



Killer Defense: Evaluating Individual Defensive Contributions on the Penalty Kill

Frithjof Sanger
Carnegie Mellon University

Ian A. Pérez
University of Arizona

Christina Vu
Texas Christian University

Introduction

Motivation: In hockey, a powerplay gives teams a number advantage and a key chance to score. Since power plays only last a short duration, maintaining possession of the puck is important for the offense, while the defense must disrupt pass attempts.

We decided to analyze power play pass completion probabilities and the impact of individual defenders.

Main questions:

- What factors influence pass completion probability?
- Which defenders are more disruptive on the penalty kill?

Data: Big Data Cup 2022 and 2024 by Stathletes.

- Play-by-play data: 29 games
- Player tracking data (Powerplay): 6 games

Player tracking data cleaned by Alon Harell, Robyn Ritchie, and Phil Shreeves in their Big Data Cup 2022 submission.

Discussion

Limitations: Limited sample of games and tracking data is only available for power plays.

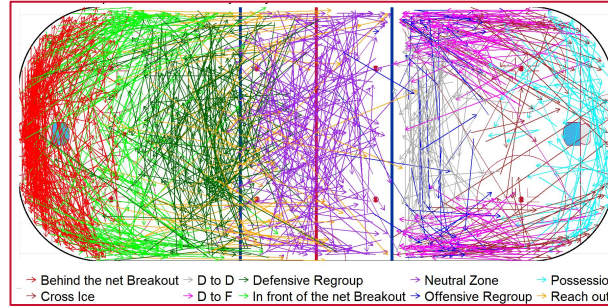
Future work: Analyze more defensive actions, such as:

- Impact of every defender on ice
- Evaluating pass value through expected goals (xG):
 - Higher weigh to dangerous passes defended
 - Analyze defender positioning and actions on dangerous passes

Acknowledgements

Special thanks to our advisor Sam Ventura, our instructor Quang Nguyen and the program TAs for their guidance.

Exploratory Data Analysis



Passes following a Puck Recovery clustered by pass types.

Results & Conclusions



Animation of Finland's power play goal against Switzerland.

- Passes are represented as arrows and labeled with the model's completion probability prediction
- Goal-scoring shot is represented by a green dot.

Methods

Data Processing:

- k-means clustering to identify pass type
- Resampling to handle class imbalance

Pass Completion Probability Model

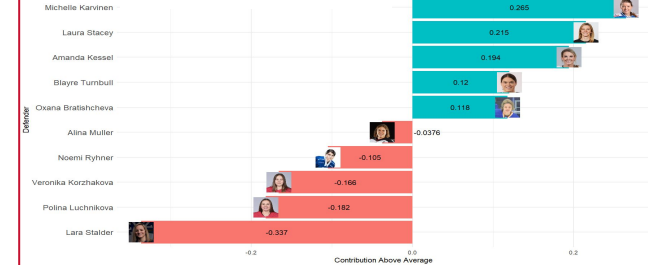
- Feature Engineering:

- Pass distance
- Closest defender distance from passing lane
- Whether a defender was in front of the passer or receiver
- Number of defenders in range of the passing lane
- Pass type

- Modeling:

- Fit and train an XGBoost model

Top 5 and Bottom 5 Defenders by Contribution Above Average



Contribution:

- The difference in pass completion probability when the nearest defender to the passing lane is removed from the ice.

Contribution Above Average:

- The difference between the average contribution of individual defenders and the average contribution of all defenders.

- Built an XGBoost model which assigns success probabilities to passes
- The most impactful variables for pass completion are: Distance of Closest Defender to Passing Lane, Pass distance and Pass type.