Jones' Reliability Index

For use with FTC Robotics match data

Alex Jones
Team 6226 Bambusa
alex.theboy.jones@gmail.com
www.ftclist.org

Abstract. A metric that assesses individual robot performance across a competition would allow for a better understanding of a robots abilities and a simpler alliance selection process. This metric is difficult to compute due to the difference in alliance partner skill across every match. We propose an equation that factors out alliance partner performance and scoring variance in order to accurately calculate the reliability of a single robot and predict future ability.

1 - Introduction

The main two metrics in FTC Robotics are QP and RP. QP, or qualifying points, is a tally of the amount of wins a team gets across a competition, two points for a win, one point for a tie. RP, or ranking points, is a backup metric to break ties in QP that is calculated by adding all of the points the losing alliance scores in matches that the robot competes in.

These two metrics provide a basic overview of what alliances are able to win the games or score the most points together, but they do not tell us anything about individual robot performance. A circumstance exists where a team is elevated to the top of the QP and RP rankings because of exemplary alliance partner performance alone. This begins to become an issue when teams are selecting alliance partners for the finals matches. Because the alliances are selecting individual teams, the QP and RP calculations are not the most accurate measures of performance, as they can be skewed by alliance partners.

The alternative metric that we propose focusses on individual robot performance over a set of games throughout a competition. This metric is most useful for instances such as alliance selection, where a team must evaluate which

other teams will score points reliably each round and give them the best chance of success in the elimination matches.

2 - Structure

Reliability is a calculation that is done in a variety of different fields, such as electrical engineering, where reliability is the probability that an item will fail at any given time. In this instance, we are attempting to quantify the reliability of a single robot, with only the match data from a competition to use in our calculations.

We first separate reliability into three relevant categories:

1. **Score average**, the typical scores received by an alliance a robot was part of.
2. **Score variance**, the difference in scores received from one match to the next.
3. **Alliance partner weight**, the strength of the alliance partners a team was paired up with.

These categories are chosen because they are all calculable with the data set of matches we are given, and they each show a different aspect of total robot performance. **Score average** is the baseline value of performance of a team with its alliance partner included. Teams with a higher **Score variance** will have inconsistent scores from one match to the next, which shows relative unreliability. Finally in order to differentiate from metrics like QP, we have the **Alliance partner weight** which tells us about the performance of a team's alliance partners compared to all of the teams in the competition. In an abstract form, if we are able to subtract off the **Score variance** and **Alliance partner weight** from the **Score average** of a team, we should be able to get a decent figure of reliability.

3 - Equations

In order to apply what we have come up with in the previous section, we simply need to come up with equations to quantify each of the three measurements stated in the previous section.

The first, **Score Average**, is simple, we just take the average score of all the matches a team has been in.

$$SA = \frac{\sum_{i=0}^{n} S_i}{n},\ where\ n\ is\ the\ number\ of\ scores\ and\ S_0 \to S_n\ are\ a\ teams\ match\ scores$$

Next, we have the **Score variance**. The early calculations of this metric followed a standard deviation equation, but after some testing, we found that this part of the equation was weighing much more heavily than the other parts because the equation increases quadratically with higher variance, whereas the other parts of the metric were performing linearly. We have modified this equation to act in a linear way because of this.

$$SV = \frac{\sum|S-\overline{S}|}{n},\ where\ n\ is\ the\ number\ of\ scores\ and\ S_0 \to S_n\ are\ a\ teams\ match\ scores$$

Lastly, we have the **Alliance partner weight** equation. For this one we just need to average the scores of each match a team's alliance partners were in and subtract the average of all of the match scores. A positive number tells us that the teams alliance partners were above average, and negative below average.

$$APW = \frac{\sum_{i=0}^{m}(\overline{W}-\overline{A_i})}{m},\ where\ m\ is\ the\ number\ of\ alliance\ partners, \overline{W}\ is\ the\ average\ score\ across\ all\ matches,\ and\ \overline{A_0} \to \overline{A_m}\ are\ alliance\ score\ averages.$$

Putting all of the components together, our final equation is

$$JRI = SA - SV - APW$$

With this equation we are first taking the baseline **Score Average**, subtracting out the **Score variance**, then subtracting out the **Alliance partner weight**. This gives us the JRI, or **Jones' Reliability Index**, which should give us an accurate measure of an individual team's performance given competition match data alone.

