**Expected Total Threat: A Hockey Markov Model to Value All Puck Actions**

1. Introduction

Consider [this possession](https://www.youtube.com/watch?v=kexF0BkIt6I) in the November 13, 2015 UND vs. Miami University college men’s ice hockey game. There are many events in this sequence, all of which contribute in some way to the goal scored. First, UND’s backcheck generates a turnover in their defensive zone. The puck is recovered by Nick Schmaltz (#8) who carries the puck through the neutral zone and controls the puck as he enters the offensive zone. Schmaltz passes the puck cross-ice to Brock Boeser (#16), who promptly passes to Drake Caggiula (#9), who drops the puck to Troy Stecher (#2). Having generated a 3-on-1 deep in the slot, Stecher passes to Boeser who passes cross-crease to Schmaltz for an easy tap-in goal.

The question of how to assign value to each of these events is non-obvious. Prior public research such as Ryan Stimson’s Passing Project and Corey Sznadjer’s research on zone entries have tracked the value of passing and carrying, which can help us understand the value of the fact that Schmaltz’ entry was controlled or acknowledge the fact that the precise tape-to-tape puck control of UND increases their chances of scoring. But how do we assign individual value to each of those passes beyond simply saying “they increased the chances of scoring”? And how can we account for Schmaltz’s efforts in moving the puck both up-ice and Caggiula’s value in the pass to Stecher, setting up the 3-on-1?

This can be answered with the use of an expected threat model; that is, a model that assigns value to on-puck events based on the location-based value the event generates. While an expected threat model is not a novel idea, both in hockey (Forstner & SportLogiq, 2020; de Lara, 2019) and soccer (Singh, 2019), it has not been implemented at the junior level before. Additionally, an expected threat model can help address a central question posed by Stathletes with the scouting dataset; i.e., given a shortened season of 40 games, how can an NHL team properly contextualize and evaluate a potential draft pick or current prospect’s contributions, development, and potential to have an impact at the NHL level?

1. Model Description and Previous Implementations

One solution to this is a model we will call xTT, short for **Expected Total Threat**. xTT is a Markov-based model that takes as inputs all actions on the ice and their locations and assigns value to each location on the ice based on historical data showing the value of plays emanating from that location. xTT divides the rink into “zones” to describe the value in each zone, and for a given zone at some and a total of vertical and horizontal zones, is equivalent to:

which we can break down into three more digestible component parts; the **shot value**, the **move value**, and the **loss value**.

Firstly, describes the “shot value” of a zone; simply the probability of shooting in a zone weighted by the probability of scoring in that zone,.

Secondly, describes the “move value” of a zone via passing or skating; we write this as the probability of moving from our zone times the xTT that would be generated by moving to any other zone with a total of zones, weighted by the likelihood of moving to each other zone.

Finally, defines the “loss value” of a zone, the probability of giving the puck away in a zone weighted by the xTT an opponent would gain from possessing the puck in that zone.

In general, the xTT generated by an action is simply the xTT of the end location of the action minus the xTT of the starting location of an action. A player’s total xTT is simply the sum of the xTT of all on-puck actions by that player, providing a nice succinct number to describe a player’s on-puck contributions.

What makes xTT valuable, as noted by Karun Singh (who implemented what I believe to be the original expected threat model in soccer), is that it “rewards actions independent of the outcome of the possession” as well as rewarding “threatening positions that can lead to high-xG shooting positions with high likelihood”. xTT is calculated by using historical data, but by assigning value on a granular event-level basis, it effectively models the value of each and every event of a possession regardless of how that possession ends. Since this is a model in which the value of every zone depends on the value of all other zones, it may seem difficult to implement. However, as Singh again notes, this model can be formed over several iterations (beginning with an expected threat of zero) until sufficient convergence is achieved, and the xTT in the model is the xTT of the previous iteration.