Linköping Hockey Analytics Conference (LINHAC 2022)

Identifying Completed Pass Types and Improving Passing Lane Models

David Radke, Tim Brecht and Daniel Radke

David R. Cheriton School of Computer Science, University of Waterloo

Abstract. The implementation of a puck and player tracking (PPT) system in the National Hockey League (NHL) provides significant opportunities to utilize high-resolution spatial and temporal data for advanced hockey analytics. In this paper, we develop a technique to classify pass types in the tracking data as either Direct, 1-bank, or Rim passes. We also address two fundamental limitations of our previous model for passing lanes by modeling 1-bank indirect passes and the expected movement of players. We implement our pass classification and extended passing lane models and analyze 198 games of NHL tracking data from the 2021-2022 regular season. We study the types of completed passes and introduce a new passing metric that shows about 59% of completed 1-bank passes have an equal or more open indirect passing lane than the direct lane. Furthermore, we show that our expected movement addition reduces receiver location error in over 94% of completed passes.

Keywords: Hockey · Passing · Metrics · Passing Lanes · Tracking Data

1 Introduction

Ball and player tracking systems have revolutionized soccer and basketball analytics with extensive implications for scouting, coaching, player development, and fan engagement. Recently, the National Hockey League (NHL) deployed a puck and player tracking (PPT) system that records the location of the puck and every player with high resolution and frequency (60 and 12 times per-second for the puck and players respectively). Traditional methods of performance evaluation in hockey have relied mostly on offensive events like goals and shots despite these representing only a small fraction of the actual game play. Hockey has lagged behind other sports in advanced analytics due to technical challenges caused by the fast pace, small puck, white-colored ice, and other hardware challenges [14, 13, 3]. However, the new tracking system broadens the scope of potential metrics, analysis, and performance evaluations in hockey.

Most of the game play in hockey involves puck possession and passing between teammates. Previously, we developed a model to quantify the availability of passing lanes for completed passes, which associated smaller values with more difficult (or less open) passes [8]. While that model effectively calculates the available space between a passer and receiver, it assumes passes can only be direct (i.e., they are not banked off of or around the boards) and it treats player

locations as static with respect to the time the pass was initiated. In reality, players often use the boards to complete passes when direct passing lanes are small or unavailable and pass to where their intended receiver is expected to be instead of where they are at the time the pass is initiated. In this paper, we make the following contributions:

- We develop a model to classify completed passes in PPT data to be either Direct, 1-bank, Rim, or Other passes. This is required to apply passing lanes models that are appropriate for different types of passes.
- We extend our passing lane algorithm [8] to 1) model the available passing lanes for 1-bank indirect passes and 2) include the expected movement of all players while the pass is made.
- We analyze passes using PPT data from 198 games from the 2021-2022 NHL regular season and devise a new metric for comparing completed 1-bank indirect passes with the alternative direct passing lane to the same receiver. We also examine the improvement in receiver location accuracy of our expected player movement model.

2 Related Work

Numerous passing models have been developed for football (soccer) and basketball using tracking data. These are typically used to analyze aspects of the game, such as pass disruptions to defensive formations [5], the expected value of passes [2, 4], and the number of outplayed opponents by passes [11]. The focus of that work is on the impact of passes instead of the actual difficulty or risk associated with a pass, which could provide insight into decision making, skills, and player risk profiles. Expected pass completion models (xPass) have gained popularity in soccer and are used to estimate the probability of passes being completed, or the difficulty of a pass, using physics [9], logistic regression [7], Graph Neural Networks [12], and supervised machine learning [1]. While these models can give significant insight into a player's decision making and passing ability, they rely on data for incomplete passes or ball control which may be difficult to determine in hockey.

To model the availability of passing lanes without relying on data for incomplete passes, Steiner et al. [11] calculate the angle from the direct pass line to the nearest opponent, where smaller angles correspond to less available passes. This model is limited by not including opponents behind the passer or receiver and not scaling for pass length. In response to these limitations, in previous work [8] we defined four key requirements for a passing lane model: **1)** always assign a real numbered value, **2)** incorporate the area surrounding the passer and receiver, **3)** be asymmetric with respect to pass direction, and **4)** scale with respect to the pass length. Our passing lane model presented in [8] assigns a value to each pass (in $\mathbb{R}^+$) that defines how *open* a passing lane is and simultaneously satisfies all four requirements without requiring data for incomplete passes. In this paper, we extend this passing lane model in three ways. We classifying different types of passes from the PPT data, calculate passing lanes for 1-bank indirect passes, and model the expected movement of players.

2.1 Background

Puck and Player Tracking Dataset Location data is collected through tracking technology that is inserted into the sweater of each player (back of the right shoulder) and embedded into pucks. Location information contains x, y, and z-coordinates to record locations in 3-dimensional space. The x and y locations are relative to center ice (which is 0,0) and the z locations are relative to the surface of the ice. The PPT data is recorded at 60 locations per second for the puck and 12 locations per second for each player on the ice, resulting in a total of about 734,400 location readings of main interest in a 60 minute game. Additional location data is obtained once a second for players that are deemed to be off of the ice. The tracking data is accompanied by event data including shots, goals, faceoffs, hits, and completed passes among others. These event labels contain information about the time of the event and the identities of the players involved.

