VALUING ACTIONS AND RANKING HOCKEY PLAYERS WITH MACHINE LEARNING

Oliver Schulte
School of Computing Science
Simon Fraser University
Vancouver, Canada
OVERVIEW

Success Probabilities

Defining Success Probabilities

A fundamental problem in sports analytics

Action Values

Success Probabilities \(\downarrow\) Action Values

Player Ranking

Action Values \(\downarrow\) Player Ranking

Learning Success Probabilities

Supervised (Classification)

Unsupervised (Reinforcement Learning)
WHAT IS SUCCESS?

- An outcome (binary event) that a team wants to bring about.
- Can be defined according to the interest of the analyst/coach/player.
SUCCESS PROBABILITY TICKER

• Assigns to each time $t$ in the match an estimated probability of future success
• Example: Win probability in NHL (Pettigrew 2015)

S. Pettigrew, “Assessing the offensive productivity of NHL players using in-game win probabilities,” 2015. 9th Annual MIT Sloan Sports Analytics Conference
Y-axis: the chance of scoring the next goal
SCORE IN POSSESSION

http://www.lukebornn.com/sloan_epv_curve.mp4
FROM SUCCESS PROBABILITIES TO ACTION VALUES
ACTION VALUES

• Success probabilities can be used to evaluate players and actions
  • $\text{impact(action)} = \left[\text{success probability after action} - \text{success probability before action}\right]$
  • “We assert that most questions that coaches, players, and fans have about basketball, particularly those that involve the offense, can be phrased and answered in terms of EPV [i.e. expected future success].” Cervone, Bornn et al. 2014.
WHY SUCCESS PROBABILITIES

• Advantages of success impact for action values?
  1. **Look-ahead**: consider the medium-term indirect impact of an action
     • pass → assist → shot → goal
  • **Context-awareness**: incorporate match context
  • Distinguish high probability of success vs. **actual** success
     • (cf. Xgoal vs. actual goal)
  • All actions are evaluated on the **same scale** (probabilities)
  • Example: *Olympics 2010 Golden Goal*
THE IMPACT OF AN ACTION

- The impact of an action $a_t$ performed at time $t>0$ is the difference in successive success probabilities:

$$\text{impact}(a_t) = p_t - p_{t-1}$$

* = THOR baseline
FROM ACTION IMPACT TO PLAYER RANKINGS
RESULTS 2013-2014 SEASON NHL

<table>
<thead>
<tr>
<th>Name</th>
<th>Goal Impact</th>
<th>Points</th>
<th>+/-</th>
<th>Salary</th>
</tr>
</thead>
<tbody>
<tr>
<td>Jason Spezza</td>
<td>29.64</td>
<td>66</td>
<td>-26</td>
<td>$5,000,000</td>
</tr>
<tr>
<td>Jonathan Toews</td>
<td>28.75</td>
<td>67</td>
<td>25</td>
<td>$6,500,000</td>
</tr>
<tr>
<td>Joe Pavelski</td>
<td>27.20</td>
<td>79</td>
<td>23</td>
<td>$4,000,000</td>
</tr>
<tr>
<td>Marian Hossa</td>
<td>26.12</td>
<td>57</td>
<td>26</td>
<td>$7,900,000</td>
</tr>
<tr>
<td>Patrick Sharp</td>
<td>24.43</td>
<td>77</td>
<td>12</td>
<td>$6,500,000</td>
</tr>
<tr>
<td>Sidney Crosby</td>
<td>24.23</td>
<td>104</td>
<td>18</td>
<td>$12,000,000</td>
</tr>
<tr>
<td>Claude Giroux</td>
<td>23.89</td>
<td>86</td>
<td>7</td>
<td>$5,000,000</td>
</tr>
<tr>
<td>Tyler Seguin</td>
<td>23.89</td>
<td>84</td>
<td>16</td>
<td>$4,500,000</td>
</tr>
</tbody>
</table>

Player performance =

- total impact of all actions performed

*Jason Spezza*: high goal impact, low +/-.
- plays very well on poor team (Ottawa Senators 2013).
• 2015-16 NHL season
• Johnny Gaudreau and Mark Scheifele drew salaries below what their GIM rank would suggest.
• Later they received a $5M+ contract for the 2016-17 season.
In the 2016-17 season, we find many underestimated players:
- high impact but low 2015-16 salary.
- The percentage of players who are undervalued decreases in the 2016-17.
- (3) This suggests that GIM provides an early signal of a player's value.
CORRELATIONS WITH STANDARD STATS

