-side.

for them to shoot upon receiving a pass, and from a more central location. Consequently, off-side players on the power play are more dangerous shooting threats. Conversely, On the PK, left-shooting players are more likely to play on the left side and right-shooting players are more likely to play on the right side. Potential shooting advantages gained by being on the off-side are less important when defending.

Figure 6 shows the allocated roles for a sample player over the course of a possession on both the PP and PK. As the players move around the offensive zone, their role changes. The frequency of role changes is shown in Figure 7. On the PP, certain pairs of spatially adjacent roles are observed to engage in role changes much more often than others. For example, players in the Center Point role do not frequently change roles during possessions. Meanwhile, on the PK, most combinations of role changes are frequently observed, except for interchanges between the Center role and either Back role.

4 Formations

Formations are a common term used in team sports like soccer to describe player positioning relative to each other. In this work, formations refer to the combination of roles allocated to a team's players at any given time during a possession. There are

$$\binom{K}{P} = \frac{P!}{K!(P-K)!} \tag{9}$$

possible formations, with 126 possible formations when using nine PP roles and 15 possible formations when using six PK roles. Since player roles can change frequently during a possession, a team's formation is a dynamic measure of its structure, constantly changing as the game progresses. Thus, teams are referred to as being *observed* in a formation. It may be convenient to say that a team *uses* a formation, but this implies that some tactical choice is made at every

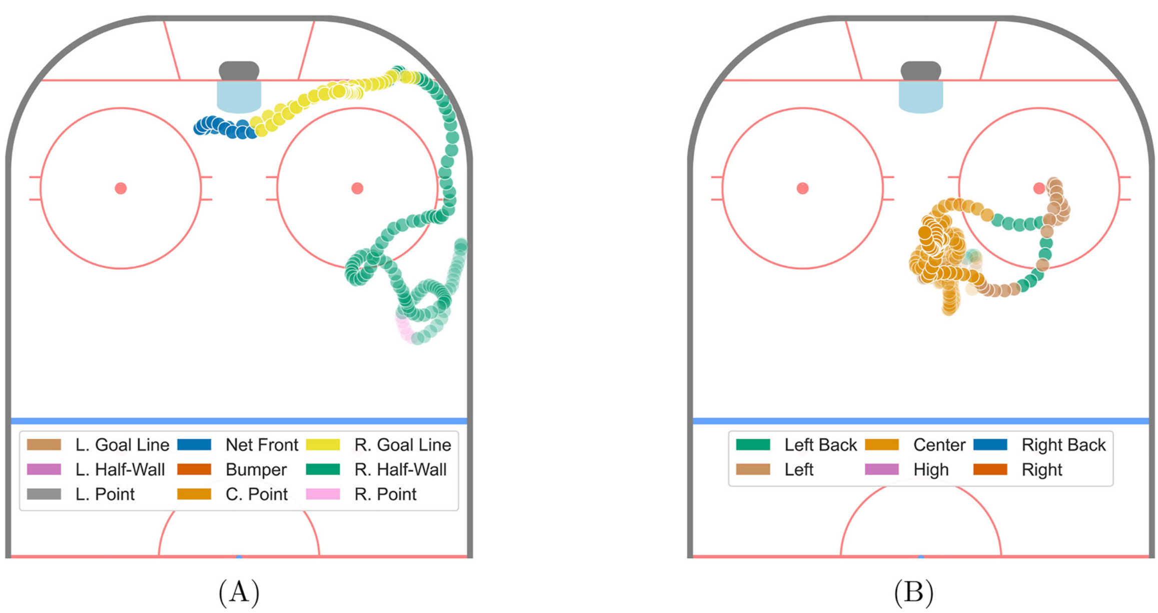
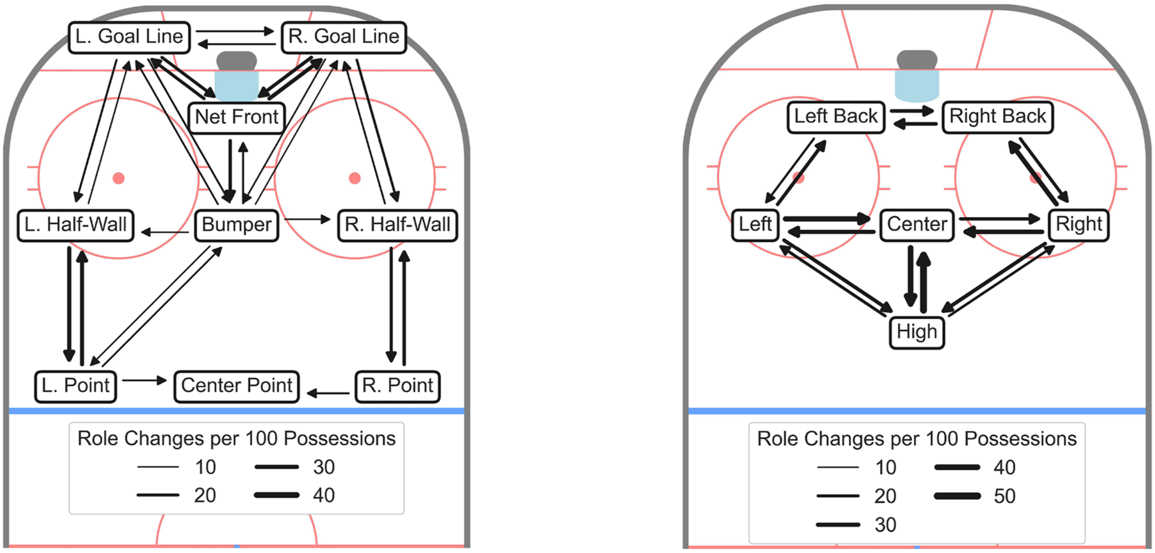


Figure 6: Role allocation over the course of a possession for a sample player on both the PP and PK. The most transparent markers indicate the beginning of the possession and the transparency decreases as time progresses. As the possessions progress and the player moves around the ice, their role changes. (A) Power play. (B) Penalty kill.



