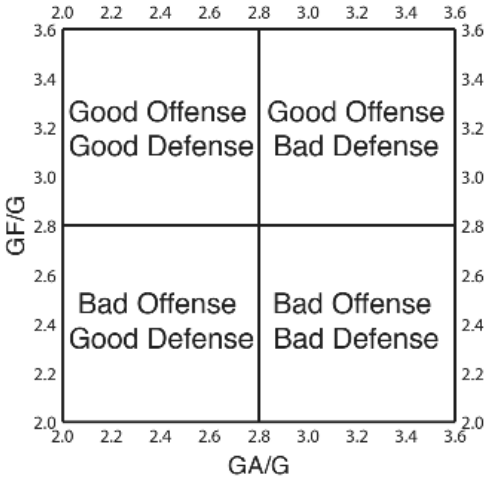
Predicting Matchups in the NHL Based on Goals Per Game

Max Bublitz
April 4th 2019

Hockey is a game that has a ton of lucky bounces, and therefor is a sport that is very susceptible to variation. In a competitive league like the NHL, any team can beat any other team on any given night. Obviously, some teams are better than others, and in the long run a better team will have a better winning record than a worse team. The goal of my project is to determine if goal differential (using goals for per 60 minutes and goals against per 60 minutes) can accurately predict the winner of an NHL game.

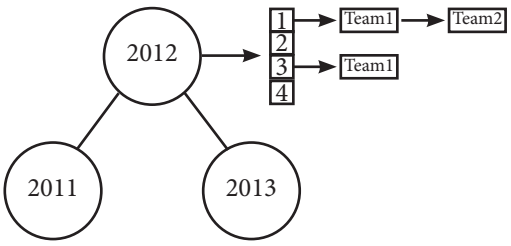
In recent years, the average goals allowed and scored by a team per game is around 2.7. Comparing teams goals for and goals against to 2.7 will show if that team scores or allows more goals than average. For example, if a team scored 2.9 goals a game and allowed 3 goals a game, than they are better offensively than the average team but are also worse defensively.



When comparing a teams goals for and against per game to the average, a team can fall into one of four possible categories:

1. Above average offense and above average defense
2. Above average offense and below average defense
3. Below average offense and above average defense
4. Below average offense and below average defense

You can plot a teams goals for and against per game on a graph to visually understand the categories they fall in. The categories are separated by lines that correspond to the average goals per game. A graph has been made on the left to visualize the four categories and how they are created.



The goal of my project is to assign each team to a category and see if certain categories have more success against others. I will read in files with information from teams from the past few years. I will store the years in a tree, the categories in a hash table and the teams in a linked list. This data structure is depicted on the left of this paragraph. The numbers represent the categories.

Once all the teams from all the years have been read in, I will read from a file that has the resutls of every game from that season and see how each team from each category faired against other categories. If two teams from the same category are playing each other that game will be ignored. After all of the matchups from each year have been analyzed I will look at how each category did against the others and see if there is enough evidence to conclude that certain categories do better against others.