- GIM: our ranking (goal impact metric)
- Takeaway: high correlation with standard stats
  - e.g. 0.93 with points

\[
\begin{array}{cccccccc}
\text{methods} & \text{Assist} & \text{Goal} & \text{GWG} & \text{OTG} & \text{SHG} & \text{PPG} & \text{S} \\
\hline
+/- & 0.236 & 0.204 & 0.217 & 0.16 & 0.095 & 0.099 & 0.118 \\
GAR & 0.527 & 0.633 & 0.552 & 0.324 & 0.191 & 0.583 & 0.549 \\
WAR & 0.516 & 0.652 & 0.551 & 0.332 & 0.192 & 0.564 & 0.532 \\
\hline
EG & 0.783 & 0.834 & 0.704 & 0.448 & 0.249 & 0.684 & 0.891 \\
SI & 0.869 & 0.745 & 0.631 & 0.411 & 0.27 & 0.591 & 0.898 \\
GIM-T1 & 0.873 & 0.752 & 0.682 & 0.428 & 0.291 & 0.607 & 0.877 \\
GIM & \textbf{0.875} & \textbf{0.878} & \textbf{0.751} & \textbf{0.465} & \textbf{0.345} & 0.71 & \textbf{0.912} \\
\hline
\text{methods} & \text{Point} & \text{SHP} & \text{PPP} & \text{FOW} & \text{P/}GP & \text{TOI} & \text{PIM} \\
\hline
+/- & 0.237 & 0.159 & 0.089 & -0.045 & 0.238 & 0.141 & 0.049 \\
GAR & 0.622 & 0.226 & 0.532 & 0.16 & 0.616 & 0.323 & 0.089 \\
WAR & 0.612 & 0.235 & 0.531 & 0.153 & 0.605 & 0.331 & 0.078 \\
\hline
EG & 0.854 & 0.287 & 0.729 & 0.28 & 0.702 & 0.722 & 0.354 \\
SI & 0.869 & 0.37 & 0.707 & 0.185 & 0.655 & \textbf{0.955} & \textbf{0.492} \\
GIM-T1 & 0.902 & 0.384 & 0.736 & 0.288 & 0.738 & 0.777 & 0.347 \\
GIM & \textbf{0.93} & \textbf{0.399} & \textbf{0.774} & \textbf{0.295} & \textbf{0.749} & 0.835 & 0.405 \\
\end{array}
\]
DRILLING DOWN ON PLAYER STRENGTHS

- Where and when do players achieve their highest impact?
- Drill-down analysis compares to average player
- Example: Erick Karlsson manages puck receptions in a high impact region 37.5% of the time.
  - Average player only 14.6%
- Uses rink discretization

LEARNING SUCCESS PROBABILITY MODELS

Try this at home
MACHINE LEARNING APPROACH

Deployment

Match Context at time $t$ → Context Feature Vector at time $t$ → Classifier → Success Probability at time $t$

Training

Observed Match Contexts → Observed Context Feature Vectors → Machine Learning → Classifier
EVENT DATA

• Illustrate approaches with event data
  • available from nhl.com
  • also pre-crawled
• Less work on tracking data
<table>
<thead>
<tr>
<th>Number of Teams</th>
<th>30</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of Players</td>
<td>2,233</td>
</tr>
<tr>
<td>Number of Games</td>
<td>1,140</td>
</tr>
<tr>
<td>Number of Events</td>
<td>3.3M</td>
</tr>
</tbody>
</table>
SPECIFYING LOOK-AHEAD

• Suppose we start with event data

• For each time $t$, add a success column $Y_t$ depending on whether the team succeeded after time $t$.
  • E.g., did they score the next goal?