Passing Lane Model To the best of our knowledge, the passing lane model in [8] is the only attempt to quantify the availability of passing lanes in hockey. The model uses the spatial locations of players in PPT data to estimate the available space between a passer p and any receiver r. The model utilizes event labels in the tracking data which have been identified by the data collection company, SportsMEDIA Technology (SMT)[1].

For each passing event, the passing lane model constructs a teardrop-like passing lane shape (shown in Figure 1) between the x, y locations of a passer p and receiver r that simultaneously satisfies all four requirements listed in Section 2. The size of this lane is determined by the locations of opposing players, representing the space between p and r *without* opponents (i.e., the *open* space). The passing lane size and shape is described by a positive real-numbered value γ, where larger γ values represent a wider lane and more open pass.

To determine the value of γ for a pass, we initialize $\gamma = 0$ (the direct line from p to r, pr in Figure 1). In this paper, we relabel pr to be $\vec{pr}$ to use vector notation. Increasing γ expands the passing lane shape until the edge of the lane contacts the location of an opponent. For example, increasing γ in Figure 1 grows the passing lane from the blue, to the green, to the yellow shaded regions. Since opponent 1 (o_1) is contacted first by the growing shape, the passing lane from p to r is represented as the blue shaded region. The resulting γ value is determined to be the passing lane value (for efficiency, we implement binary search instead of unidirectional growth). In Figure 1, $\gamma = 0.6$ since it is the smallest γ value with respect to each opponent (i.e., o_1 was contacted by the growing passing lane first). While γ has no direct correspondence with completion percentage, values of γ can be compared across time, locations, or players. We refer the reader to [8] for a more detailed description about the original passing lane model.

[1] www.smt.com/hockey

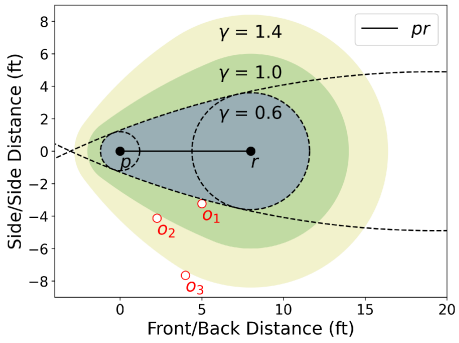


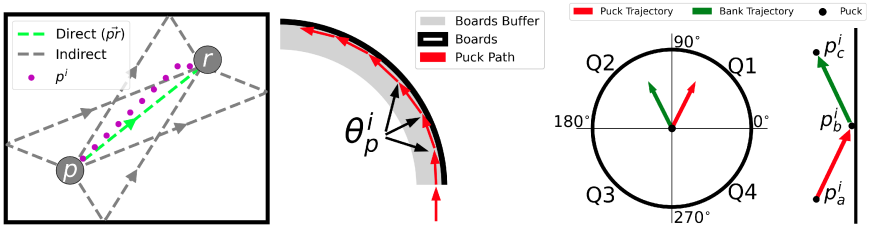
Fig. 1: Passing lane diagram from [8]. This example shows three passing lane shapes regulated by a parameter γ. The passing lane grows until the edge contacts the nearest opponent. The passing lane in this example has value $\gamma = 0.6$ and is the blue shaded region (the others are included as examples if o_1 or o_2 did not exist). In this work, we relabel the direct passing line pr as $\vec{pr}$.

3 Completed Pass Classification

The PPT data includes event labels to identify instances of a completed *pass*; however, there are multiple ways to pass the puck in hockey that should be modeled differently. A passer p can pass directly to r (a *Direct* pass), bank off a flat section of boards (a *1-bank* pass), or rim the puck around a curved corner of the surrounding boards (a *Rim* pass). We construct a model to identify these types of completed passes from the PPT data and overcome several challenges in the process. For example, there may exist some noise in the exact location of the puck (potentially from puck fluttering or position accuracy) or the time labels associated with passes. We found that the trajectory of passes cannot be assumed to compose perfectly straight lines, even for Direct passes. The puck may also contact the boards between consecutive readings of its location (i.e., the puck is traveling towards the boards at time t but traveling away from the boards at time $t+1$). Thus, the puck location never truly *contacts* the boards in the data. Our model uses a sequential filtering approach to differentiate between completed passes that are Direct, 1-bank, and Rim passes, and leave more fine-grained classification for future work.

Let $\mathcal{P}$ be the set of all passes in a game, $P^i \in \mathcal{P}$ be a single pass, and p^i be the set of x, y puck locations for P^i (origin at center ice). Our classification algorithm identifies: 1) Direct ($\mathcal{P}^d$), 2) Rim ($\mathcal{P}^r$), and 3) 1-bank ($\mathcal{P}^1$) passes in that order, each time reducing the set of possible passes to consider (starting at $\mathcal{P}$). The remaining unclassified passes compose a fourth class, *Other*, which we discuss in detail later. Since two consecutive readings close to the boards may represent actual contact with the boards (a challenge described above), all three phases use a value d_b, a distance from the boards, to construct a buffer that is used to determine puck readings that are sufficiently close to the boards.

