Elo Ratings in the NHL
Potential player and team applications

Roman Parparov, morehockeystats.com

Elo ratings – prediction of outcomes based on ratings difference

For two competitors with Elo ratings R_1 and R_2, the expected result X_1 for R_1 on scale [0,1] is:

$$X_1 = \frac{1}{1 + 10^{\frac{R_2-R_1}{400}}}$$

And then the expected result X_2 for R_2 is:

$$X_2 = 1 - X_1$$

If R_1 or R_2 are not defined, assign an initial value (**2500** in our case)

Then given the actual results A_1 of competitor R_1 and A_2 of competitor R_2, their new ratings become:

$$R_1' = R_1 + C \times (A_1 - X_1)$$
$$R_2' = R_2 + C \times (A_2 - X_2)$$

where C is the volatility coefficient (usually between 10 and 32 in chess)

$$R_1' + R_2' \text{ should be the same as } R_1 + R_2$$

A higher rating always leads to a higher expected score, therefore is more difficult to maintain. We assign initial ratings of 2500, considering the NHL players grandmasters of the game.

Two straightforward applications in hockey

Penalty Shots

We analyzed the data since 1999, although the majority comes since the introduction of the post-game shootout in 2005.

We're able to estimate the probability of a score/save for a particular encounter of a player vs. a goalie.

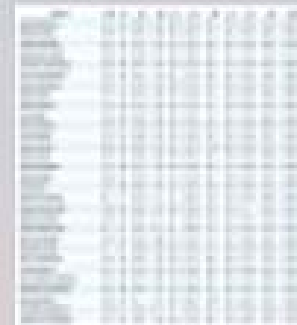Goaltender record and simulation – **Jimmy Howard:**

Skater record and simulation – **Logan Couture:**

Goaltender has about 68% chance of stopping the shot, not 50%. Therefore the starting ratings are not 2500 for both but 1935 for skater and 2565 for goalie.

Face-offs

Face-off data is available since 2002. Using the Elo ratings we can rank the players according to their performance in the circle. We can also project their potential performance against each other.

Season-end record – **2017/18:**

Career record, rating and stick last/first – **Jonathan Toews:**

The face-off is a high frequency event, therefore we change the volatility coefficient from the usual 20 down to 4. We might also need to adjust for the stick first/last on ice.

Other usages and problems

**Team vs. Team**, predict the game outcome or span performance.

Issues:
- Changing team composition: it's not the same team even from game to game
- Extra OT point makes harder to fit results into 0..1 span ($R_1 + R_2$ sum will change!)
- Does goal differential factor in?

Player, predict game or span (individual stat or overall).
Issues:
- What is the expected result?
- The actual result doesn't really fit into 0..1
- Broom: a lot of slight underperformers and a few huge overperformers
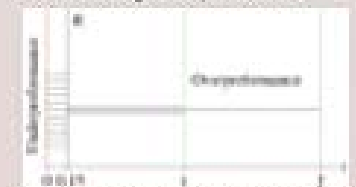
- Goaltenders might be an easier case (svp vs. expected svp)

Teams of players:
Team rating = f(Array Of Player Ratings)

Issues:
- What is f()? Do we use TOI?
- Do we adjust for home/away or for rest?