Introducing Roundnet Player Rating

Jun 30, 2017

Skyler Boles dives for a ball during a game at the San Diego Tour Stop this season.

Note: This piece was written by Max Model.

Several weeks ago, we posted an article on The Rally that broke down some statistics from this year's College National Championship.

The focus of that article was visualizing trends in serving percentage, aces, and defensive breaks. Since then, I have undertaken the more general project of analyzing what roundnet statistics can tell us about a player's value. I am excited to share my first model for Roundnet Player Rating (RPR) and the improvements I am working on.

Quick overview

My goal was to create a statistical metric that attaches a numerical value to a player's performance based on data from the game, essentially scoring a player's performance.

This is similar to what Billy Bean did in baseball. For all of you non-Moneyball fans, Billy Bean, a general manager working with a low team salary cap, used statistics and run-scoring metrics to recruit and hire undervalued and unknown players that performed well in his system. He turned the baseball industry upside down.

I began this project by researching sabermetrics, which is the statistical analysis of baseball records for the purposes noted above. I came upon one of the most famous equations, the Base Runs algorithm, which allows you to input a season's worth of baseball batting data to accurately estimate how many runs a team or player will score over the course of that season. This is the general Base Runs formula:

$$\frac{A * B}{B + C} + D$$

A = H + BB - HR

B = (1.4 * TB - .6 * H - 3 * HR + .1 * BB)

C = AB - H

D = HR

---

where

- **H** - Hits
- **BB** - Walks

- **HR** - Home runs
- **TB** - Total bases
- **AB** - At-bats

For this project, I researched a number of metrics, like Base Runs, to learn about their structures. Like other algorithms, Base Runs has a general equation showing the relationship between four parameters A, B, C, and D, as well as four independently modeled and differently weighted equations for each parameter. Each equation for A, B, C, and D models a different aspect of baseball stats - D accounts for home runs, C for unsuccessful at bats, and so on.

A similar format is found for a football quarterback's passer rating:

$$\text{Passer Rating} = \left(\frac{(a+b+c+d)}{6}\right) \times 100$$

$$a = \left(\frac{\text{COMP}}{\text{ATT}} - .3\right) \times 5$$

$$b = \left(\frac{\text{YDS}}{\text{ATT}} - 3\right) \times .25$$

$$c = \left(\frac{\text{TD}}{\text{ATT}}\right) \times 20$$

$$d = 2.375 - \left(\frac{\text{INT}}{\text{ATT}} \times 25\right)$$

---

where

- **COMP** - Completions
- **ATT** - Attempts
- **YDS** - Passing Yards
- **TD** - Touchdowns
- **INT** - Interceptions

For passer rating, each parameter (A, B, C, and D) accounts for a different aspect of a quarterback’s game. A accounts for completions, B for yards, C for touchdowns, and D for interceptions. Each parameter weights these individual statistics differently, and thus each parameter has a different formula. I followed this basic structure with my model for roundnet statistics.

The process

But before I could do that, I needed data. I arrived at the basic spreadsheet below and tallied each player's statistics while watching video of a variety of games.

GAME INFO: ________________ vs ________________

Game #: __________

Winner: __________

Location: __________

INPUT DATA:

Player	Serves Made	Serves Missed	Aces	Aced	Put Aways	Spikes Returned	Missed Spikes	Missed Sets	Tough Touch	DTouch NR	DTouch R

To quickly clarify the meaning of several of these categories:

- A "Put Away" is any spike a player hits that the defense does not successfully return to the net, no matter why it doesn't get returned.
- "Spikes Returned" are spikes successfully returned to the net by the opposing team.
- A "Tough Touch" is a judgment made on any serve return or set, one on which one would have expected a better touch and which indirectly leads to the loss of the point.
- A "DTouch NR" is a defensive touch that, for whatever reason, does not successfully get returned to the net.
- A "DTouch R" is a defensive touch that is successfully returned to the net, no matter how it happens. If a player gets a DTouch and their partner hits it on the net on the second hit, that player gets credit for the DTouch R, while their partner will receive credit for either the Put Away, Spike Returned, or Missed Spike.