1) Direct Passes: We identify completed Direct passes using two characteristics: 1) they may never be close to the boards and 2) have relative straight trajectories when compared with the possible indirect passes to the receiver. Our algorithm has two phases. First, if no points in p^i are within distance d_b of the boards, P^i is classified as to be Direct. Since Direct passes *may* also happen close to the boards, if any points in p^i are within distance d_b of the boards, we proceed to the second phase. If not identified as a Direct pass in the first phase, we determine the five possible paths for p to pass to r, ignoring corners (i.e., the direct path $\vec{pr}$, and off of both side-boards and both end-boards). Figure 2a shows this procedure in an example box (not-to-scale). The purple dots represent p^i, which has some change in direction near the receiver (i.e., likely contacting the receiver's stick before being considered *received*). To mimic actual puck behavior, we remove any of the five passing paths that contact the net or a rounded corner since the puck would not follow the projected trajectory following contact. We estimate the error from P^i to each projected path using the total Euclidean distance from each of the points p^i to each of the five possible paths. If $\vec{pr}$ has the least error, the pass is considered Direct.



(a) Direct (b) Rim (c) 1-bank

Fig. 2: (a) Project 5 ways for p to pass to r (excluding corners). "Direct" pass if the path with least error (Euclidean distance from p_i) is $\vec{pr}$ and/or puck is never within distance d_b of the boards. (b) Calculate puck direction changes in the rink corners. "Rim" passes have more than three direction changes in a corner that are greater than threshold t^θ. (c) Identify where the puck trajectory direction changes quadrants of the Unit Circle. "1-bank" passes have at most 3 of these points within distance d_b of the boards for specific changes of direction.

2) Rim Passes: The set of remaining completed passes are those not classified as Direct (i.e., indirect). Some indirect passes may be rims, where p directs the puck around a curved corner of the boards so the puck contacts the boards multiple times (Figure 2b). Our intuition to classify "Rim" passes is that the puck 1) changes direction multiple times and 2) these changes in direction are close to the corner boards. To calculate general puck direction vectors (and reduce change in direction noise), we average every 10 readings for the puck locations for passes $|p_i| > 10$ (red arrows in Figure 2b; shorter passes are not averaged). We calculate the difference in direction between adjacent vectors θ_p^i (in degrees),

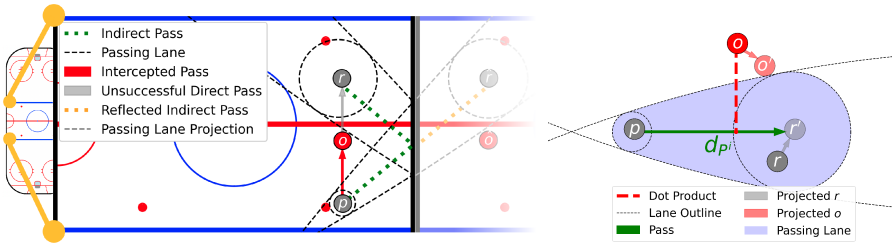
and define a threshold t^{θ} to determine if a direction change is sufficiently large. Since direction changes can have several causes (e.g., deflection from a stick, player, or referee) we only consider those direction changes that occur within d_b of the corner boards. In our implementation, a Rim pass is determined to have *greater than* three direction changes greater than $t^{\theta} = 4^{\circ}$ within distance d_b of the corner boards. We choose three points since 1-bank passes should contain at-most three points with specific direction changes (explained next).

3) 1-bank Passes: From the remaining completed passes, we determine the set of 1-bank passes, where the puck only contacts a straight segment of the boards once. Since we are only detecting a single change in direction, the model for Rim passes is unable to be adapted since a change in the puck's direction of travel could happen for any number of reasons (i.e., deflections from sticks, inaccuracies in device readings, or inaccuracies with time labels associated with the pass). Therefore, we build on the intuition of detecting significant types of direction changes since a puck contacting the boards once will completely change its direction of travel. Our model draws on concepts from the quadrants of the Unit Circle in Trigonometry (Figure 2c left). To reduce the noise in the puck's trajectory we use an average of 10 consecutive readings (we do not use averages for short passes). For example, a sequence of 30 points could result in the three points p_a^i, p_b^i, and p_c^i shown on the right side of Figure 2c. We then calculate vectors between these points to determine the general direction of the puck (red and green vectors in Figure 2c). In the right of Figure 2c, the red vector (from point p_a^i to p_b^i) represents the puck traveling towards the boards (at 60°), and the green vector (from point p_b^i to p_c^i) represents the puck traveling away from the boards after the contact (now at 120°). Note that angles are relative to 0° which is the line perpendicular to the boards in this example.

We plot these vectors for the puck traveling to and from the boards on the Unit Circle shown on the left of Figure 2c. For a 1-bank pass, our model identifies the three points (p_a^i, p_b^i, and p_c^i) that comprise two consecutive vectors (red and green) where their directions appear in different quadrants of the Unit Circle $(0^{\circ}, 90^{\circ}, 180^{\circ}, 270^{\circ})$. It is not possible for a puck to contact a straight segment of boards and continue in the same quadrant of the Unit Circle. Therefore, 1-bank passes are classified if *three or fewer* points associated with puck direction vectors that are within distance d_b of the boards where the direction changes due to the boards (in the example in Figure 2c the angle of the vectors changes quadrants from Q1 to Q2).

4 Passing Lanes for 1-bank Passes

The original passing lane model only considers the direct line from p to r, which only represents Direct passes. For example, the model would consider the passing lane from p to r in Figure 3a extremely small (red arrow) because there is an opponent o directly on the path from p to r. However, a 1-bank pass can avoid the opponent o and is more open than the Direct pass. Our goal in this section is to model such passing lanes.



