

Killer Defense: Evaluating Individual defensive contributions on the penalty kill

by Fritz Sanger, Ian A. Pérez, Christina Vu



Introduction

Penalties

- Offending player is removed from the ice for 2-5 minutes.
 - Advantaged team is “on the Power Play”
 - Disadvantaged team is said “on the Penalty Kill”
 - Significant scoring rate increase

Which defenders are the most disruptive on the penalty kill?

What factors influence pass completion probability?


Data Description



Play by play and Player tracking data from 29 Olympics and Rivalry Series games

- Play-by-play data
 - Number of skaters
 - Time
 - Event (pass, shot, puck recovery, etc)
- Player-tracking data (from TV Broadcast)
 - Jersey number
 - x and y-coordinates
 - Frame number

Manipulation:

- Merged Play-by-Play and Player tracking data
- 

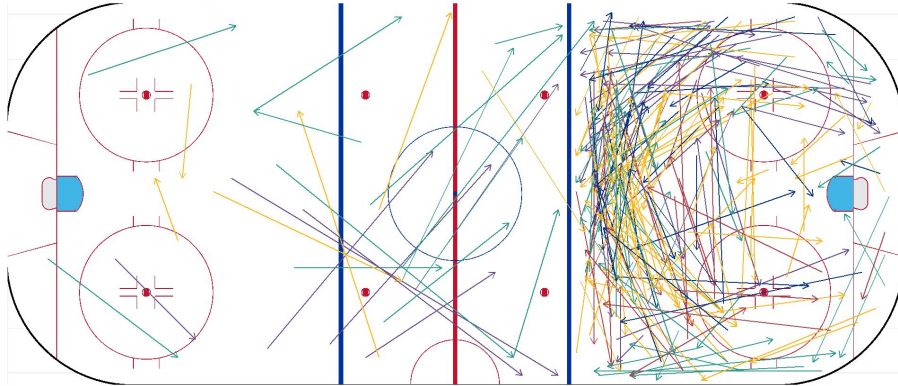
Exploratory Data Analysis (EDA)