(A) (B)

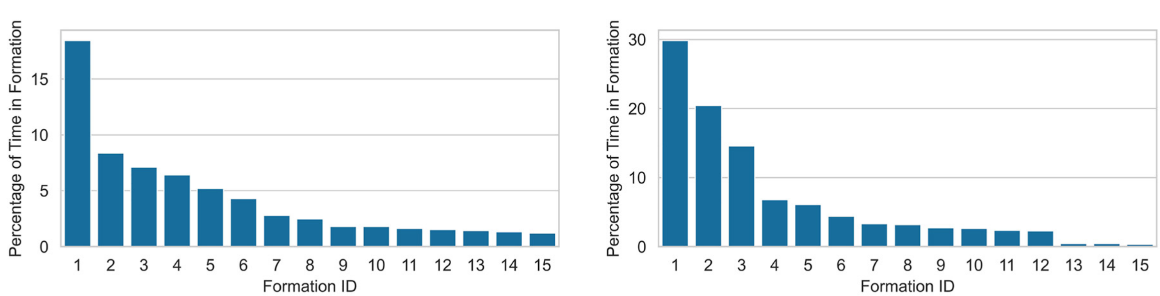
Figure 7: Rate of role changes on both the PP and PK. Only role changes observed at a rate of more than 10 and 15 changes per 100 possessions are displayed for the PP and PK, respectively. The size of the arrow is proportional to the rate of role changes. (A) Power play. (B) Penalty kill.

instant to be in a given formation. Continuous player motion makes it unlikely that this is the case.

Nonetheless, formations do provide insight into high-level team tactics, especially when the data is aggregated over many possessions. To uniquely distinguish formations, they are identified using an integer ID, such as PP1 and PK1. Formations are ranked by how frequently they are observed. Their integer ID refers to their rank. For example, PP1 refers to the most often observed PP formation, PP2 the second most often observed formation, and so on. To get a global view of how frequently each formation is observed, the percentage of frames that they are each observed in is computed and displayed in Figure 8. On the PP, only one formation is observed in over 10 % of frames, while only 19 of 126 possible formations are observed in over 1 % of frames. On the PK, fewer formations means individual formations are more frequently observed, but three formations still appear much more often than the others.

4.1 Power play formations

The three most commonly observed formations on the PP are shown in Figure 9, with additional formations shown in Appendix B. Formations are visualized using kernel density estimation to model the distribution of the location of the players allocated to each role in the formation. Formation PP1, observed in 18.4 % of frames, is referred to as the base PP formation, as most of the other frequently observed formations are slight deviations from it. Formations PP2 and PP3 in Figure 9 are examples of this. Formation PP2 is identical to Formation PP1 with the exception of the Net Front player moving to the Right Goal Line. Formation PP3



(A) (B)

Figure 8: Percentage of time the top 15 PP and PK formations are observed. One formation is especially prominent on the PP, compared to three on the PK. (A) Power play. (B) Penalty kill.

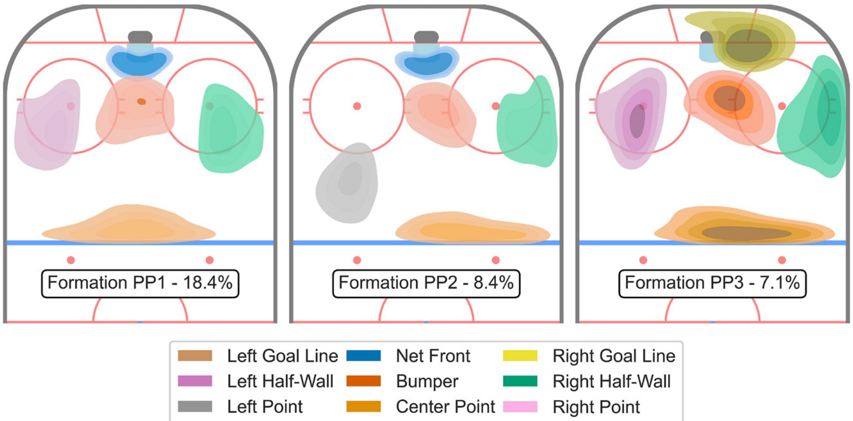


Figure 9: Three most commonly observed formations on the power play with percentage of time each formation was observed. Most formations are slight deviations from the base formation, PP1, including PP2 and PP3, shown here.

Table 2: Number of times each PP formation is the first-, second-, and third-most used formation for a team. The base formation is the most observed formation for 25 of the 32 teams in the dataset, indicating some degree of homogeneity in power play tactics.

Usage rank	Formation ID								
	PP1	PP2	PP3	PP4	PP5	PP6	PP7	PP8	PP12
First	25	4	0	1	0	1	1	0	0
Second	6	6	9	5	6	0	0	0	0
Third	0	10	4	9	6	0	0	2	1

similarly sees the Left Half Wall player from Formation PP1 moving to the Left Point. Not shown is Formation PP4, where the Net Front player moves to the Left Goal Line. Formation PP2 is likely more common than Formation PP4 because a majority of NHL players are left-shooting, meaning playing on the Right Goal Line makes them a better passing option.

Table 2 provides a summary of which formations are the first-, second-, and third-most often observed formations for each team in the dataset. The base formation is the most observed formation for 25 of the 32 teams in the dataset. By this specific measure, power play tactics may seem similar across teams. However, the second and third most used formations varies widely across teams, as displayed in Table 2. This indicates that there indeed exists stylistic differences across teams. For example, a team that prefers Formation PP2 as its first- or second-most often observed formation may rely on shots from the Left Point to generate scoring opportunities, while a team preferring Formation PP3 may attempt to score by moving the puck to the Right Goal Line and creating chances through shorter-distance shots.