(a) 1-bank Passing Lane (b) Expected Movement

Fig. 3: (a) We calculate passing lanes for 1-bank passes by reflecting receiver r and opponent o about the boards. (b) We fit the passing lane to the expected movement of r and o using their locations, velocities, and expected pass distance.

Using the theory of geometric reflections, a 1-bank pass off the boards is geometrically equivalent to a Direct pass *through* the boards to a reflected representation of r ($\hat{r}$). This assumes the angle of incidence is equal to the angle of reflection which we acknowledge may not be completely accurate due to puck spin, fluttering, board imperfections, and variables such as drag and energy loss. However, our model is an approximation of the available passing lane instead of modeling the exact trajectory of the pass. We *reflect* all players besides the passer about the boards so that the 1-bank pass can be modeled as a Direct pass (in Figure 3a, green dotted line can be modeled as the orange dotted line extension). In the example in Figure 3a, we keep p at it's location and reflect r and o to locations $\hat{r}$ and $\hat{o}$ respectively. Using $\hat{r}$ as the location of r, we calculate the passing lane with respect to the nearest opponent (also considering their reflections). We acknowledge that 1-bank passes should be considered more difficult than Direct passes. This is accounted for in our model, as γ scales with the pass length and a 1-bank pass would be longer than the Direct pass.

To consider an in-game example when both teams are at even strength (no penalties), consider a passer p. Given any potential receiver r, p has the option to make a Direct pass, or bounce the puck off either side-boards or end-boards (e.g., shown in Figure 2a). Some of these lanes will make more sense than others, since a player is unlikely to pass the puck off their defensive end-wall when in the offensive zone. Since γ decreases as the length increases, excessively long 1-bank passes will receive very low γ values. We calculate γ for all five passing options from p to r with respect to all opponents. The largest γ value is the *most* open passing option for p to pass the puck to r. We expect a similar reflection-based methodology may work for Rim passes, but leave this for future work.

5 Expected Player Movement

To compute γ, the passing lane algorithm in [8] uses the locations of players taken at the time the pass is initiated. The asymmetry of the passing lane shape

accounts for opponents closer to r having more time to react to a pass (i.e., skate towards the pass and/or move their stick in an attempt to intercept or disrupt the pass). Furthermore, receiver r will most often not be stationary and receive the pass at a different location than where they were when the pass was initiated. We expand the previous passing lane model to include the expected location of all players when computing the passing lane. In Section 6, we demonstrate how our new model improves the expected location of the actual pass reception. Our method calculates 1) the approximate distance of a pass P^i (d_{P^i}), 2) the expected speed of the pass ($s_{\vec{pr}}$), 3) the duration of the pass (t_{P_i}), and 4) the expected locations of receiver r and opponents o (r' and o'). Visualized in Figure 3b, we fit the passing lane from p to r' with respect to o'.

The approximate pass distance (d_{P^i}) is calculated using the Euclidean distance from p to where the pass is estimated to be received (r'), defined later. Given the approximate pass distance, we train a linear regression model on previous passes π to produce the expected speed of a pass with distance d_{P^i}, so that $\pi(d_{P^i}) \rightarrow s_{\vec{pr}}$ produces a positive real number. We calculate the duration of a pass as $t_{P^i} = \frac{d_{P^i}}{s_{\vec{pr}}}$ and use r's velocity vector (which includes direction) at the time of the pass $\mathbf{v_r}$ to determine their expected location, $r' = r + t_{P^i}\mathbf{v_r}$. This assumes r will continue in the current trajectory for t_{P^i} time.

Since opponents only have until the puck passes their location to disrupt the pass, we only project o's movement for time t_o, the time until the puck passes their location along the pass trajectory $\vec{pr}$. For this computation, consider the example in Figure 3b where o is located between p and r. Taking the dot product of o with respect to p and the direct passing line $\vec{pr}$ determines the perpendicular location of o onto $\vec{pr}$ (the red dashed line in Figure 3b). The expected time for the puck to reach this intersection is calculated to be t_o. If o is behind p, $t_o = 0$, and if o is behind r, $t_o = t_{P^i}$. We solve $o' = o + t_o\mathbf{o_r}$, where $\mathbf{o_r}$ is the velocity vector for o at the time the pass is made. Given the locations of p, r' and o', we calculate the passing lane using the algorithm from [8], shown as the blue shaded region in Figure 3b. In this example we only show one r and o for simplicity; however, we can calculate passing lanes for any r with respect to all opponents. This allows us to determine the receiver with the largest passing lane (i.e., the most open player). The tracking data does not include information on stick location, although this could be included if collected or estimated in future work.

6 Analysis

We implement our pass classification algorithms and passing lane extensions to analyze completed passes using a combination of the raw tracking data and labeled event data. Our dataset is from 198 games played in November of the 2021-2022 NHL regular season. We utilize the pass event labels in the dataset to determine when a pass was made; however, this dataset does not contain labels for passes that were not completed. Additionally, the automated labeling of events is a difficult problem; thus, the dataset may be missing some completed passes and/or include labels for events that are not actually completed

passes. This dataset is still considered unofficial by the NHL, and may differ from other datasets that contain complete and/or incomplete passes (e.g., a hand labeled dataset). In this paper, we utilize the event labels provided in the dataset while including techniques to handle some, but not all, inaccuracies. We analyze features of our classification algorithm, 1-bank passing lanes, and expected movement extensions in isolation to identify interesting passing behavior and hypothesize about the potential performance of our models in the absence of ground truth data.