Pass completion% depends on the area of the rink

Pass Success by Rink Zone

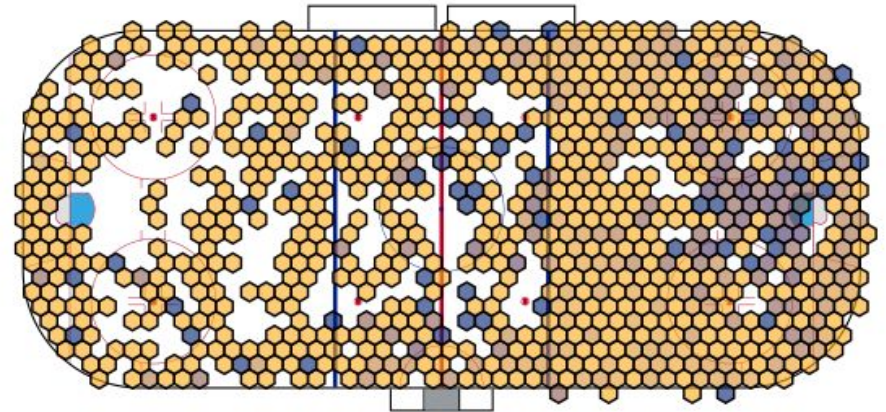
Minimum of 25 Passes Recorded

All Powerplay Passes by Team

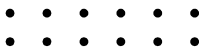


→ Canada → Finland → Russia → Switzerland → United States

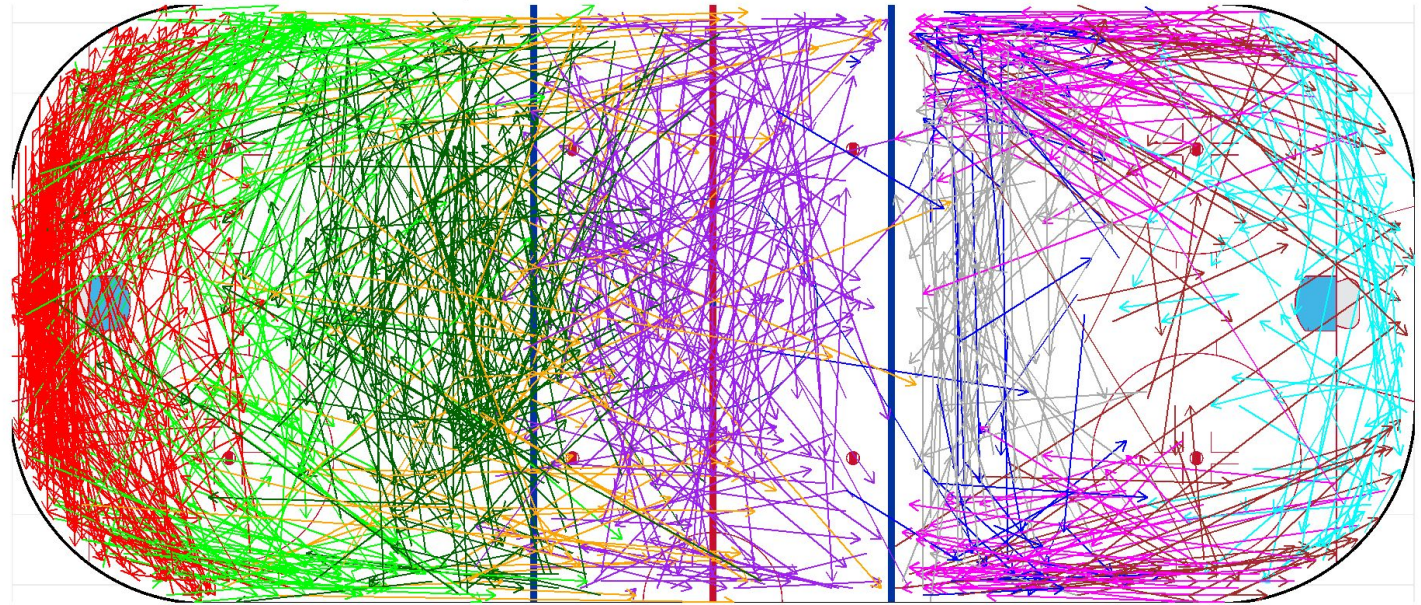
→ Focus in Offensive Zone



Pass Success%
0.00 0.25 0.50 0.75 1.00



Passes following a Puck Recovery

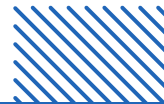
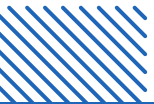


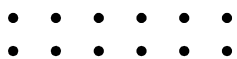
Methods

→ Behind the net Breakout → D to D → Defensive Regroup → Neutral Zone → Possession
 → Cross Ice → D to F → In front of the net Breakout → Offensive Regroup → Reach out

Identify pass patterns →

- k-means Clustering
- 10 clusters
- Normalized coordinates





Pass-completion Probability Model

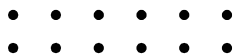
- Predict the probability of a successful pass.
- Analyze factors impacting pass success during power plays.

XGBoost

- Easier for inference
 - > Want to study how each predictor affect the response.