4.2 Penalty kill formations

Figure 10 shows the three most commonly observed PK formations, with additional formations relegated to Appendix B. Once again, a base PK formation, Formation PK1, is observed a majority of the time. Formations PK2 and PK3 are slight deviations from the base formation, where the player in the High role is replaced by a player either in the Left or Right role, respectively. Interestingly, Formation PK2, where a player is in the Left role, is observed 4.9 % more often than Formation PK3, where a player is in the Right role. The player in the Left role is positioned to counter a shooting threat coming from the player in the Left Point role on the PP. This combination, observed in Formations PP2 and PK2, is more common than the similar combination on the right, observed in Formations PP3 and PK3. This symmetry illustrates how the positioning of teams on the PK is a reaction to the positioning of teams on the PP. As was done for the PP, the preferred formations by team are shown in Table 3. There is more homogeneity among teams on the PK. This is to be expected, as there are fewer possible formations a team can occupy. Furthermore, as previously mentioned, a team on the PK is often reacting to the positioning of the team on the PP. What Table 3 reveals, then, is how a team would approach an average opponent.

5 Applications

5.1 Grouping similar power play possessions

A relevant application of this framework is to organize special teams possessions. Grouping possessions that are by some metric similar may be valuable to coaches as it may allow them to gain a better understanding of the tactical

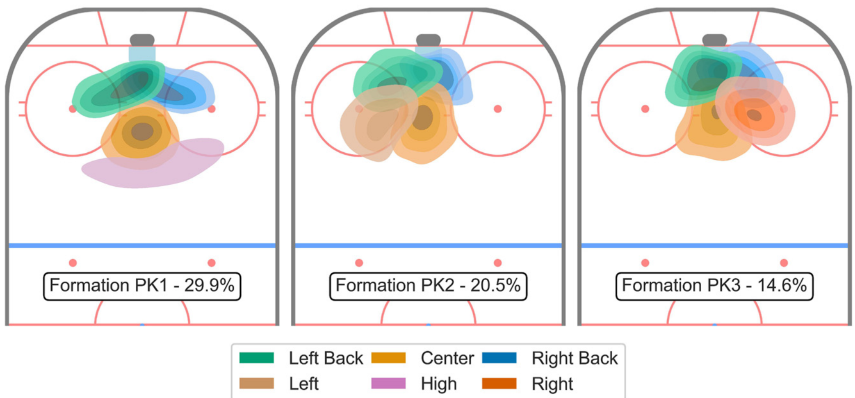


Figure 10: Three most commonly observed formations on the penalty kill with percentage of time each formation was observed. Formations PK2 and PK3 are slight deviations from the base formation, PK1.

Table 3: Number of times each PK formation is the first-, second-, and third-most used formation for a team. These results can be interpreted as how a team would approach an average opponent, given that the team on the PK is often reacting to the positioning of the team on the PP.

Usage rank	Formation ID							
	PK1	PK2	PK3	PK4	PK5	PK6	PK7	PK8
First	25	6	0	0	0	0	1	0
Second	3	20	4	2	0	1	0	2
Third	2	3	20	4	1	1	1	0

choices a team makes in certain scenarios. To organize the special teams possessions, topic modelling techniques are used. Topic modelling is type of statistical model for discovering latent topics in a collection of documents. Topic modelling has previously been used to group similar basketball possessions (Miller and Bornn 2017), and can be adapted to organize special teams possessions in hockey.

5.1.1 Methodology

As in Miller and Bornn (2017), latent Dirichlet allocation (LDA) (Blei et al. 2003), a specific topic modelling technique, is employed to group similar special teams possessions. In natural language processing, LDA is commonly used to identify themes across a collection of documents and classify them accordingly. Topics are defined as collections of words that share a theme. The end result of LDA is the distribution of words in each topic and the distribution of topics across all documents. In this application, each special teams possession is treated as a separate document. The results from the role allocation procedure are used to generate a set of words that describe the possession, which is collectively referred to as the vocabulary.

When LDA is used in natural language processing, a common approach is to represent documents with feature vectors that contain the frequency of each word in the vocabulary within the document. There are many ways of defining possible vocabularies for special teams possessions. In Miller and Bornn (2017), the vocabulary is made up of action-action pairs. A similar concept is implemented here, where the vocabulary is made up of role-role pairs, and the feature vector consists of the amount of times each role-role pair is observed during a possession. By using this input, LDA identifies spatially similar topics where players generally occupy the same areas of the ice. To further categorize the possessions by deployment, roles are distinguished by the handedness of the player occupying the role, resulting in four possible role-role pairs for each initial pair. Incorporating handedness enables discrimination between spatially similar possessions that may differ strategically, meaning they may correspond to a different set of probable actions. To summarize, each possession $m = 1, \ldots, M$ in the dataset is represented as a matrix $\mathbf{y}^{(m)}$ of dimension $1 \times 2K(2K-2)$, where

$$
\begin{aligned}
y_i^{(m)} &= \text{\# of times role} - \text{role pair } i \\
&= \left(r_1', r_2'\right) \text{ is observed in possession } m,
\end{aligned}
\tag{10}
$$

K is the number of roles and $\left(r_1', r_2'\right)$ is a role-role pair with handedness.