6.1 Pass Classification and Statistics

We first analyze general features of the completed passes in our dataset and how our classification model differentiates them. We observed $d_b = 2.5$ feet (ft) captures multiple adjacent puck readings for Rim and 1-bank passes that are close to the boards in most cases. Table 1 shows the results of our pass classification algorithm on the set of all completed passes. Since our tracking dataset only includes completed passes, the information may be biased towards successful events and not necessarily reflect the game as a whole (i.e., what was attempted and failed). As shown in Table 1, our model identifies 84.4% of the passes labeled complete in this dataset to be Direct, 10.2% to be 1-bank, 2.6% to be Rim, and 2.8% to be Other. Forwards as a whole tend to complete slightly more passes (49.5%) than defence (47.2%); however, when considering there are typically two defence and three forwards on the ice, a defensive player on average completes 43% more passes than a forward. We consider the relatively small percentage of unclassified completed passes (Other) to be acceptable, but is something we plan to examine in future work. After manually inspecting a significant number of these unclassified passes, we believe that most are either mislabeled as passes or consist of edge-cases that are difficult to identify (e.g., inaccurate timestamps resulting in odd changes in trajectory by a player).

Type	Direct	1-Bank	Rim	Other	Total
Forward %	41.9	4.8	1.2	1.5	49.5
Defence %	39.7	5.1	1.2	1.1	47.2
Goalie %	2.8	0.3	0.2	0.1	3.4
Avg/Game %	84.4	10.2	2.6	2.8	100.0

Table 1: Completed pass categorizations. Note that this data is based on events labeled as completed passes. Actual values may differ if labels are incorrect, missing and/or if incomplete passes are included.

Figure 4 compares the paths of completed 1-bank passes made by defence (left) and forwards (right). Darker green represents more 1-bank passes in that region. We see that most of the completed 1-bank passes initiated by defence are

behind their own net or off the defensive half-walls. In contrast, the majority of completed 1-bank passes from forwards are made behind the offensive net or off the offensive half-walls (likely passing back to defence).

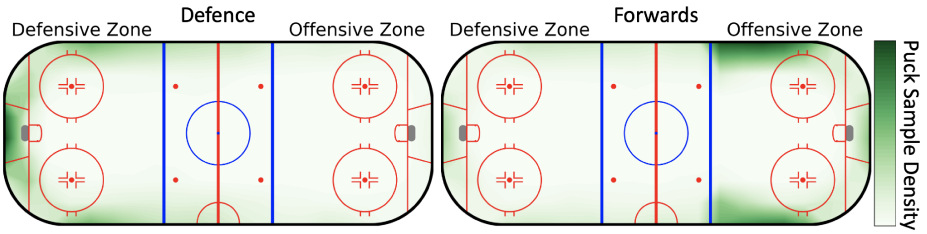
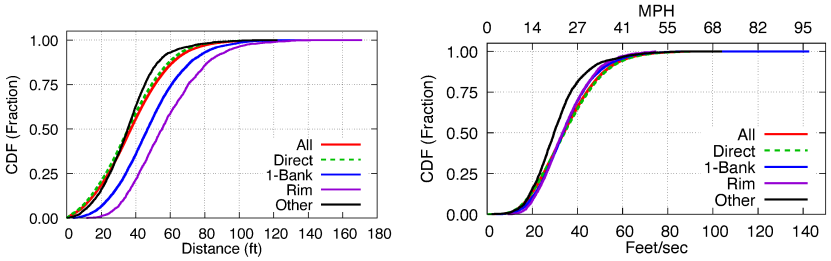


Fig. 4: Heatmap of completed 1-bank passes; defence (left), forwards (right).

Figures 5a and 5b show Cumulative Distribution Functions (CDF) for pass distance (puck travel distance) and pass speed for completed passes, calculated by comparing the total puck travel distance with the duration of the pass. For example, half (0.5) of all completed passes had a distance of about 38 ft or less, shown in Figure 5a. Interestingly, while the distances for completed indirect passes (1-bank and Rim passes) are typically longer than completed Direct passes, we observe almost no distinct difference between these classes for pass speed (Figure 5b). Thus, we hypothesize that players pass the puck harder towards the boards for indirect passes than they would for a Direct pass, to account for the expected energy loss from the boards. The distributions of this data may be different when considering all pass attempts.

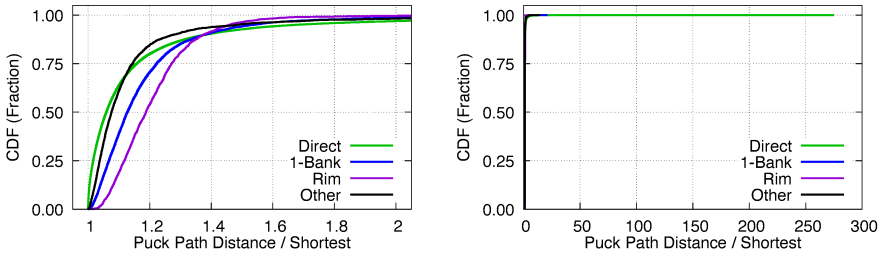


(a) Pass puck travel distances. (b) Pass speeds.