Data Preparation

- Class Imbalance: 90% success rate -> Resampling



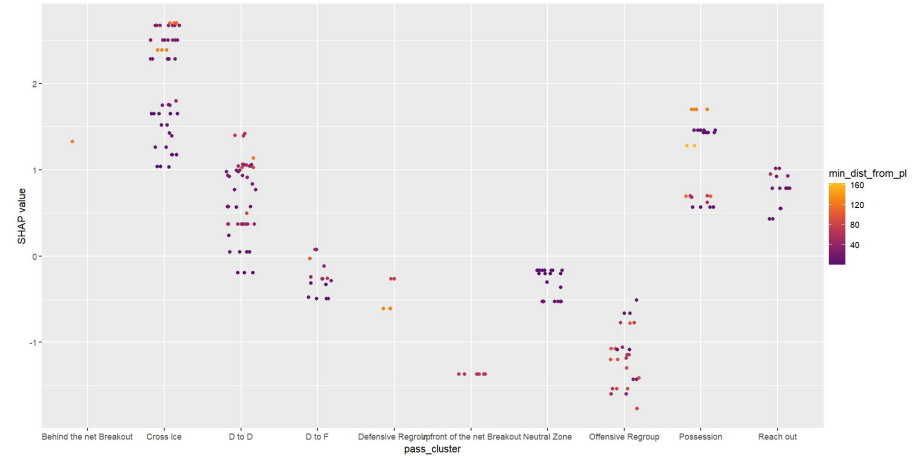
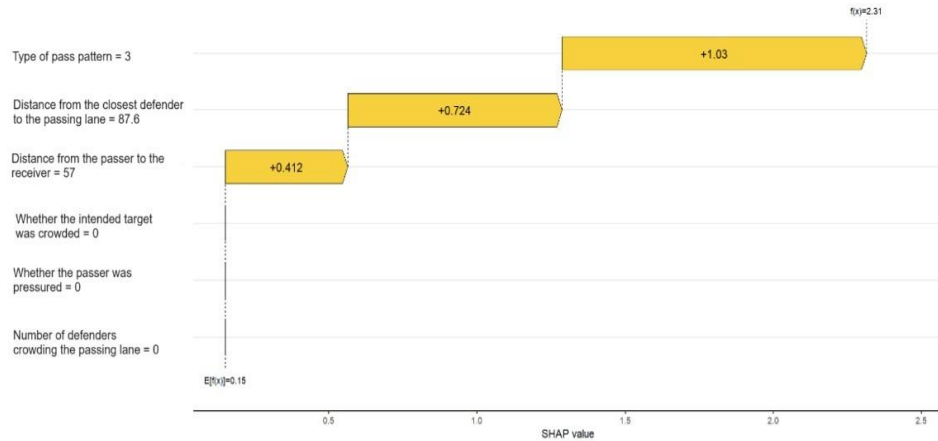
Pass-completion Probability Model

Feature Building and Predictor variables

1. Distance from the closest defender to the passing lane
2. Number of defenders crowding the passing lane
3. Whether the passer was pressured
4. Whether the intended target was crowded
5. Distance from the passer to the receiver
6. Type of pass pattern

Pass-completion Probability Model

SHAP: how each variable contributes to the model's predictions.



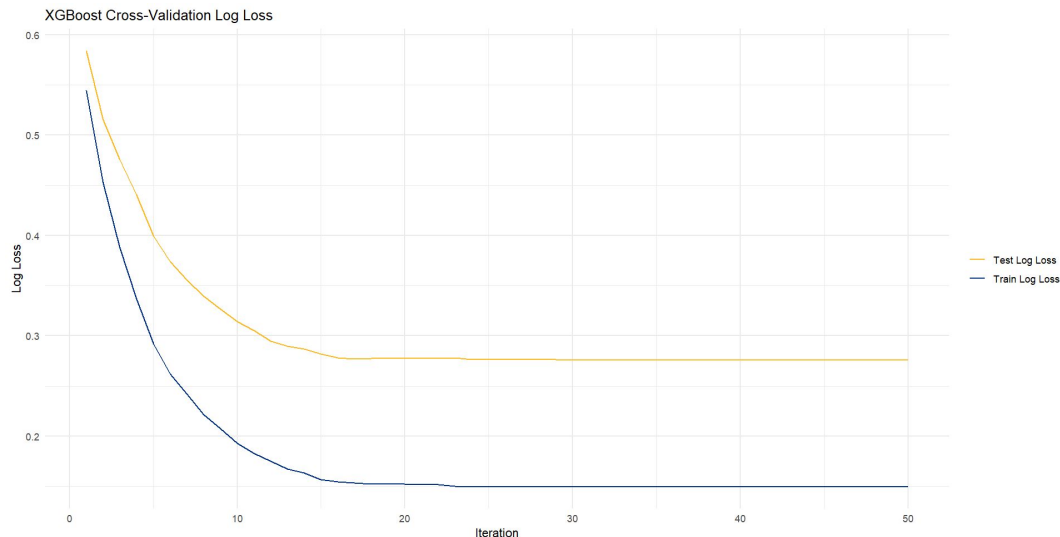
- Only 3 of 6 impact the model.
- All positive impact on success.

- Impact of pass type varies.
- Shorter passes have less impact.

Pass-completion Probability Model

Model Evaluation

- 5-fold Cross Validation



- ROC Curve → AUC: 95.61%

- Confusion Matrix → Accuracy: 89.23%

Results

Expected Pass-Completion Probability (PP)

- The XGBoost model's prediction on success probability of a pass.