5.1.2 Results

LDA is applied to the PP possession dataset using 50 topics. PK possessions are omitted for brevity. The number of topics is selected using a cross-validation procedure, using the log likelihood on the held-out dataset as the evaluation metric. This procedure revealed that 50 topics provided

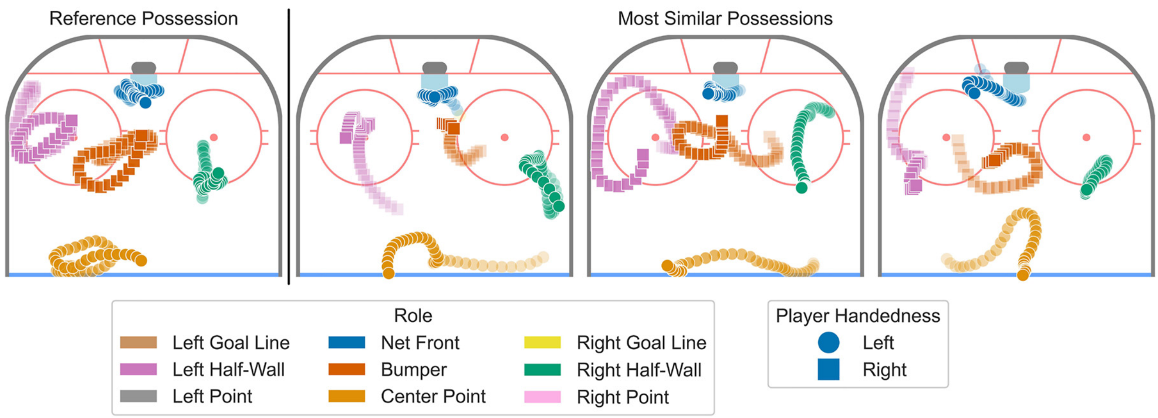


Figure 11: Reference possession and the three possessions most similar to it. These possessions all see the same five roles occupied by a unique player for the duration of the possession. Furthermore, the handedness of the players in each role is the same across possessions.

a reasonable compromise between model complexity and performance. Figure 11 shows a reference possession along with the three most similar possessions. Similarity is measured by comparing the probability that each possession belongs to each topic. The distance metric used is cosine similarity. These possessions are similar in the sense that they have high probabilities of belonging to Topic 19, which is defined by having left-shooting players occupying the Center Point, Right Half-Wall and Net Front roles and right-shooting players occupying the Left Half-Wall and Bumper roles. By most measures, an analyst would characterize these possessions as similar. In all four possessions displayed in Figure 11 the base formation is maintained for the duration of the possession, no role changes occur, and the handedness of the players in each role is the same. This example demonstrates that the role-based representation can successfully be used to group similar possessions, which, when applied on a large scale, could lead to improved tools for assessing PP tactics and performance.

5.2 Player assignments

The roles discussed so far are useful for analyzing player positioning and its evolution during a possession. However, continuous player motion can result in frequent role changes that may not necessarily imply a change in strategy. For certain applications, such as analyzing player deployment over longer time periods, a more stable method of describing player positioning may be preferred. To increase stability, assume that, for a given possession, players may be assigned to a specific area of the ice. This leads to new role definitions, termed *assignments*, that remain constant for the duration of a possession. To determine a player's assignment, the player role distributions during each possession are clustered using a K-means algorithm. This yields a set of assignment definitions, which are a weighted combination of the role definitions from Section 3.1.2.

As the desired output is an assignment per player per possession, each player's observed role occupation distribution per possession is computed and used as the input to the K-means algorithm. Specifically, let M be the number of possessions in the dataset and P be the number of players on the ice in the game situation of interest, meaning $P = 5$ for the power play and $P = 4$ for the penalty kill. Let $\mathcal{X} = (\mathbf{x}_{11}, \ldots, \mathbf{x}_{1P}, \ldots, \mathbf{x}_{MP})$ be the set of observations, where $\mathbf{x}_{mp} = \left[x_{mp,1}, \ldots, x_{mp,K}\right]^{\mathrm{T}}$ and $x_{mp,k}$ is the proportion of time player p spent in role k during possession m. This yields a total of MP observations, one observation per player per possession, as required.

The observations are clustered using manually initialized K-means. The number of clusters is set to $K = P$, the number of players on the ice. This allows for the possibility that each player has a unique assignment, a scenario which would yield the most interpretable results. In practice, this is not always the case, as will be discussed. The resulting cluster centers represent the assignment definitions for the PP and PK, as shown in Figure 12. These cluster centers can be interpreted as the average deployment of a player over a possession. For example, a player in the Low assignment in a possession would be expected to spend approximately 54 % of the possession in the Net Front role, 21 % of the possession in the Right Goal Line role and 15 % of the possession in the Left Goal Line role.

Matching the number of assignments to the number of players on the ice leads to assignment definitions that neatly match the base formations seen in Figures 9 and 10. It is important to note that these assignments are no longer unique for each possession because no linear assignment problem is formulated at this stage. On average, the number of players in each assignment is close to 1 for each

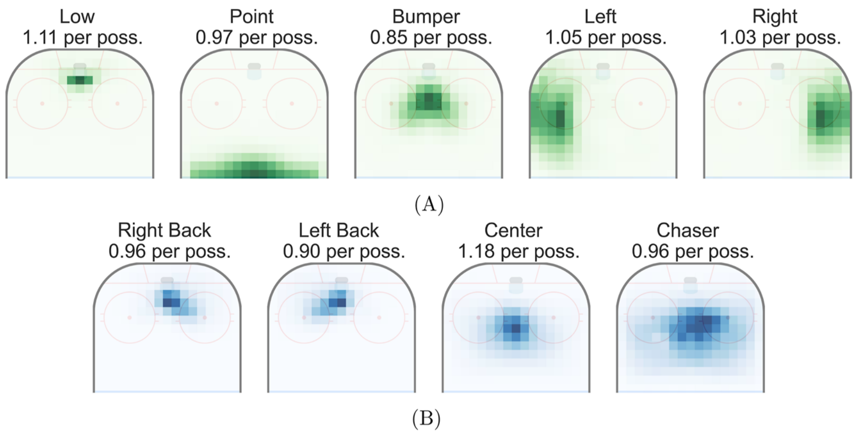


Figure 12: Definitions of the assignments on both the PP and PK. Each assignment can be interpreted as the average deployment of a player over a possession. In both cases, they align with the base formations, PP1 for the PP and PK1 for the PK, shown in Figures 9 and 10, respectively. (A) Power play. (B) Penalty kill.