After I collected data, I chose four parameters and weighted each of these parameters as follows:

Hitting - 20%

Defense - 30%

Serving - 30%

Cleanliness - 20%

I attributed only 20 percent apiece to **Hitting** and **Cleanliness** because they are essential parts of the game but are not what generate the break points that are needed to win. **Defense** and **Serving** receive 30 percent weight each because they contribute more weight to winning games.

The next step was to create the formula for each parameter. I'll go through them one by one.

Preston Bies of 2 Guys hits a ball during a game against Origin Impact this season.

Hitting (20%) = 20 - 20 * (Spikes Returned/Total Spikes)

This is a decreasing function, similar to the D parameter for interceptions in passer rating. A player begins with a set value of points (20) and their point total drops fractionally for each of their spikes that is returned. This category includes all spikes, those on normal put aways and those during rallies.

Defense (30%) = DTouchNR + 0.4 * **Hitting** * DTouchR

To calculate a player's score on **Defense**, +1 point is awarded for each unsuccessful defensive touch, such as a hand or body block that goes in the opposite direction of a player's partner, or even a nice touch that for whatever reason a player's team fails to convert. I decided to attach a small value to each DTouchNR because some credit should be given for being in the right spot at the right time and getting a touch, even though it is unsuccessful and amounts to no tangible value in a game of roundnet.

DTouchR is where things get interesting. This portion of the **Defense** category is affected by **Hitting** score, because after all, to successfully convert a DTouch into a DTouchR, a player must successfully spike it on the net. If a player's **Hitting** score is mediocre, say a 10 out of 20, then the defense is more likely to return their spikes after their successful defensive touch. Therefore, even though they may have made a great play on the ball, their play is essentially worthless numerically because the odds of it being returned again are higher.

That being said, with a perfect **Hitting** score of 20, the maximum value of a DTouchR is 8 points. I weighed it this way so that three to four successful defensive touches constitutes, statistically speaking, a great defensive game, and is deserving of almost all the defensive points allotted.

Note that there is no limit to how high your defensive score can go, as there is essentially no limit to how well you can play defensively. This formulation of the **Defense** metric is preliminary; refer to the end of the article to read about what I am trying next.

Peter Jon Showalter of cisek_showalter goes to hit a serve at the San Francisco Tour Stop.

Serving (30%) = 5.5 * Aces + 15 * (Serves Made/Total Serves)

Out of the allocated 30 points for serving on the 100 point scale, 15 are designated to serving percentage and the rest to aces.

You may ask, why does an ace only give a player 5.5 points while a DTouchR is worth a maximum of 8 points? My justification for this is because some of the value of an ace is already awarded in the serving percentage portion of the serving category. Serving percentage only matters on the premise that serves are effective enough to cause mistakes or increase the chance of getting defensive touches - if a player's serve is completely ineffective and they simply plop the ball on the net, their serving percentage becomes less important in the grand scheme of their RPR.

That being said, a player is awarded points for their serving percentage based on the assumption that their serve is at least relatively effective. An ace is considered a bonus to the serving score category. Think of it like this: a player's serve being effective enough to get an ace is considered a slight bonus, as it is already assumed to be decently effective.

Also note, similar to the **Defense** category, this section has no score limit. However, it is scaled such that a recording a solid serving percentage and roughly three aces awards a player all 30 serving points, as that is considered a great serving game.

Cleanliness (20%) = 20 - 5 * (Missed Set + Missed Spike) - 2 * (Tough Touch + Aced)

This category accounts for sets and mistakes. At the top of competitive roundnet, both setting and limiting mistakes are necessary to be good. **Cleanliness** is a decreasing function for that exact reason: a player is 20 points at the outset of a game and a healthy number of points is deducted for each mistake made.

Five points are deducted for each directly missed set or spike, as they result in a lost point. Two points are deducted for tough touches or getting aced - a tough touch indirectly causes the loss of the point and getting aced is often times not the serve returner's fault but goes to the server’s credit.