Defender Contribution

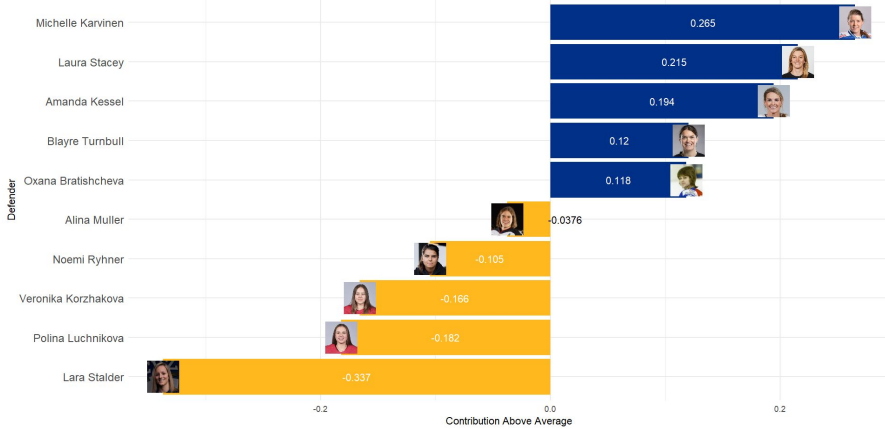
- $PP(\text{nearest defender}) - PP(\text{second nearest defender})$.

Defensive Contribution Above Average

- $\text{average}(\text{Defender Contribution}) - \text{average}(\text{All Defender Contribution})$.

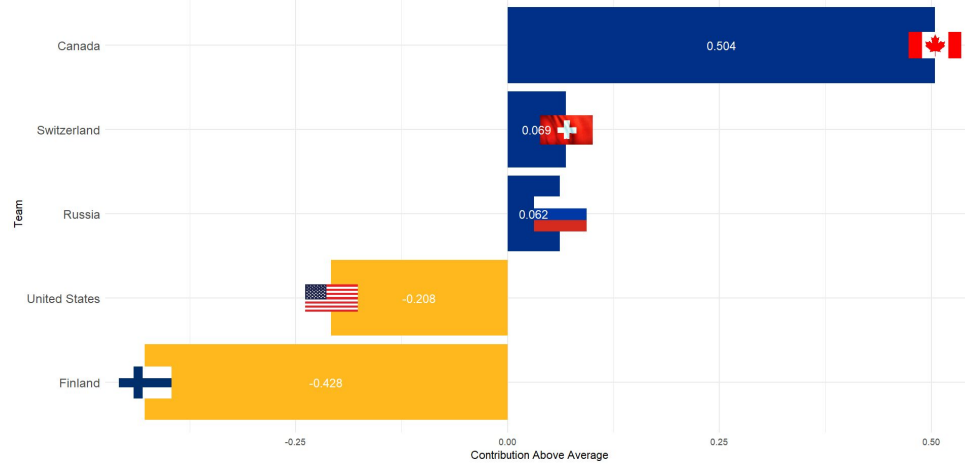
Defensive Contribution Above Average

Top 5 and Bottom 5 Defenders by Contribution Above Average



Note: 4 out of 5 worst defenders are Forwards

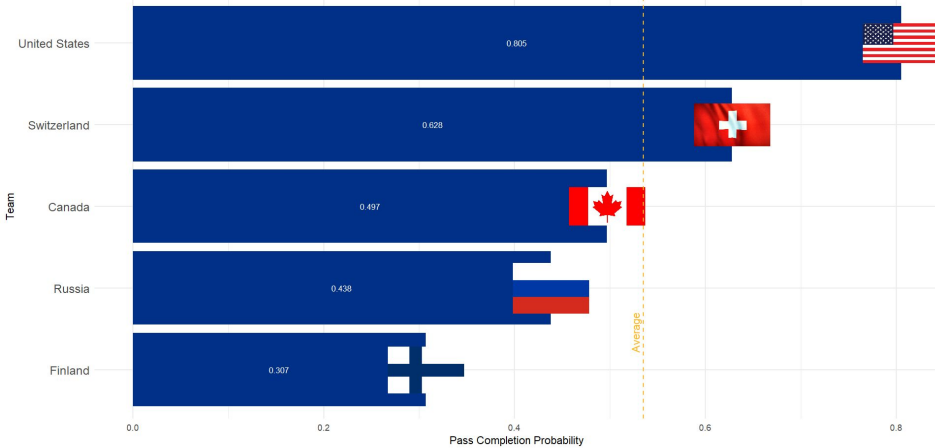
Team Defensive Contribution Above Average