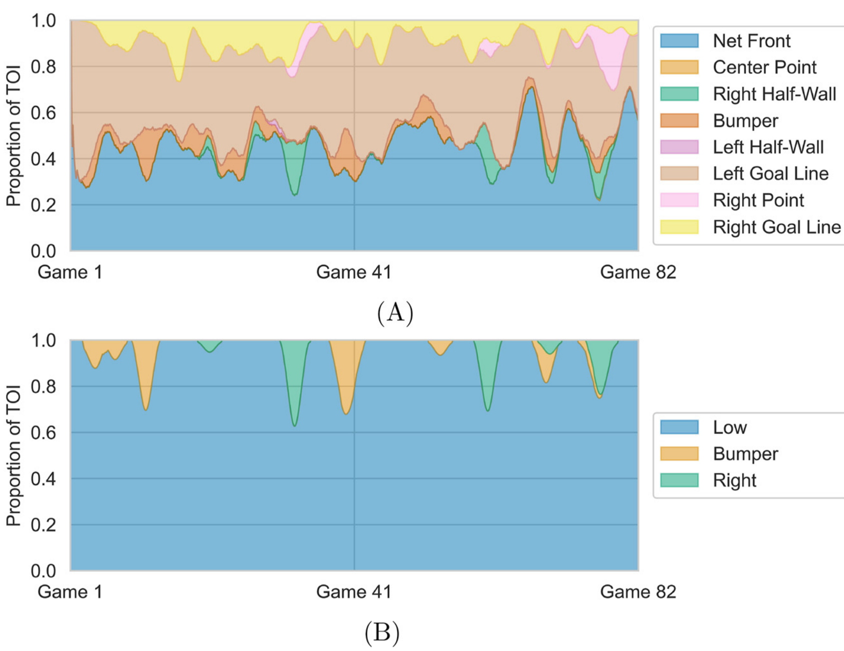


Figure 13: Deployment for a sample player over the course of the season. Assignments, while less granular, provide a more stable measure of player deployment than roles, while preserving the necessary information. (A) Role. (B) Assignment.

possession. However, there are some exceptions. For instance, Figure 12A shows that, on average, 1.11 players are in the Low assignment per PP possession, while only 0.85 players are in the Bumper assignment per PP possession.

The assignments offer a more consistent way of measuring player deployment, as illustrated in Figure 13. Figure 13A displays a rolling average of the proportion of time a sample player spent in each role on the PP. Over the course of the season, the player was primarily tasked with playing the Net Front and the Left Goal Line, with some time on the Right Goal Line. By recreating this figure with assignments, as done in Figure 13B, it becomes apparent that this player was deployed in the Low assignment for the majority of the season. Incidentally, this Low assignment is primarily a combination of the Net Front, Left Goal Line and Right Goal Line roles. Assignments provide an alternative, less detailed method of assessing player deployment, which may prove more well-suited for some analyses.

6 Conclusions

Leveraging player tracking data, this work proposes a novel method for describing player positioning during special teams possessions. Player roles for both the PP and PK are defined using non-negative matrix factorization techniques. These roles provide a low-dimensional and interpretable representation of player positioning. The newly created dataset is utilized for two applications. Firstly, a method of grouping similar possessions using LDA is presented, which effectively characterizes possessions by combining the role-based representation with player handedness. Secondly, assignments are introduced, providing a more stable measure of player deployment over longer time periods.

These are only two of many proposed ways of using the proposed framework to gain actionable insights into special teams play. Notably, outcomes, such as shots and goals, were not considered in this analysis. By expanding the dataset to multiple seasons, it may be possible to quantify how special teams positioning relates to outcomes. For example, individual player statistics could be adjusted for role deployment. Similarly, a regression-based model to estimate a player's role-dependent impact on outcomes could be developed, both on the power play and penalty kill. With such a model, a player's ability in each role could be assessed. This could be used to improve inter-player comparisons, between different players who have similar deployment, and intra-player comparisons, to monitor player performance as their deployment changes. Additionally, the evaluation of units can be improved by comparing the performance of the same five players across changing role allocations.

The efficacy of team-level positioning could be investigated as well, by determining if there is a relationship between a team's preferred formations and performance. However, this may be complicated by the general homogeneity of formations across teams. A more granular look at the subtle differences between teams may be necessary. Coupling an analysis of formations with the previously-described role-dependent impact model, player deployment could be optimized. For example, in the event a key player is forced to miss time, it would be possible to determine whether another player would seamlessly integrate themselves in their role, or whether the team-level tactics must be altered to better integrate a replacement player.

Acknowledgments: The authors would like to thank the National Hockey League for providing access to the data. Additional acknowledgement is extended to the developers of scikit-learn which was used to implement several of the presented models.
Research ethics: Not applicable.
Author contributions: The authors have accepted responsibility for the entire content of this manuscript and approved its submission.
Competing interests: The authors state no conflict of interest.
Research funding: This research has been funded by the Carolina Hurricanes.
Data availability: Not applicable.