Fig. 5: (a) CDF of pass length (distance traveled) for each type of completed pass. (b) Speed of completed passes for each type (distance traveled divided by total duration of the pass).

Figure 6a shows a CDF for the extra distance the puck traveled for each type of pass. We calculate this as $\frac{d_{puck}}{d_{p,r}}$, where d_{puck} is the actual distance the puck traveled and $d_{p,r}$ is the Euclidean distance from p to r (i.e., the shortest path for

the puck). In theory, Direct passes should have the least extra distance traveled compared to $d_{p,r}$ and Rim passes must travel corners which accumulates more distance. Figure 6a shows that Direct passes generally do travel the least extra distance, followed by 1-bank, and Rim passes.



(a) CDF Extra Distance Traveled (zoom). (b) CDF Extra Distance Traveled.

Fig. 6: (a) CDF comparing the distance the puck traveled to the shortest possible distance (Euclidean distance from p to r) for completed passes. (b) Zoomed out CDF to show the long tail, likely due to event labeling or classification errors.

We do note that the actual path of most Direct passes is longer than the shortest possible path ($d_{p,r}$). These passes are those with values on the x-axis greater than 1. Extreme examples of this can be also seen by the long tail in Figure 6b (an un-zoomed version of Figure 6a). We believe this is due to pucks being deflected by sticks or bodies (but the pass should still be considered Direct). Furthermore, inaccuracies in the timestamps of pass events also lead to add additional distances.[2] Motivated by these challenges, our first classification phase considers such passes Direct if the puck is not within distance d_b from the boards during the pass. We hypothesize that the Direct, 1-bank, and Rim ordering for extra distance in Figure 6a provides some insight into the accuracy of our classification algorithm despite these artifacts.

6.2 Passing Lanes for Indirect vs Direct Passes

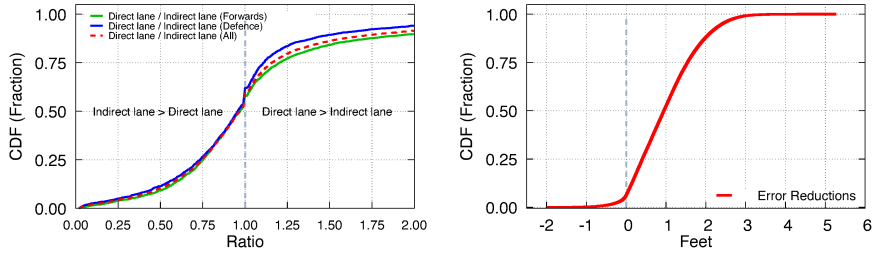
We now analyze our addition to the passing lane model for calculating 1-bank passing lanes. Without any ground truth for how open a passing lane is, our goal is to analyze how our passing lane model captures 1-bank passing behavior by comparing Direct and indirect passing lanes for completed 1-bank passes. For this analysis, we only consider the set of completed 1-bank passes for the reason that a more open indirect passing lane does not always indicate a better play and depends on the context of the game. For example, a player will likely opt

[2] By manually inspecting a significant number of these cases, we observed the timestamp at the end of the pass may occur after the pass was received and the receiver changed directions.

for a Direct pass on a 2-on-1 offensive rush instead of a 1-bank pass, even if the 1-bank is technically more open.

For each completed 1-bank pass, we calculate the value of the indirect 1-bank passing lane γ_i as well as the direct passing lane γ_d for p to pass to receiver r and define a new metric, $\gamma\text{-}ratio = \frac{\gamma_d}{\gamma_i}$. If the $\gamma\text{-}ratio < 1$, the indirect passing lane was more open than the direct lane, otherwise the Direct pass was actually more open. Figure 7a shows a CDF of the $\gamma\text{-}ratio$ for completed 1-bank passes separated by player position for forwards and defence. We observe that about 59% of 1-bank passes were completed when the 1-bank passing lane was equal to or more open than the direct passing lane size (the $\gamma\text{-}ratio \leq 1$). There is little difference between the behavior of forwards and defence when the $\gamma\text{-}ratio < 1$; however, when the $\gamma\text{-}ratio > 1$, defence tend to make more 1-bank passes when both lanes are similar (i.e., the $\gamma\text{-}ratio$ closer to 1).

Note that in Figure 7a the x-axis is centered around 1 and is limited to a maximum of 2, since if γ_d is much larger than γ_i, the $\gamma\text{-}ratio$ grows instead of trending to zero. For an in-game scenario, Figure 8 (left) in Section 7 shows how our model captures the 1-bank passing lane from Player #86 (who has possession of the puck) up to Player #3, whereas our previous model [8] does not. For this pass, the $\gamma\text{-}ratio = \frac{0.23}{0.46} = 0.5$ and completing this pass increases the subsequent passing lane to Player #28 from $\gamma = 0.3$ to 0.98 (right side of the figure).



(a) The $\gamma\text{-}ratio$ by position. (b) Expected movement improvement.

Fig. 7: (a) CDF of the $\gamma\text{-}ratio$ to show the fraction of completed 1-bank passes where the 1-bank passing lane was more open (< 1) or direct lane was actually more open (> 1). (b) Location of r error improvements with expected movement.