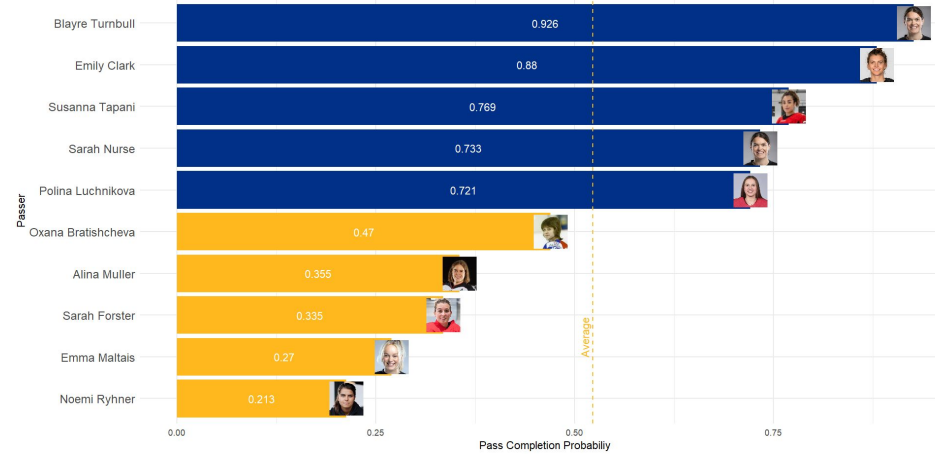
Canada won the 2022 Olympics

“Offensive” Contribution by Pass Completion

Average Pass Completion Probability by Team



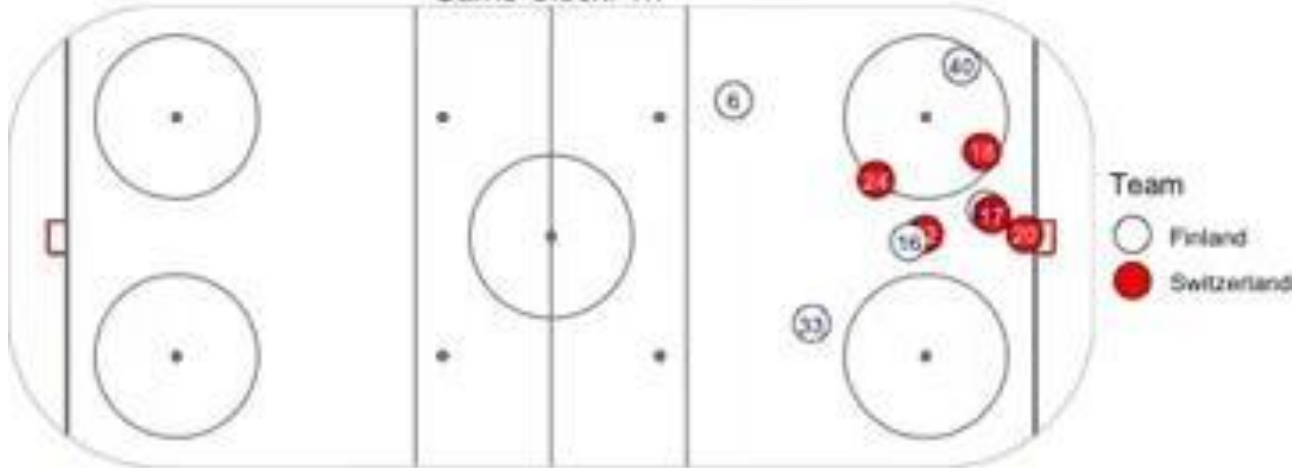
Top 5 and Bottom 5 Passers by Expected Pass Completion Probability



Sarah Nurse was the Assist leader.

Finland Power Play Goal

Game Clock: 1:7



**Expected Pass
Completion
Probability**

Discussion

Conclusion

- **Expected Pass Completion Probability**
 - Speaks to decision-making and aggressiveness of approach
- **Defensive Contribution**
 - Depends on other defenders
 - ie. The defensive “load/responsibility” of a player

Limitations

- Small sample size

Future Work

- Expanding the model to all defenders
- Weighing contribution by pass value