Appendix A: Gaussian blurring of player occupancy matrices

Let $\mathbf{O}^{(n)} \in \mathbb{R}_{+}^{L_x \times L_y}$ be the raw player occupancy matrix for player n, where L_x and L_y are the number of bins in the x and y directions, respectively. An example of the raw player occupancy matrix for a sample player is shown in Figure 14A. These occupancy matrices are smoothed using Gaussian blurring to reduce the impact of noise and the arbitrary binning of the ice surface. A two-dimensional Gaussian filter is convolved with $\mathbf{O}^{(n)}$ to yield a smoothed player occupancy matrix, $\tilde{\mathbf{O}}^{(n)}$. The Gaussian filter is a 5×5 matrix, denoted $\mathbf{G}$, approximating a two-dimensional Gaussian distribution with standard deviation $\sigma = 0.5$. Each element of $\tilde{\mathbf{O}}^{(n)}$ is computed as

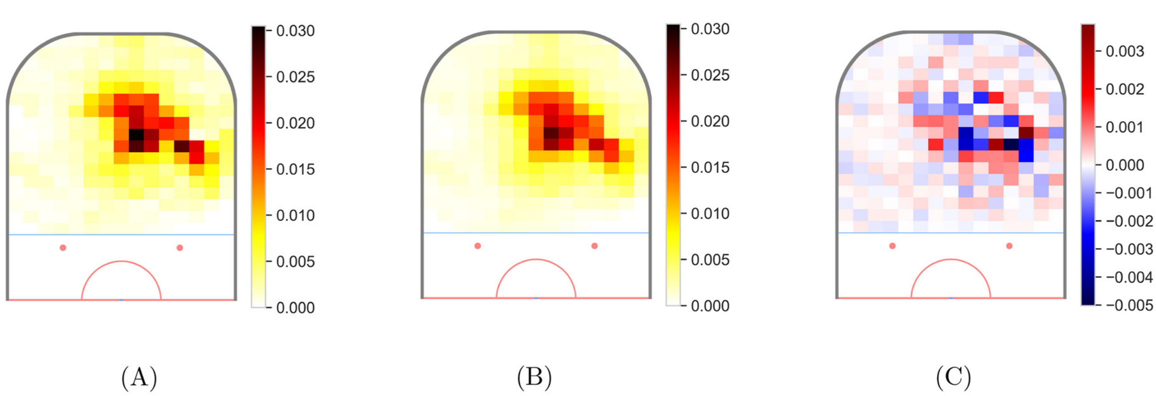


Figure 14: Raw and smoothed player occupancy matrices for a sample player. Applying a Gaussian blur to the raw player occupancy matrix diminishes the impact of the small sample size and the arbitrary binning of the ice surface. The difference between the raw and smoothed player occupancy matrices is shown in (C). (A) Raw matrix $\mathbf{O}^{(n)}$. (B) Smoothed matrix $\tilde{\mathbf{O}}^{(n)}$. (C) Difference, $\mathbf{O}^{(n)} - \tilde{\mathbf{O}}^{(n)}$.

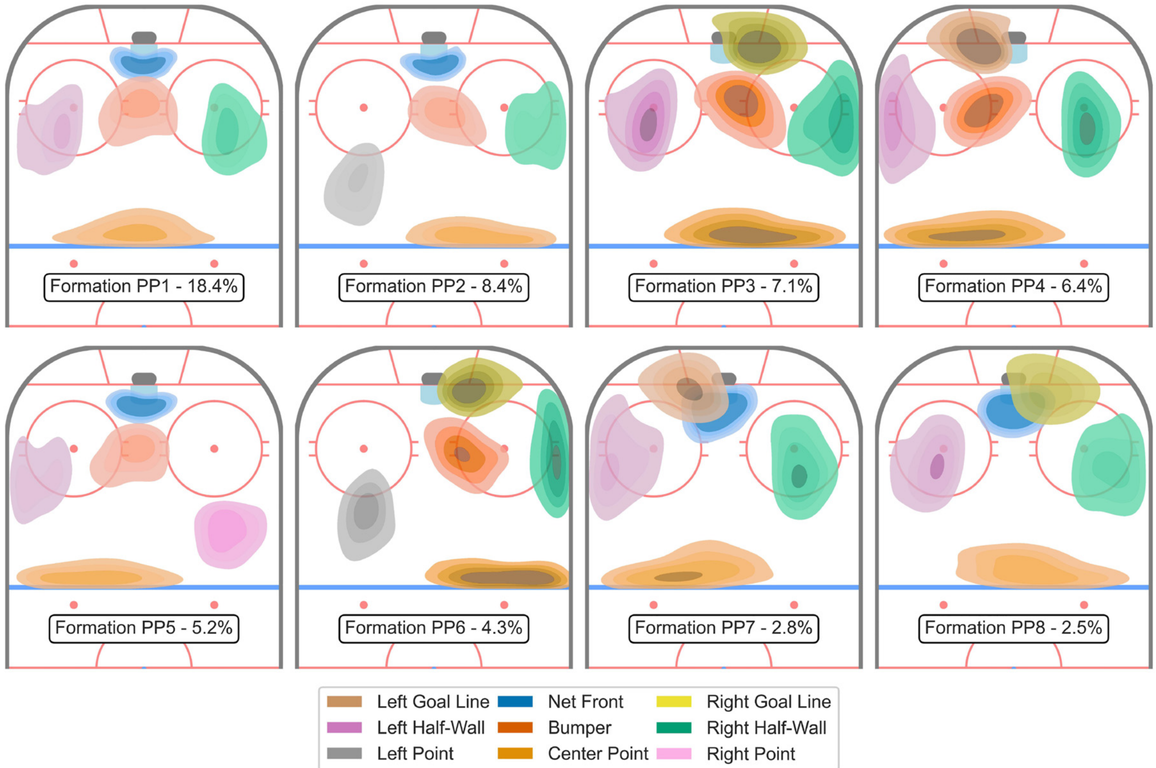


Figure 15: Eight most frequently observed PP formations.

