Deep Reinforcement Learning in Ice Hockey for Context-Aware Player Evaluation

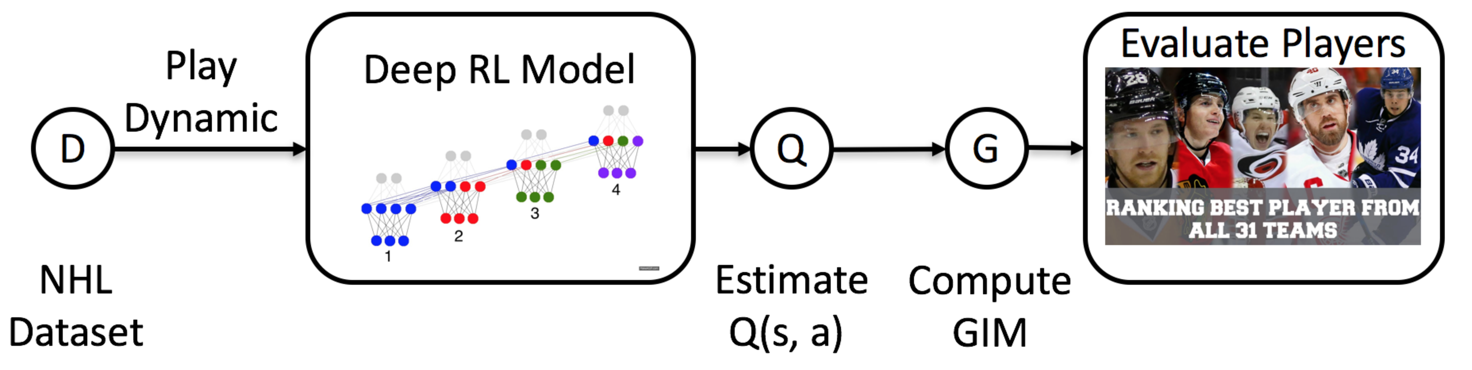
Guiliang Liu and Oliver Schulte
Simon Fraser University, Burnaby, BC, Canada

Overview

- **Goal:** Evaluate the National Hockey League (NHL) players with video tracking datasets.
- **Motivation:** (1) Most datasets have partial observability, (e.g. cover only half of court at a time). (2) Traditional sports models consider limited actions and game context.
- **Our Work:** (1) We apply Deep Reinforcement Learning (DRL) to learn an action-value Q function. (2) We develop a novel Game Impact Metric (GIM) that aggregates the impact of players' actions.

Problem Formulation

- **System Flow** for Player Evaluation.



- **Dataset:** The data contains game events and player actions for the entire 2015-2016 NHL season, which contains 3,382,129 events, covering 30 teams, 1140 games and 2,233 players.

Name	Type	Range
X Coordinate of Puck	Continuous	[-100, 100]
Y Coordinate of Puck	Continuous	[-42.5, 42.5]
Velocity of Puck	Continuous	(-inf, +inf)
Game Time Remain	Continuous	[0, 3600]
Score Differential	Discrete	(-inf, +inf)
Manpower Situation	Discrete	{EV, SH, PP}
Event Duration	Continuous	[0, +inf)
Action Outcome	Discrete	{successful, failure}
Angle between puck and goal	Continuous	[−3.14, 3.14]
Home or Away Team	Discrete	{Home, Away}

The dataset is provided by SPORTLOGiQ (http://sportlogiq.com/en/about/) using computer vision techniques.

