Michael Freeman

Dr. Chartier

MAT 395

March 20, 2020

March Madness Report: PageRank

During my research I focused on using the PageRank method to predict games in March Madness. At the beginning of this project the goals were to determine if PageRank was able to outperform Massey and Colley in certain years, and if changing the alpha value and weighting the method improved the results for the ESPN score and accuracy. In addition to these additional goals I explored the results of different weightings and different alphas by round. The alpha value is important because in the Page Rank algorithm it determines what percent of time a random event will occur. This percentage is 1 – alpha. I also used another student’s research to determine which years this method would work the best and how that would apply the years to come. All results and averages are based off of all of the data from the 2002-2018 seasons.

The first step of the project was to see which years PageRank outperformed Massey and Colley methods. To do this I found the ESPN Score and percentage correct for each method in every year from 2002-2018. I have shown the values for multiple PageRank alpha values and the best Page Rank value for each year. Figure 1 and 2 show these results using no form of weighting in select years.

A screenshot of a cell phone

Description automatically generatedA picture containing grass, green, wooden, large

Description automatically generatedA close up of a sign

Description automatically generated

Figure 1: Shows the ESPN Score for all the methods

A screenshot of a cell phone

Description automatically generatedA close up of a sign

Description automatically generatedA close up of a sign

Description automatically generated

Figure 2: Shows the percentage of games picked correctly for all the methods

Each year seemed somewhat different in terms of which method performed the best. Consistently it was the Massey method, but there are a few years where PageRank did very well in comparison to the other methods. In Figures 1 and 2 I picked the years where PageRank performed very well: 2009, 2011, and 2014 while also showing a few years where it did poorly and where it was similar to the other methods.

The next step of this project was to determine what value to use for alpha in the PageRank algorithm. At the beginning he thought was the less random the algorithm is the more accurate it would be in predicting the results. This clearly was not the case, as the highest values for alpha consistently produced worse results as seen in Figure 3.

A picture containing clock

Description automatically generatedA screenshot of a cell phone

Description automatically generatedA screenshot of a cell phone

Description automatically generated

Figure 3: Shows how the ESPN Score and Percentage changed with a change in Alpha

Based on these results an alpha value of .55 was optimal on average, but that means 45% of the algorithm was using randomness, which seemed like too much. So, based on these results I decided to pick .75 as the alpha value because it showed fairly high ESPN Scores and percentage of games picked correctly, while also lowering the randomness.

After determining that PageRank is a viable option in many years of March Madness, I continued to try to improve the method by weighting. In Figure 4 I tried many different weights for the entire season. Dividing the season into 3 or 4 segments seemed like the best idea and then I created some weights with even gaps between them and some with decreasing as the season went on, always with later games being weighted more heavily. The results of these weights are shown in Figure 4 for the years 2002-2018.

A picture containing clock

Description automatically generatedA close up of a sign

Description automatically generatedA picture containing clock

Description automatically generatedA picture containing hanging, wooden, many, white

Description automatically generated

Figure 4: Shows the results when weighting parts of the season

The best weights were clearly [.5, 1, 1.5] and [.5, 1, 1.3, 1.4]. So those were chosen to combine with the home, away, and neutral game weighting that is seen in Figure 5. The location weighting was done similarly to the season weighting, in that I tried many combinations and strength and proportion. The averages of the results are below in Figure 5.

A picture containing clock

Description automatically generatedA close up of a sign

Description automatically generatedA screenshot of a cell phone

Description automatically generatedA close up of a sign

Description automatically generated

Figure 5: Shows the results when weighting home, away, neutral games

The best weighting for location(Home, Neutral, Away) was (.5, 1, 1.5) and (.7, 1, 1.3). These two weights were then combined with the season weighting to determine if combining the two improved the model.

When the two weights were combined I got mixed results. The combinations of these two weighting mechanisms caused the ESPN Score to continue to rise, but the overall percentages went down. This means the weighting caused the algorithm to pick later games more accurately, but it struggled to get the earlier ones of less point value. The results of this are seen in Figure 6.

A screenshot of a cell phone

Description automatically generatedA picture containing green, city, track

Description automatically generatedA screenshot of a cell phone

Description automatically generatedA picture containing city, clock

Description automatically generated

Figure 6: Shows the results when combining location and season weighting

As you can see the best weighting for ESPN Score and Win Percentage was the (.7, 1, 1.3) combined with [.5, 1, 1.3, 1.4] weighting. The ESPN Score increased by about 20-30 points from only weighting it with the location and only weighting the season.

The final part of this project was finding how different weighting affected the performance of the method from a round by round standpoint. To do this I edited the MATLAB code to output a CSV file which contained the percentage accuracy of the method in each round. From this I was able to look at how the different weights and combinations of weights affected the round by round score. Also, I hypothesized maybe the alphas would change how the algorithm would perform round by round since randomness might be more present in earlier rounds of the tournament where there are more opportunities for upsets. Figure 7 shows the round by round performance of the different weighting measures and Figure 8 shows the performance based on the different alphas. Unfortunately, our results showed that one weighting performed the best across all rounds, which was (.7,1,1.3) and [.5, 1, 1.3, 1.4]. The results using the alphas was similar in that if you picked .85 as the alpha it would perform the best across all rounds except the 2nd round, where .95 slightly outperformed it. I also included the average on select years that PageRank performed well which were: 2004, 2009, 2011, 2014, 2015, and 2018. The results using those years were not significantly different for either weighting or the alphas.

A picture containing clock

Description automatically generatedA close up of a scoreboard

Description automatically generated A screenshot of a cell phone

Description automatically generatedA screenshot of text

Description automatically generated

Figure 7: Round by round weighting Figure 8: Round by round alphas

Looking at the select years and using another student’s research we found a high positive correlation between the instability of the tournament field and how well PageRank performed. Those select years, especially 2014 and 2011 had very high instability, which basically means there are not teams that are clearly separated from other teams. This causes lots of upsets which PageRank has been shown to capture better than any other method we’ve studied.