• Could also annotate whether they score the next goal in $k$ steps (Descrooset al. 2019) or fixed time interval (Shuckers and Curro 2013 THoR)

---


M. Schuckers and J. Curro, “Total Hockey Rating (THoR): A comprehensive statistical rating of National Hockey League forwards and defensemen based upon all on-ice events,” 2013. 7th Annual MIT Sloan Sports Analytics Conference
<table>
<thead>
<tr>
<th>game</th>
<th>player</th>
<th>Period</th>
<th>team</th>
<th>x</th>
<th>y</th>
<th>Manpower</th>
<th>Action</th>
<th>( Y = \text{next goal} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>849</td>
<td>402</td>
<td>1</td>
<td>15</td>
<td>-9.5</td>
<td>1.5</td>
<td>Even</td>
<td>Recovery</td>
<td>0</td>
</tr>
<tr>
<td>849</td>
<td>402</td>
<td>1</td>
<td>15</td>
<td>-24.5</td>
<td>-17</td>
<td>Even</td>
<td>Carry</td>
<td>0</td>
</tr>
<tr>
<td>849</td>
<td>417</td>
<td>1</td>
<td>16</td>
<td>-75.5</td>
<td>-21.5</td>
<td>Even</td>
<td>Check</td>
<td>1</td>
</tr>
<tr>
<td>849</td>
<td>402</td>
<td>1</td>
<td>15</td>
<td>-79</td>
<td>-19.5</td>
<td>Even</td>
<td>Pass</td>
<td>0</td>
</tr>
<tr>
<td>849</td>
<td>413</td>
<td>1</td>
<td>16</td>
<td>-92</td>
<td>-32.5</td>
<td>Even</td>
<td>Turnover</td>
<td>1</td>
</tr>
<tr>
<td>849</td>
<td>413</td>
<td>1</td>
<td>16</td>
<td>-92</td>
<td>-32.5</td>
<td>Even</td>
<td>Pass</td>
<td>1</td>
</tr>
<tr>
<td>849</td>
<td>389</td>
<td>1</td>
<td>16</td>
<td>-98</td>
<td>0</td>
<td>Even</td>
<td>Goal</td>
<td>1</td>
</tr>
</tbody>
</table>
• Context can be represented as a feature vector
  • E.g. score differential, manpower differential
• What do to about the previous match history?
• Simple Approach:
  • Fix a sliding window size $k$ (common values are 3,4,10).
  • Use previous $k$ events context for current event
# Example: Sliding Window

<table>
<thead>
<tr>
<th>game</th>
<th>player</th>
<th>Period</th>
<th>team</th>
<th>x</th>
<th>y</th>
<th>Manpower</th>
<th>Action</th>
<th>Y=next goal</th>
</tr>
</thead>
<tbody>
<tr>
<td>849</td>
<td>402</td>
<td>1</td>
<td>15</td>
<td>-9.5</td>
<td>1.5</td>
<td>Even</td>
<td>Recovery</td>
<td>0</td>
</tr>
<tr>
<td>849</td>
<td>402</td>
<td>1</td>
<td>15</td>
<td>-24.5</td>
<td>-17</td>
<td>Even</td>
<td>Carry</td>
<td>0</td>
</tr>
<tr>
<td>849</td>
<td>417</td>
<td>1</td>
<td>16</td>
<td>-75.5</td>
<td>-21.5</td>
<td>Even</td>
<td>Check</td>
<td>1</td>
</tr>
<tr>
<td>849</td>
<td>402</td>
<td>1</td>
<td>15</td>
<td>-79</td>
<td>-19.5</td>
<td>Even</td>
<td>Pass</td>
<td>0</td>
</tr>
<tr>
<td>849</td>
<td>413</td>
<td>1</td>
<td>16</td>
<td>-92</td>
<td>-32.5</td>
<td>Even</td>
<td>Turnover</td>
<td>1</td>
</tr>
<tr>
<td>849</td>
<td>413</td>
<td>1</td>
<td>16</td>
<td>-92</td>
<td>-32.5</td>
<td>Even</td>
<td>Pass</td>
<td>1</td>
</tr>
<tr>
<td>849</td>
<td>389</td>
<td>1</td>
<td>16</td>
<td>-98</td>
<td>0</td>
<td>Even</td>
<td>Goal</td>
<td>1</td>
</tr>
</tbody>
</table>

$k=2$
**EXAMPLE: SLIDING WINDOW**

\[ k=2 \]