Play Dynamic in NHL

Construct **Markov Game framework**

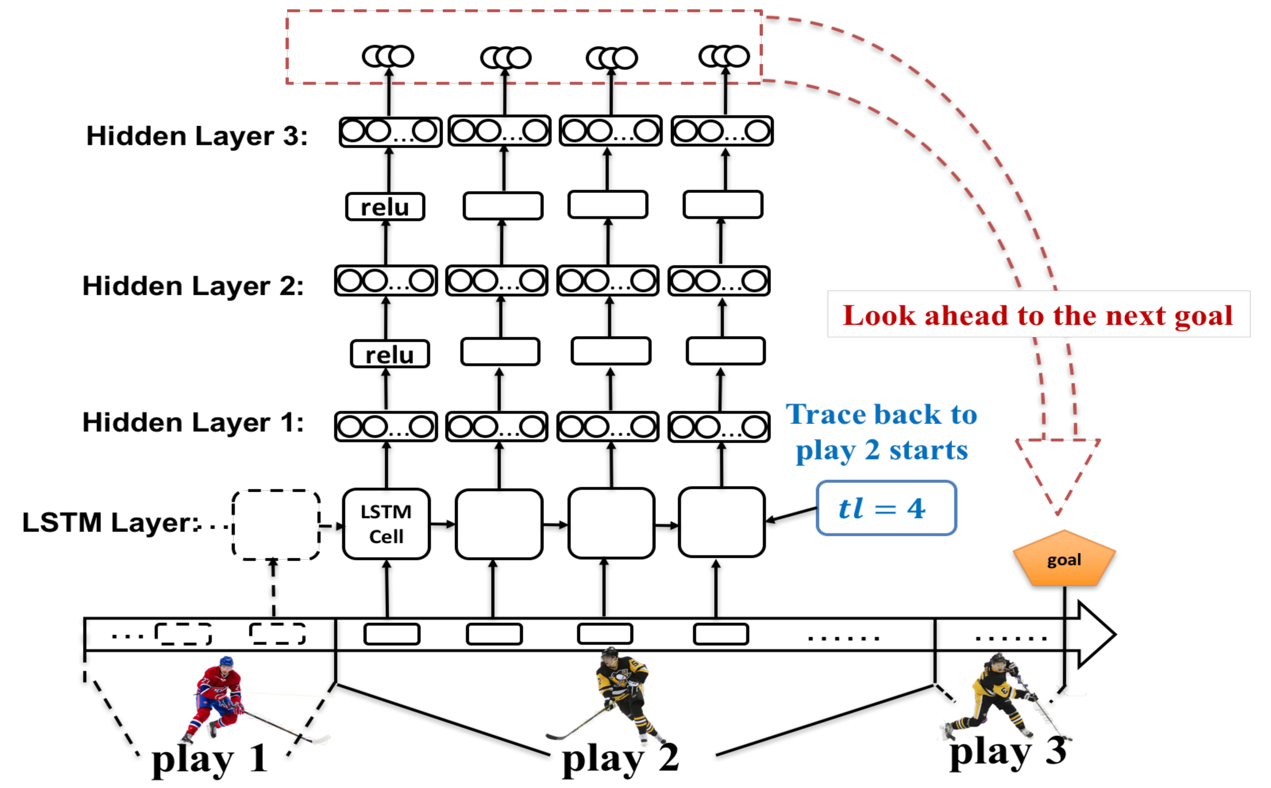
- *Reward* $\mathbf{g_t}$ is a vector that specifies which team scores
- *Action* $\mathbf{a_t}$ is one of 13 types, including shot, assist, etc.
- *Sequence (state)* $\mathbf{s_t}$: We use the complete sequence $s_t \equiv (x_t, a_{t-1}, x_{t-1}, x_0)$ as the state representation at time step t, where observation x_t is a feature vector.

A **Q function** represents the conditional probability of the home resp. away team scores the goal at the end of the current episode:

$$Q_{team}(s,a) = P(goal_{team} = 1 | s_t = s, at = a)$$

Network Architecture

- Our neural net has three output nodes: the estimated $Q_{home}(s,a)$, $Q_{away}(s,a)$ and $Q_{neither}(s,a)$.



Player Evaluation

We define $impact(s,a)$ as the change in Q-value due to a players action. The total impact of a players actions is his Goal Impact Metric (GIM).

$$impact^{team}(s_t, a_t) = Q^{team}(s_t, a_t) - Q^{team}(s_{t-1}, a_{t-1})$$

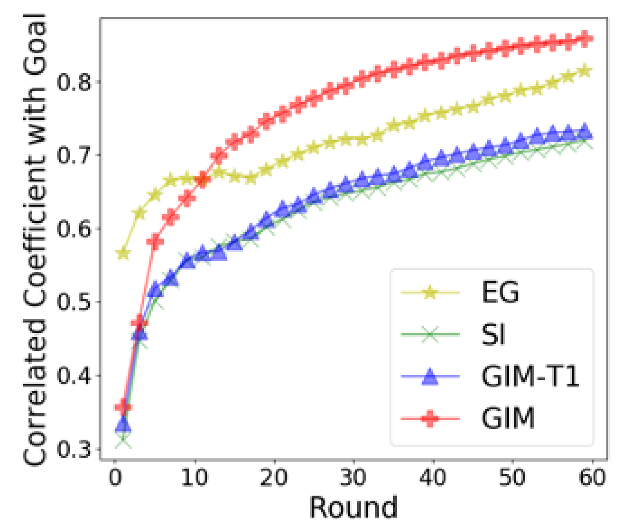
$$GIM^i(D) = \sum_{s,a} n^i_D(s,a) \times impact^{team_i}(s,a)$$