6.3 Player Movement

Our motivation for including expected player movement when model passing lanes is to better fit the shape of the passing lane to the location of the receiver when they receive the pass (and opponents to where they would be when the puck passes their location). For the set of all completed passes, we have the labeled location of the receiver at the time when the pass is considered received (r_t^*). Therefore, we calculate the difference between r_t^* and their location when

modeled with expected movement (r'_t) and without expected movement (r_t). We calculate the two location errors as the Euclidean distance between 1) r's true location and their projected location with expected movement (r^*_t and r'_t), and 2) r's true location and their location without expected movement (r^*_t and r_t). The difference of these two errors provides insight into whether or not the expected movement model better estimates the location of r when they receive the pass (i.e., there less error). Figure 7b shows a CDF for the difference between these two errors, where positive values correspond with expected movement reducing the location error by the distance along the x-axis (more accurate location of r). We find that expected movement reduces the error of r's location for over 94% of passes, in one case up to 5.3 ft, and increases error in a small fraction of passes by a small amount (at most up to 2.0 ft).

7 Potential Applications

The influx of data in professional sports has given broadcasters and fans the ability to absorb more information about an event, such as shot speed, shift length, or face-off win probabilities; this information is typically presented by overlaying graphics or augmented reality (AR) on the live video broadcast. Our passing lane model can also be used in this context to display the most available passing option for one or more teammates, or the γ of a successful pass. Furthermore, our model could provide more fine-grained metrics that may be useful in fantasy sports or gambling applications. This can increase fan engagement and enjoyment by drawing attention to player formations and passing options.

When reviewing video of games, our passing lane model would give players and coaches quantitative data for the availability of passing lanes to devise new plays or assess performance. For example, the "up-and-over" is a common powerplay sequence to shift the defence to a new side of the ice and open passing lanes to certain players, shown in Figure 8. Using our models, coaches would be able to adjust the location of offensive or defensive players to find positioning to increase passing lane sizes, or to reduce the size of an opponent's passing lanes.

While GMs are tasked with constructing rosters and assessing players, watching every game or shift of a player is often infeasible and current metrics (such as goals and points) provide only a coarse view of player performance skewed towards offense. Our passing lane model could quantify passing behavior in a game or across a season (for assessing consistency). If augmented with incomplete passes, our model could determine how often players force passes when a more open alternative is available and provide insights into the passing skills of players (e.g., whether players manage to complete passes with smaller lanes).

8 Discussion

The high fraction of the γ-*ratio* ≤ 1 (59%) shows that NHL players in our dataset typically complete 1-bank passes when the indirect lane is larger or equal to

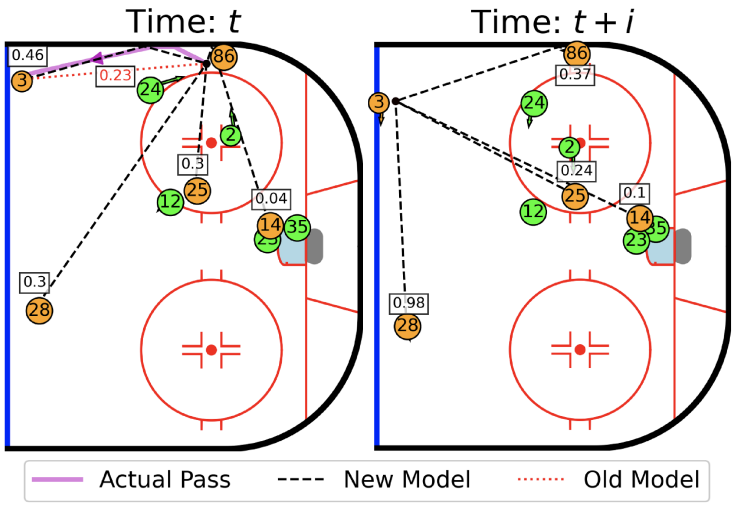


Fig. 8: Powerplay scenario for the Orange team, showing the best passing lanes to each player at times t and $t+i$. At time t (left figure), Player #86 has the puck. Our new passing lane model identifies the 1-bank lane to Player #3 as being the most open (twice as large as the direct lane). Player #86 chooses this lane for their pass (purple line). At time $t+i$ (right figure), after Player #3 receives the pass, the cross-ice lane to #28 increases from 0.3 to 0.98 (a factor of 2.3). Completing this pass is known as an "up-and-over" on the powerplay.

the direct lane defined by our model. Reducing the location error for r in the majority of completed passes (94%) shows that expected movement better aligns with where NHL players pass the puck than when it is not included. However, our analysis has several limitations that are important subjects of future work.

First, since our dataset only contains completed passes, our analysis may not accurately reflect the full behavior of all attempted passes. Another potential application of our model may be to identify incomplete passes based on the movement of the puck; however, this is beyond the scope of this paper.

Second, more accurate time labels for the start and end of passes would improve the precision and scope of future passing models. More accurate time labels would also improve the ability to calculate the speed of passes which has implications on the pass speed model we use for expected movement.

Third, future datasets could allow for more concrete evaluations such as calculating classification accuracy and lead to the development of new models. A ground truth dataset of pass types could be used to evaluate the accuracy of our classification model and allow our system to learn classification thresholds directly from data instead of observing and defining values. Furthermore, a dataset of incomplete passes could help analyze correlations between game context, γ values, and pass completion probabilities.