It should be noted that the values we receive from this equation do not have any value by themselves. Because this is an index, the numbers we receive from this calculation are only useful when comparing to other teams scores within the same pool of data. That said, the occurrences of negative numbers are incidental to the calculation process, and are only significant when comparing to other teams.

4 - Implementation

The sample implementation of this algorithm is put into use in the ftclist.org scores database, where the calculations are done for all documented competitions and regions. Where n is the number of matches, and m is the number of teams, Reliability can be computed for a single team in $O(n * m)$, which for standard competitions with between 20 to 40 matches and around 15 teams can be computed on the client side in a negligible amount of time.

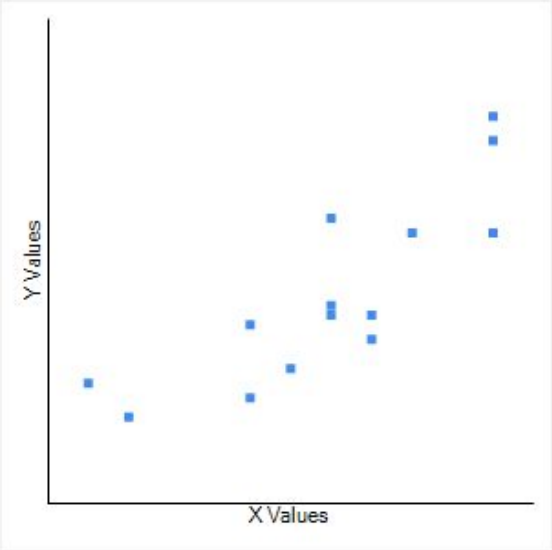
Here is an example implementation of the index in javascript where we take an array of team scores, an array of alliance average scores, and the global average for the competition as parameters and output the reliability of this team.

```
function JRI(teamScores, allianceAverages, globalAverage)
{
	var teamAverage = 0;
	for(var i = 0; i < teamScores.length; i++)
	{
		teamAverage += teamScores[i];
	}
	teamAverage /= teamScores.length;

	var teamVarience = 0;
	for(var i = 0; i < teamScores.length; i++)
	{
		teamVarience += Math.abs(teamScores[i] - teamAverage);
	}
	teamVarience /= teamScores.length;

	var allianceWeight = 0;
	for(var i = 0; i < allianceAverages.length; i++)
	{
		allianceWeight = allianceAverages[i] - globalAverage;
	}
	allianceWeight /= allianceAverages.length;

	return Math.round(teamAverage - teamVarience - allianceWeight);
}
```

5 - Analysis

In our tests with this metric, we have found a very strong correlation between QP and Reliability. Below is a graph with x values as QP and y values as Reliability.

X Values	Y Values
10	60
10	36
10	55
8	36
7	19
7	14
6	21
6	39
6	19
5	8
4	17
4	2
1	-2
0	5



With this set of data from a league meet, there is a **0.8377** correlation coefficient between QP and Reliability, which shows a strong correlation between the two. This is to be expected, because the teams that were able to score points reliably each match are more likely to win the matches they are in.

A notable difference in this dataset is the point (6, 39). The surrounding y values are more than 20 away from this point, showing a clear outlier. Looking at the individual match data for this team, we have {20, 14, 68, 81, 107, 109}, where the global average for the league meet was 46, and the alliance weight was 5, which shows only slight help from their alliance partners. The team only got 3 wins out of a possible 6 because of early failures with their robot which led them to a low QP, but their later game performance showed scores twice as high as the global average. If a team were selecting alliance partners by QP alone, they would consider this team to be 8th, and fail to see the scoring potential of this team.

6 - Conclusion

The Jones' Reliability Index has been shown to hold some notable properties in defining individual robot performance throughout a set of matches. Though it is not a perfect solution and in no way replaces the QP and RP system, it is a notable metric for FTC robotics matches. The sample implementation is at ftclist.org, along with an ever growing list of match data.

References

[0] FTC Game Manual Part 1
(http://www.firstinspires.org/sites/default/files/uploads/resource_library/ftc/game-manual-part-I.pdf)

[1] Standard Deviation
(http://www.mathsisfun.com/data/standard-deviation.html)

[2] Correlation Coefficient Calculator
(http://www.socscistatistics.com/tests/pearson/Default2.aspx)