An AUDL Data Visualization and Analysing App and its Applied Showcase

Haoyang Yan University of Wisconsin - Madison
yanhaoyang493@gmail.com

The main work I've done is developing a comprehensive and exhaustive app for the visualization and analysing of the AUDL data.

The link to the app is here: (It may take some time to initialize)

https://haoyangyan.shinyapps.io/boston_audl/

All the data and code can be found in the Google Drive shared with bostongloryanalytics@gmail.com:

https://drive.google.com/drive/u/0/folders/1h1AUKRZY9muByGindjo0QF0zyTj426Ab

The running order and the way to generate the app can be found in 'README.txt'

These links haven't been published. Please don't share it with unrelated people. Thanks!

Thank Jiaying Ji, who helped me develop the network visualization.

Thank Jeremy C. Weiss and Sean Childers, who firstly proposed the idea of Expected Point Outcome (EPO) map.

Thank Hiro Schmidt, who implemented the first version of UltiMap.

1 Introduction

The App has 4 main parts:

1. Throwing Map. Shows the throwing preference and the completion rate of any selected team or player compared to the league's average. It can also show the map of opponents of a selected team. The player list can be found in the page by filtering team.

2. Passing Network. Shows the main passing network of a selected team and year. The number of edges can be chosen.

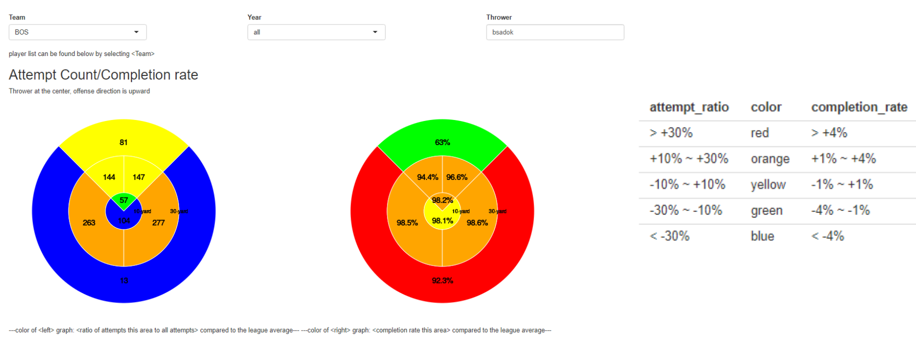
3. Field Map. Includes 4 heat maps of the field. The last 3 maps are interactive by a brush in the first map. The 4 maps show 'Expected Point

Outcome (of a possession)', 'Change of Expected Point Outcome', 'Distribution of where disc go next' and 'Completion rate'. There is a filter for selecting a team. The maps can be smoothed by a checkbox, and it may take some time for smoothing.

4. Analysis. Provides two bar plots and two tables for both offense and defense for a selected team. Two bar plots shows the yardage distribution of throwing and the average change of EPO for different yardage. Two tables compare the count and the EPO of possessions with/without a huck and forcing flick/backhand.

2 Throwing Map

To read this throwing map. I take the core player of Boston Glory, Ben Sadok, as the example.



These two layered pie charts are relative coordinated by the thrower. The thrower is at the center, and the offense direction is upward. The bounds of layers are 10-yard and 30-yard to the center, and the angle dividing forward passes and swings is 45 degree.

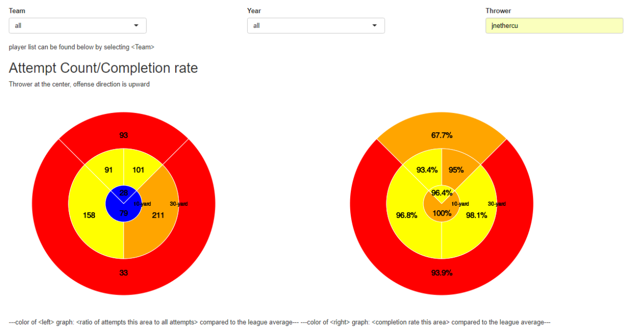
There are totally 8 zones in each graph. 2 in the short range, which are 'tiny throw'(forward) and 'dump'. 4 in the mid range, which are 'flick', 'flick swing', 'backhand swing' and 'backhand' (assume that all the players are righty). 2 in the long range, which are 'huck' and 'cross-field'.

The left graph shows the total throwing attempts of the selected team or player in different zones. The color of left graph means 'the ratio of attempts in this zone to all attempts from this player' compared with the league's average. The right graph shows the average completion rate of the selected team or player in different zones. Color means compared with the league's average. Red means higher while blue means lower.

By the way, after selecting the team, the player's id can be easily found at the bottom of the page.

From the graph above of Ben Sadok, we can get this conclusion. He is a conservative thrower, who throws more swings but less short-range and cross-field throws than the average. His mid-range and cross-field throws are pretty good, but the completion rate of huck still can be improved.

For comparation, I take one of the most characteristic handlers in the league, Jon Nethercutt, as another example.