Empirical Evaluation

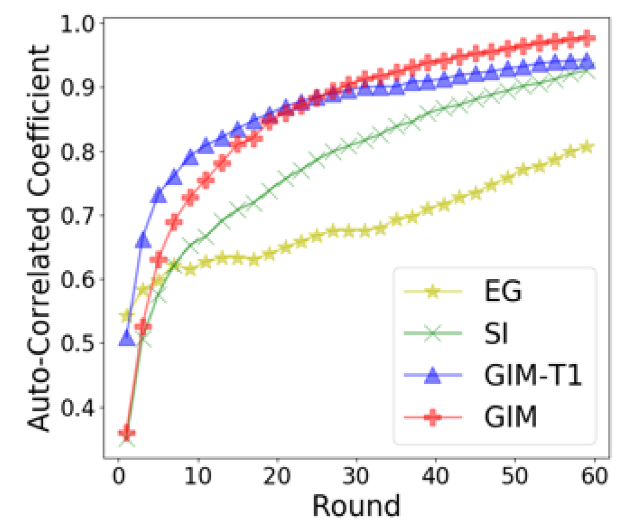
- **Comparison Metrics**
  (a) *Statistics based*: Plus-Minus, Goal/Win Above Replacement (GAR/WAR), and Expected Goal (EG).
  (b) *AI based*: Scoring Impact (SI), GIM with LSTM trace length = 1 (GIM-T1).
- **Season Totals**: Compute the correlation between evaluation metrics and standard success measures.

methods	Point	SHP	PPP	FOW	P/GP	TOI	PIM
+/-	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
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	**0.955**	**0.492**
GIM-T1	0.902	0.384	0.736	0.288	0.738	0.777	0.347
GIM	**0.93**	**0.399**	**0.774**	**0.295**	**0.749**	0.835	0.405

- **Round-by-Round Correlations**: Measure the correlation between the value computed over *the first n round* and computed over *the entire season*.

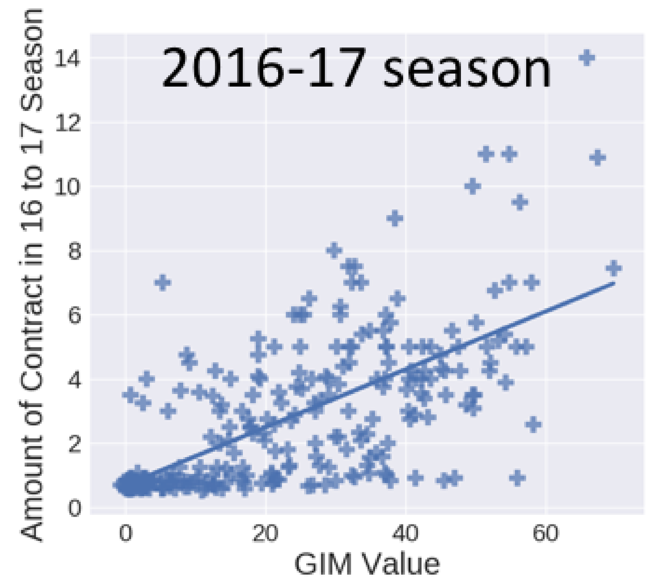


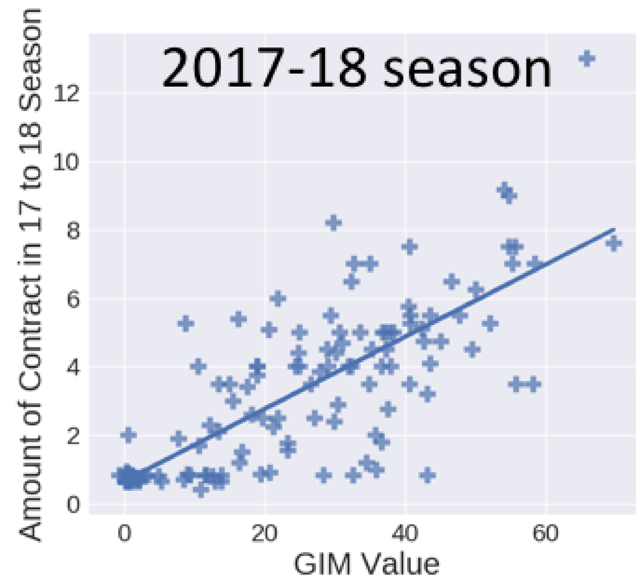
Correlation with Goal



Auto-correlation

- **Future Seasons**: Predict the players future salary with GIM metric and study the undervalued players.





the 27th International Joint Conference on Artificial Intelligence and the 23rd European Conference on Artificial Intelligence