Peter Jon Showalter dives to set a ball during a match at the Baltimore Tour Stop.

The RPR model has one other feature that arises due to the structure of roundnet games. I realized that if games go into extra points, the **Defense** score gets severely inflated. For example, the average score for **Defense** in games I watched is around 11 points, and in the 55-53 2015 Cream of the Crop game between Nashburgh and Moist, the four player's scores were all around 40.

They didn't all necessarily play stunning defense in that game. Rather the game was over twice as long as usual and that inflated the **Defense** score. This inflation does not happen to the other categories. In fact, **Hitting** and **Cleanliness** decrease slightly as the game gets longer, as you'd expect, and **Serving** goes up only moderately, nullifying the decreasing contribution of **Hitting** and **Cleanliness**. **Serving** tends not to inflate because it is very rare that teams trade aces back and forth during a game that goes into extra points.

To fix this, I added a small condition to the **Defense** score. I found that over the 20 games I watched, the average Total Points was 44, meaning the average game score was 23-21. I added a condition that if the Total Points of a game exceeds 44, then the **Defense** score is scaled by a factor of (44/Total Points) so that the **Defense** scores are scaled to the length of an average game.

To recap, RPR is calculated as follows:

RPR = **Hitting** + **Defense** + **Serving** + **Cleanliness**

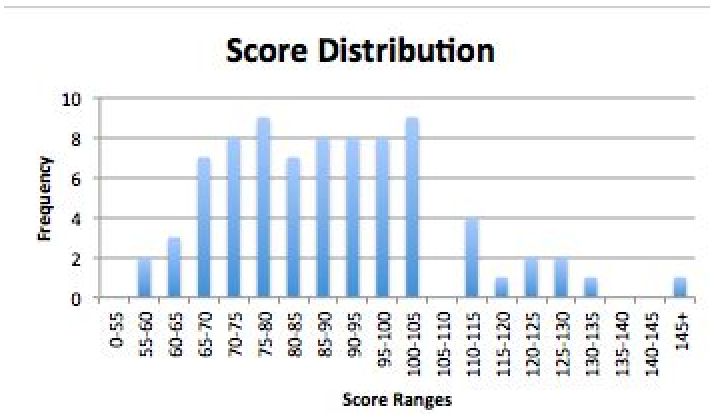
where

FORMULAS:		
Hitting	20%	20 - 20*(Spikes Returned/Total Spikes)
Defense	30%	DtouchNR + .4* **Hitting** *DTouchR
Serving	30%	5.5*Aces + 15*(Serves Made/Total Serves)
Cleanliness	20%	20 - 5*(Missed Set + Missed Spike) - 2*(Tough Touch + Aced)

What we learned

The final step in this process was watching more games, gathering more data, calculating RPR values, and interpreting what they meant. After watching 20 games, I found that the 25th percentile score was 47.7, that average score was 56.6, and that the 75th percentile score was 63.6. The lowest score was 37.1, and the highest score was 95.

Below is a histogram showing the score distribution for the 20 games I watched.



I then scaled the scores so that they were easy to interpret, and I did so in a way that resembled the way a quarterback's passer rating is scaled. In terms of passer rating, an excellent game is a score of 100 or higher, and league average is about 83.2. The maximum possible passer rating is 158.3 (which is rather arbitrary).

I scaled all RPR values by a factor (100/63.6) so that the 75th percentile score became 100. I considered any score above the 75th percentile equivalent to an excellent game, and so in the same way that a 100 passer rating is excellent, so is a score of 100 in RPR. This made the average score equal to 89, and the 25th percentile equal to a score of precisely 75.

In summary and for ease of interpretation,

- 100 is an excellent game
- 90 is an average game
- 75 is a poor game

I’ll walk through one of the games that I coded to show how the data collection and model application works.

This was game one of 2016 Nationals quarterfinals, The Rookies v. Nashburgh. Let’s break down each player’s performance and evaluate how the model reflects that.