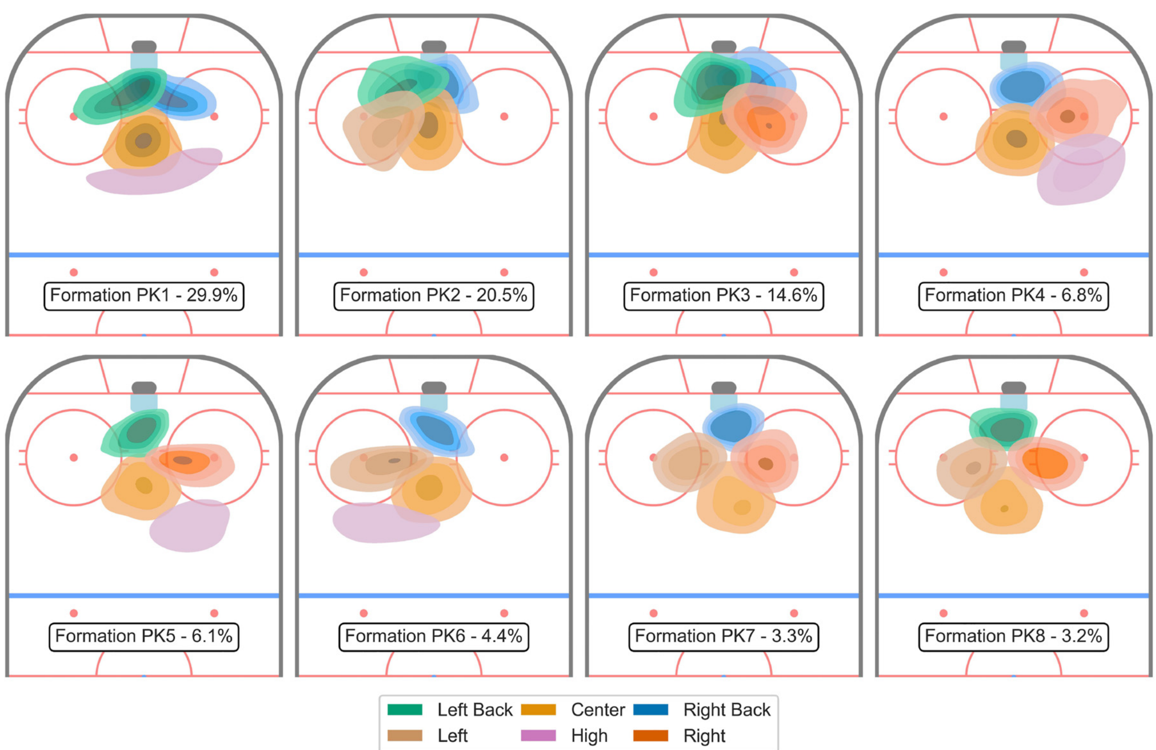


Figure 16: Eight most frequently observed PK formations.

$$\tilde{\mathbf{O}}^{(n)}[i,j] = \sum_{k=-2}^{2} \sum_{\ell=-2}^{2} \mathbf{G}[k,\ell]\mathbf{O}^{(n)}[i+k, j+\ell], \qquad (11)$$

where $i = 1, \ldots, L_x$ and $j = 1, \ldots, L_y$. The `scipy.ndimage.gaussian_filter` Python function is used to implement this, where the default parameters are used. The smoothed player occupancy matrix is shown in Figure 14B. The difference between the raw and smoothed player occupancy matrices is shown in Figure 14C.

Appendix B: Additional formations

The 8 most frequently observed formations on the PP and PK are shown in Figures 15 and 16, respectively.

References

Anzer, G. and Bauer, P. (2022). Expected passes: determining the difficulty of a pass in football (soccer) using spatio-temporal data. *Data Min. Knowl. Discov.* 36: 295−317.

Askari, F., Ramaprasad, R., Clark, J.J., and Levine, M.D. (2022). Interaction classification with key actor detection in multi-person sports videos. In: *IEEE/CVF conference on computer vision and pattern recognition workshops (CVPRW)*. IEEE, New Orleans, LA, USA.

Bialkowski, A., Lucey, P., Carr, P., Yue, Y., and Matthews, I. (2014) “Win at home and draw away”: automatic formation analysis highlighting the differences in home and away team behaviors. In: *MIT sloan sports analytics conference*, Boston, MA, USA.

Bialkowski, A., Lucey, P., Carr, P., Matthews, I., Sridharan, S., and Fookes, C. (2016). Discovering team structures in soccer from spatiotemporal data. *IEEE Trans. Knowl. Data Eng.* 28: 2596−2605.

Blei, D.M., Ng, A.Y., and Jordan, M.I. (2003). Latent dirichlet allocation. *J. Mach. Learn. Res.* 3: 993−1022.

Cane, M. (2017a). *Measuring the importance of structure on the power play*. Hockey Graphs, Available at: https://hockey-graphs.com/2017/02/14/measuring-the-importanceof-structure-on-the-power-play/ (Accessed 13 April 2021).

Cane, M. (2017b). *Second units and zone entries: why teams should go all-in on the 4 forward power play*. Hockey Graphs, Available at: https://hockey-graphs.com/2017/03/07/second-units-and-zone-entries-why-teams-should-go-all-in-on-the-4-forward-powerplay/ (Accessed 15 April 2021).

Cervone, D., D’Amour, A., Bornn, L., and Goldsberry, K. (2016). A multiresolution stochastic process model for predicting basketball possession outcomes. *J. Am. Stat. Assoc.* 111: 585−599, .

Chu, D., Thomson, J., Reyers, M., and Wu, L. (2020). Route identification in the national football league. *J. Quant. Anal. Sports* 16: 121−132.

Dutta, R., Yurko, R., and Ventura, S.L. (2020). Unsupervised methods for identifying pass coverage among defensive backs with NFL player tracking data. *J. Quant. Anal. Sports* 16: 143−161, .

Fani, M., Neher, H., Clausi, D.A., Wong, A., and Zelek, J. (2017) Hockey action recognition via integrated stacked hourglass network. In: *IEEE conference on computer vision and pattern recognition workshops (CVPRW)*, Honolulu, HI, USA, July.

Fernández, J., Bornn, L., and Cervone, D. (2021). A framework for the fine-grained evaluation of the instantaneous expected value of soccer possessions. *Mach. Learn.* 110: 1389−1427.

Gregory, S. (2019) Ready player run: off-ball run identification and classification. In: *Barça sports analytics summit*, Barcelona, Spain, Available at: https://static.capa.biliaserver.com/frontend/clients/barca/wp_prod/wp-content/uploads/2020/01/40ba07f4-ready-player-run-barcelona.pdf (Accessed 21 February 2023).