<table>
<thead>
<tr>
<th>Manpower</th>
<th>Action</th>
<th>Y=next goal</th>
</tr>
</thead>
<tbody>
<tr>
<td>Even</td>
<td>Turnover</td>
<td>1</td>
</tr>
<tr>
<td>Even</td>
<td>Pass</td>
<td>1</td>
</tr>
<tr>
<td>Even</td>
<td><strong>Goal</strong></td>
<td>1</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>MP(-2)</th>
<th>Action(-2)</th>
<th>MP(-1)</th>
<th>Action(-1)</th>
<th>Manpower(0)</th>
<th>Action(0)</th>
<th>Y=next goal</th>
</tr>
</thead>
<tbody>
<tr>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>Even</td>
<td>Turnover</td>
<td>1</td>
</tr>
<tr>
<td>*</td>
<td>*</td>
<td>Even</td>
<td>Turnover</td>
<td>Even</td>
<td>Pass</td>
<td>1</td>
</tr>
<tr>
<td>Even</td>
<td>Turnover</td>
<td>Even</td>
<td>Pass</td>
<td>Even</td>
<td><strong>Goal</strong></td>
<td>1</td>
</tr>
</tbody>
</table>
TRAINING A CLASSIFIER

- Given a list of pairs (context vector, success target), can simply run any classifier (R/Weka/Scikit-Learn)
  - E.g. logistic regression, gradient boosted decision trees
- Excerpt from logistic model tree
ALTERNATIVE: RECURRENT NEURAL NETWORK

- Can also use RNN (e.g. Long Short-Term Memory)
  + Natively handles sequence data; no need for preprocessing
  - Less familiar than classification packages
  - Results more difficult to interpret
- Try to get the best of both worlds with *mimic learning*:
  1. Train a neural network to be accurate
  2. Train a tree model to mimic the neural network -> interpretable

REINFORCEMENT LEARNING
BIG PICTURE: SPORTS ANALYTICS MEETS REINFORCEMENT LEARNING

- Reinforcement Learning: Learning to act
- Studied since the 1950s
  - Many models, theorems, algorithms, software.
- **Learning success probabilities is one of the fundamental problems of reinforcement learning**
A problem with the classifier approach is that it (implicitly) assumes that all events are independent.

The probability of scoring of success at $t+1$ is treated as independent of the probability of success at $t$.

Ignores temporal dynamics.

In fact success chances are highly correlated.

RL aims to exploit temporal connections.
STATE TRANSITION PROBABILITIES

• Step 1: Estimate the probabilities of getting one from match state to the other

• Basketball Demo

• In our discrete NHL models, we estimated state transition probabilities for 1.3M states

• Step 2: Estimate the chances of reaching a success state using dynamic programming
MULTI-STEP TRANSITION PROBABILITIES

To compute $P_{t+1}(s_1|s_0)$: the probability of reaching state $s_1$ from $s_0$

\[ P_{t+1}(s_1|s_0) = 0.1 \times 0.2 + 0.3 \times 0.5 + 0.6 \times 0.3 \]
DYNAMIC PROGRAMMING

• **Input**: State Transition Probabilities
• **Output**: Probability of Future Success for every match state
• For lookahead \( L = 1, \ldots \)
  • Compute probability of success in \( L+1 \) steps using
    1-step state transitions and \( L \) step success probabilities from previous lookahead
  • Terminate at convergence or at fixed bound
• For the NHL, our computation converged at \( L = 13 \)
• **Xthreat Visualization**
LEARNING SUCCESS PROBABILITIES: OVERVIEW

1. estimate transition probabilities
2. apply dynamic programming

- Hockey
  Schulte et al. 2017
  Ljung et al. 2019

- NFL
  Chan and Puterman 2019

- continuous data space + time
  discretize

- raw data
  classifier or reinforcement learning

- sliding window
  Decroos et al. 2019
  recurrent NN

- neural net
  temporal difference learning
  model-free

- Hockey
  Liu et al. 2018

- Soccer
  Fernandez et al. 2019
EXTENSIONS AND FUTURE DIRECTIONS

• Who is acting? → Context
  • E.g. a shot by Scheifele is different from shot by average NHL player
    ➢ Player embeddings (Liu et al. 2020)
• Evaluate lines/formations
  • Pairs of player evaluation by Ljung et al. 2018
• Try to infer team objectives (Inverse reinforcement learning)
  Luo et al. 2020
• Future: Determine optimal actions for a given match context

CONCLUSION

• Estimating Success Probabilities is a basic problem in hockey analytics
• Can be applied to solve other problems
  • Action values
  • Player ranking
• Machine learning models can include rich match contexts to provide useful success probabilities
• Reinforcement learning is especially suitable for handling complex dynamic domains like ice hockey
THANK YOU!

Kurt Routley  Zeyu Zhao  Guiliang Liu  Pascal Poupart