GAME INFO:	Rookies		vs	Nashburgh				RESULTS:					
								Player	Hitting	Defense	Serving	Cleanliness	TOTAL:
	Game #:	1						Scott	20	9	16	18	97.5874221
	Winner:	Nashburgh						Joel	17.7777778	17.2222222	16.4090909	15	102.868127
	Location:	Nationals 2016 Quarters						Tyler	16.6666667	6.66666667	10	18	79.5156772
	TOTAL PTS:	40						Ryan	15	2	16	16	75.9013283
INPUT DATA:													
Player	Serves Made	Serves Missed	Aces	Aced	Put Aways	Spikes Returned	Missed Spikes	Missed Sets	Tough Touch	DTouch NR	DTouch R	Total Serves	Total Spikes
Scott	7	3	1	1	6	0	0	0	0	1	1	10	6
Joel	8	3	1	0	7	1	1	0	0	3	2	11	9
Tyler	6	3	0	1	5	1	0	0	0	0	1	9	6
Ryan	7	3	1	1	6	2	0	0	1	2	0	10	8

First is **Hitting**. Not one of Scott Wilson’s spikes was returned, so he gets the full 20 points. Joel Graham, Tyler Cisek, and Ryan Fitzgerald each had at least one spike returned and thus sported a lesser Put Aways/Total Spikes rate, so their **Hitting** scores slip mildly. It was a relatively average game across the board with respect to hitting.

Nashburgh was much more active on the defensive side of this game. During the course of the game, Wilson had a successful DTouchR and Graham had two. Three defensive breaks in a game are crucial. This renders Wilson a score of 9 and Graham a score of 17 out of 30 for **Defense**, so Graham had an above average game in terms of his defense.

Alternatively, Fitzgerald had next to nothing going on defense that game with zero DTouchR and was awarded a mere two points for his two DTouchNR. Cisek had just one defensive break the entire game. This category is what appeared to differentiate Nashburgh and The Rookies in this game.

All four players made between six and eight serves and missed three, which is relatively average or just below average. On top of that, all but Cisek each had an ace, raising their serving scores each to about 16 out of the possible 30 points. They did nothing special serving-wise this game, but they did put serves on the net at a decent rate while each getting an ace, and their **Serving** scores reflect that.

Finally, the **Cleanliness** section. Graham missed a spike at one point in the game, dropping his hitting score by five points. Otherwise, with each player getting aced or having at least one Tough Touch, the **Cleanliness** scores are quite similar across the board.

Scott Wilson of Nashburgh follows through after hitting a ball at 2016 Nationals.

Looking at each player as a whole, Wilson and Graham clearly dominated the game. Wilson's hitting was untouchable, he had a defensive break and an ace, served consistently, and played clean. Graham was not as clean, but had an extra defensive break to make up for it. They both scored around 100 for RPR, which is in the league of the 75th percentile and is considered a strong game.

Alternatively, Cisek and Fitzgerald had little going on defense and from the six-foot stripe, and they did not hit or play as clean as Nashburgh. This combination gave them RPR scores of 79 and 75, respectively, which are both in the 25th-30th percentile, or below average games.

Future directions

After completing this preliminary model, my next steps are to modify the equations. For starters, I plan to add a “Great Sets” category to the **Cleanliness** section, worth +2 points for each "Great Set." I also plan to change the **Defense** section rather drastically, differentiating between first defensive touches and those following the first, as the first defensive touch is by far the most difficult, and those that follow, like those in a rally, are generally easier to attain.

More than just attaching a numerical value to how a player performs in a single game, this metric is a way of tracking a player or team’s collective performance over the course of a season. It is theoretically possible to quantify who the best roundnet player was in the 2015 or 2016 season, or to determine who the best player is in each region this season. We can calculate who plays better on sand or in the month of July, or who shines the most in a specific round of bracket play.

This is a good start to incorporating statistics in roundnet, but there is a long way to go. We'd like to hear from you, too. What are your thoughts on this model?