Gudmundsson, J. and Horton, M. (2017). Spatio-temporal analysis of team sports. *ACM Comput. Surv.* 50.

Hochstedler, J. and Gagnon, P.T. (2017) American football route identification using supervised machine learning. In: *2017 MIT sloan sports analytics conference*, Boston, MA, USA.

Kasan, S. (2008). *Off-ice officials are a fourth team at every game*. NHL, Available at: https://www.nhl.com/news/off-ice-officials-are-a-fourth-team-at-every-game/c-388400 (Accessed 31 March 2020).

Kovalchik, S.A. (2023). Player tracking data in sports. *Annu. Rev. Stat. Appl.* 10: 677−697.

Kovalchik, S.A. and Albert, J. (2022). A statistical model of serve return impact Patterns in professional tennis, arXiv:2202.00583, February, arXiv: 2202. 00583 [stat] (Accessed 14 November 2023).

Kuhn, H.W. (1955). The Hungarian method for the assignment problem. *Nav. Res. Logist. Q.* 2: 83−97.

Lee, D.D. and Seung, H.S. (2000). Algorithms for non-negative matrix factorization. *Adv. Neural Inf. Process. Syst.* 13: 556−562.

Liu, G. and Schulte, O. (2018) Deep reinforcement learning in ice hockey for context- aware player evaluation. In: *International joint conference on artificial intelligence*.

Lucey, P., Bialkowski, A., Carr, P., Morgan, S., Matthews, I., and Sheikh, Y. (2013) Representing and discovering adversarial team behaviors using player roles. In: *IEEE conference on computer vision and pattern recognition (CVPR)*.

Lucey, P., Bialkowski, A., Carr, P., Yue, Y., and Matthews, I. (2014) “How to get an open shot”: analyzing team movement in basketball using tracking data. In: *MIT sloan sports analytics conference*.

Macdonald, B. (2012). Adjusted plus-minus for NHL players using ridge regression with goals, shots, fenwick, and corsi. *J. Quant. Anal. Sports* 8: 1−24.

McCurdy, M.B. (2022). *The magnus prediction model, version 6*. HockeyViz, Available at: https://hockeyviz.com/txt/magnus6ST (Accessed 10 January 2023).

Miller, A.C. and Bornn, L. (2017) Possession sketches: mapping NBA strategies. In: *2017 MIT sloan sports analytics conference*, Boston, MA, USA.

Miller, A., Bornn, L., Adams, R., and Goldsberry, K. (2014) Factorized point process intensities: a spatial analysis of professional basketball. In: *International conference on machine learning*, Beijing, China, pp. 398−414.

Nandakumar, N. and Jensen, S.T. (2019). Historical perspectives and current directions in hockey analytics. *Annu. Rev. Stat. Appl.* 6: 19−36.

Parnass, A. (2016a). *How can we quantify power play performance in formation?* Hockey-Graphs, Available at: https://hockey-graphs

.com/2016/04/25/how-can-we-quantify-power-play-performance-in-formation/ (Accessed 7 June 2022).

Parnass, A. (2016b). *ZEFR rate: a new and better way to evaluate power plays*. Hockey Graphs, Available at: https://hockey-graphs.com/2016/04/18/zefr-rate-a-new-and-betterway-to-evaluate-power-plays/ (Accessed 13 April 2021).

Radke, D., Radke, D., Brecht, T., and Pawelczyk, A. (2021) Passing and pressure metrics in ice hockey. In: *Artificial intelligence for sports analytics (AISA) workshop at IJCAI '21*, Montreal, QC, Canada.

Radke, D., Brecht, T., and Radke, D. (2022). Identifying completed pass types and improving passing lane models. In: *Linköping hockey analytics conference* . LINHAC, Linköping, Sweden.

Ritchie, R., Harell, A., and Shreeves, P. (2022) Pass evaluation in women's Olympic ice hockey. In: *International ACM workshop on multimedia content analysis in sports*, Lisbon, Portugal (Accessed 26 October 2022).

Schulte, O., Zhao, Z., Javan, M., and Desaulniers, P. (2017) Apples-to-Apples: clustering and ranking NHL players using location information and scoring impact. In: *MIT sloan sports analytics conference*.

Shaw, L. and Glickman, M. (2019) Dynamic analysis of team strategy in professional football. In: *Barça sports analytics summit*, Barcelona, Spain, Available at: https://static.capabiliaserver.com/frontend/clients/barca/wpprod/wp-content/uploads/2020/01/56ce723e-barca-conference-paper-laurie-shaw.pdf.

Tora, M.R., Chen, J., and Little, J.J. (2017) Classification of puck possession events in ice hockey. In: *IEEE/CVF conf. on computer vision and pattern recognition workshops (CVPRW)*, Honolulu, HI, USA.

Vats, K., Fani, M., Clausi, D.A., and Zelek, J. (2021) Puck localization and multi-task event recognition in broadcast hockey videos. In: *IEEE/CVF conference on computer vision and pattern recognition workshops (CVPRW)*, Nashville, TN, USA.

Vats, K., Walters, P., Fani, M., Clausi, D.A., and Zelek, J.S. (2023) Player tracking and identification in ice hockey. In: *Expert systems with applications*, p. 213.

Wei, X., Sha, L., Lucey, P., Morgan, S., and Sridharan, S. (2013) Large-scale analysis of formations in soccer. In: *International conference on digital image computing: techniques and applications (DICTA)*, Hobart, TAS, Australia.

Yurko, R., Matano, F., Richardson, L.F., Granered, N., Pospisil, T., Pelechrinis, K., and Ventura, S.L. (2020). Going deep: models for continuous-time within-play valuation of game outcomes in American football with tracking data. *J. Quant. Anal. Sports* 16: 163−182.