Fourth, extensions to the current passing lane model could explore a series of different directions. Rim passing lanes are a natural extension of this work. We could further improve the passing lane model to include more advanced methods of expected movement, such as predicting a player's movement with machine learning (i.e, ghosting) [6], physics-based approaches used in soccer [10], or considering handedness, reach, and stick length. When modeling the expected speed of a pass, a future iteration may consider personalized pass profiles by observing previous passes only by a specific player, their location, position, orientation (augmented from a visual dataset since this is not in the PPT data), or type of pass (i.e., Direct, 1-bank, or Rim). Another potential pass classification could be drop passes, which have significantly different dynamics (player movement and puck speed) than most Direct passes. Furthermore, future work can leverage the z coordinate of the puck to analyze who makes *saucer* passes and where, a common pass in hockey that elevates the puck off the ice.

Finally, we would also like to conduct a sensitivity analysis to determine if our classifications are sensitive to d_b, t^θ, and other variables.

9 Conclusions

The new PPT system implemented by the NHL has opened the door for a broader scope of hockey analytics to better model higher resolution events of the game. In this paper, we present an algorithm to classify different types of passes from PPT data and extend the passing lane model in [8] to include 1-bank indirect passes and the expected movement of players. Our model estimates that 1-bank passes comprise about 10.2% of all completed passes in our dataset and make up the majority of non-Direct passes completed. We present *gamma-ratio*, a metric to model the relationship between direct and indirect passing lanes available to a passer. Our model calculates the indirect passing lane to be equal or more available than the direct lane for approximately 59% of completed indirect passes. Furthermore, we show that including the expected movement of players reduces the error in modeling the location of the receiver when they receive the puck for over 94% of completed passes. As PPT systems continue to expand and improve, the impact of algorithms to leverage this type of data will only increase.

Acknowledgments

This research is partially funded by the Natural Sciences and Engineering Research Council of Canada (NSERC), an Ontario Graduate Scholarship, a Cheriton Scholarship, and the University of Waterloo President's Graduate Scholarship. We thank Brant Berglund, Christopher Baker, Keith Horstman, Neil Pierson, and Russell Levine from the National Hockey League Technology, Stats and Information Team for their participation in fruitful discussions and their insights related to this work. In particular we would like to thank Christopher for his timely and insightful comments on drafts of the paper. We thank Jonah

Eisen, Neel Dayal and Oguzhan Cetin from Rogers Communications and Colin Russell and Aaron Pereira from the University of Waterloo, for their help in getting this project off the ground. We especially thank Neel Dayal for his efforts in creating the relationship with the NHL, providing us with access to the dataset and the talented group at the NHL. We also thank Alexi Orchard for her feedback and useful discussions on drafts of this work.

References

1. Anzer, G., Bauer, P.: Expected passes. Data Mining and Knowledge Discovery pp. 1–23 (2022)
2. Bransen, L., Van Haaren, J.: Measuring football players' on-the-ball contributions from passes during games. In: International Workshop on Machine Learning and Data Mining for Sports Analytics. pp. 3–15. Springer (2018)
3. Douglas, A., Kennedy, C.: Tracking in-match movement demands using local positioning system in world-class men's ice hockey. Journal of Strength and Conditioning Research (2019)
4. Fernández, J.: Decomposing the immeasurable sport: a deep learning expected possession value framework for soccer (2019)
5. Goes, F., Kempe, M., Meerhoff, L., Lemmink, K.: Not every pass can be an assist: a data-driven model to measure pass effectiveness in professional soccer matches. Big Data **7 1**, 57–70 (2019)
6. Lindström, P., Jacobsson, L., Carlsson, N., Lambrix, P.: Predicting player trajectories in shot situations in soccer. In: International Workshop on Machine Learning and Data Mining for Sports Analytics. pp. 62–75. Springer (2020)
7. Power, P., Ruiz, H., Wei, X., Lucey, P.: Not all passes are created equal: Objectively measuring the risk and reward of passes in soccer from tracking data. In: Proceedings of the 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. pp. 1605–1613 (2017)
8. Radke, D.T., Radke, D.L., Brecht, T., Pawelczyk, A.: Passing and pressure metrics in ice hockey. AI for Sports Analytics Workshop (2021)
9. Spearman, W.: Beyond expected goals. In: Proceedings of the 12th MIT Sloan Sports Analytics Conference. pp. 1–17 (2018)
10. Spearman, W., Basye, A., Dick, G., Hotovy, R., Pop, P.: Physics-based modeling of pass probabilities in soccer. In: Proceeding of the 11th MIT Sloan Sports Analytics Conference (2017)
11. Steiner, S., Rauh, S., Rumo, M., Sonderegger, K., Seiler, R.: Outplaying opponents—a differential perspective on passes using position data. German Journal of Exercise and Sport Research **49**, 140–149 (2019)
12. Stöckl, M., Seidl, T., Marley, D., Power, P.: Making offensive play predictable-using a graph convolutional network to understand defensive performance in soccer. In: Proceedings of the 15th MIT Sloan Sports Analytics Conference (2021)
13. Vats, K., Fani, M., Walters, P.B., Clausi, D.A., Zelek, J.: Event detection in coarsely annotated sports videos via parallel multi receptive field 1d convolutions. 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW) pp. 3856–3865 (2020)
14. Walkters, P., Fani, M., Clausi, D., Wong, A.: Bendernet and ringernet: highly efficient line segmentation deep neural network architectures for ice rink localization. Journal of Computational Vision and Imaging Systems (2020)