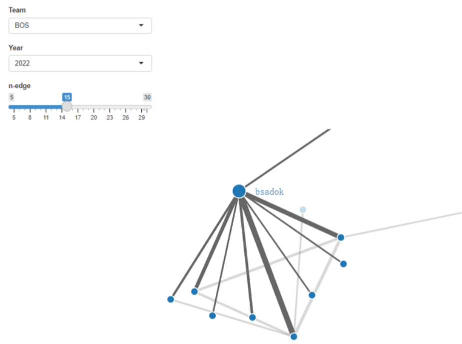


From the graph above of Jon Nethercutt, we can see that he's definite a risk-preferred thrower, who throws lots of long-range throws, including huck and cross-field, and he's really good at it.

The throwing map can also show all opponents' information of a selected team, which may give us the idea about its defense preference and pattern. Here is the map of Boston Glory's opponent.

It shows that Boston did a great job on limiting its opponents' short-range throws and backhand (probable breakside?) throw. However, Boston's opponents throw much more cross-field throws than average (zone defense?) with a high completion rate, and the completion rate of huck is also high. In summary, Boston is probable a team with tight and radical defense style.

3 Passing Network



The graph shows the main Passing Network of Boston Glory. The number of edges can be adjusted by the slider input. The width of edges means the count of total connections between two players, and we can see the player's id by clicking the node.

From this graph we can know that Ben Sadok is obvious the core player, and Cole Davis-Brand is the second most important player on the handling side.

4 Field Map

In this part, I have to introduce an important concept called the Expected Point Outcome (EPO). By dividing the field by each 5-yard on both x and y axis, we get 24*11=264 blocks. From the 16k throws (including pass, goal, throwaway, drop, callahan) in the whole dataset, I detect that for the possession this throw in, whether it ends with a goal. By calculating the ratio of 'with a goal' to the total counts of throws group by the block where the throw is from, I get the Expected Point Outcome of each block. I also add a filter of teams to generate different EPO maps.

$$EPO_i = \frac{\sum if(score\_possession)\{T_i\}}{\sum T_i}$$

i: block, T: throws set, if true return 1

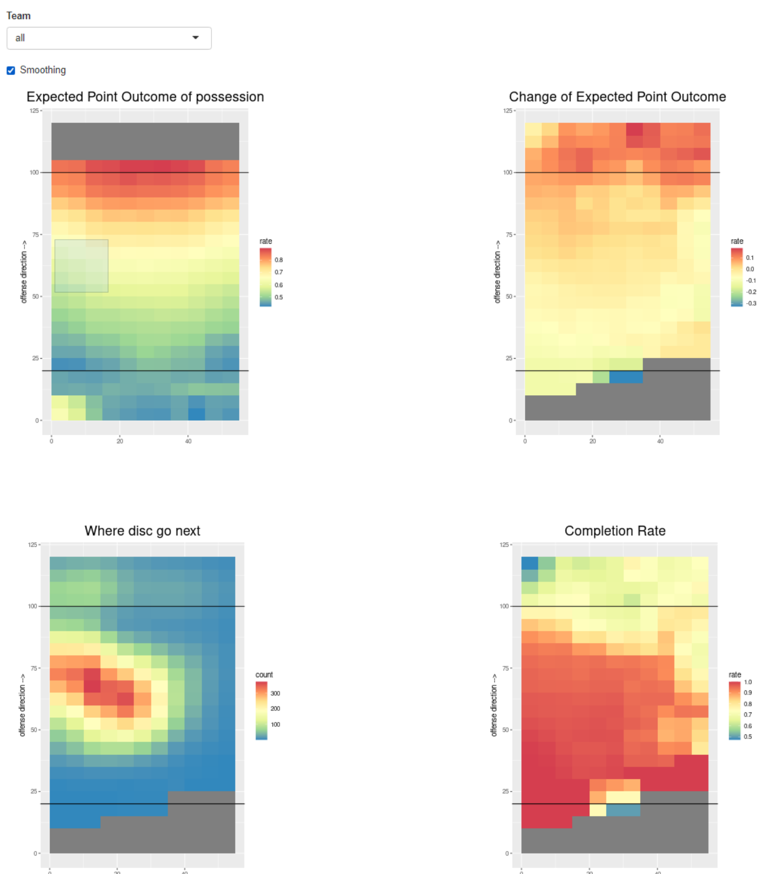
The next question is, given a brush in the map (which means disc already in that area), throwing to where is a probable good choice? To answer this question, I develop two interactive maps of 'Distribution of where disc go next' and 'Completion rate' by brushing an area.

To find out the best balance between risk and return, I finally plot the 'change of EPO' map. It shows, given an area where disc is in, what's the expected change of EPO when attempting to throw somewhere. Considering both completion and turn, the formula should be as follow.

$$
\begin{aligned}
C\_EPO_{ij} &= \mathbb{P}(comp) * EPO_j + \mathbb{P}(turn) * EPO_j^* - EPO_i \\
&= \mathbb{P}(comp) * EPO_j + \mathbb{P}(turn) * (1 - EPO_j^{-1}) * conv - EPO_i
\end{aligned}
\tag{1}
$$

EPO_j^{-1}: EPO of the mirrored block of j
$conv$: Average conversion rate of the league, which means the EPO of a possession without any prior information

Considering that there may be few samples from block i to j, the maps have some fluctuation. The first version of UltiMap developed by Hiro used k-nn algorithm to reduce randomness. It cannot calculate real time with a filter and brush due to the large dataset. So, I choose a simpler way of getting average of the 9 surrounding blocks to smooth the map.



5 Analysis

It is easy to understand from the app. This margin is too narrow to contain.