12-2013

College Football Rankings: An Industrial Engineering Approach

Brent H. Wiles
University of Arkansas, Fayetteville

College Football Rankings:
An Industrial Engineering Approach

An Undergraduate Honors College Thesis

in the

Department of Industrial Engineering
College of Engineering
University of Arkansas
Fayetteville, AR

By

Brent H. Wiles

This thesis is approved.

Thesis Advisor:

Dr. C. Richard Cassady

Thesis Committee:

Dr. Chase Rainwater

Dr. Kelly Sullivan

University of Arkansas – College of Engineering
Honors College Graduation Certification

Name: Wiles, Brent H.
Last Name First Name Middle Initial

Student ID #: 010522585 Department: INEGBS

College of Engineering requires 12 hours of honors credit and honors thesis (6 of the 12 hours of must be engineering):

Engineering (6 hours)

Course (Dept. & Number)	Course Title	Semester	Grade
INEG 400VH	Honors Thesis	Fall 2012	A
INEG 4433H	Honors Systems Engineering and Management	Fall 2012	A

Any Honors (6 hours)

Course (Dept. & Number)	Course Title	Semester	Grade
MATH 2574H	Honors Calculus III	Fall 2010	B
PHYS 2054H	Honors University Physics I	Fall 2009	A
GNEG 1111H	Honors Intro to Engineering I	Fall 2009	A
GNEG 1121H	Honors Intro to Engineering II	Spring 2010	A

Thesis Title: College Football Rankings: An Industrial Engineering Approach

College Football Rankings: An Industrial Engineering Approach

Signatures below indicate that the thesis has been completed by the student and accepted by the Thesis Advisor and Department.

[signature] 12/11/13

Thesis Advisor (Print and Sign) Date

[signature] 12/11/13

Departmental Coordinator (Print and Sign) Date

This form must be submitted to the Engineering Dean's Office (BELL 3189) at least one week before dead day of the student's graduating semester.

Table of Contents

Introduction

In modern day America, there are few people that would argue with the idea that football is the most popular sport being played today. Even though baseball is considered “America’s Pastime”, today it is football that reigns supreme, as evidenced by the National Football Leagues annual revenue of over $9.5 billion. That is nearly a 36% increase over the revenue of Major League Baseball (Bery, 2013). However, while the professional leagues are generating these dynamic revenues, many would argue that the best form of football at present is taking place a day before the pros hit the turf, on the college gridiron. While the game of college football is generating much of it’s own buzz, many times it is the ranking of the teams that is the subject of much controversy and discussion, with the Bowl Championship Series’ (BCS) rankings being at the very heart of the conversation.

Officially beginning in 1998, the BCS is a series of 5 championship games that highlights 10 of the top college football teams in the country (BCS, 2013). The biggest of these 5 games is the BCS National Championship Game which determines who will be crowned as the best team in the country at the end of each season. The teams that play in these games are determined by the final BCS rankings for that football season. Now some may argue that with 27 non-BCS bowls available, why does it matter so much about the rankings? With 64 teams qualifying for bowl games at the end of each season, surely every deserving team will get the opportunity to play in a bowl? So what is it that makes these 5 games in particular such an object of controversy? The most likely answer is one that determines so much in today's society: money. While the notoriety and publicity that comes with qualification is important to the schools, the biggest difference comes with the payout. In 2012, each team in each of the BCS bowls received a payout of $17 million, with an extra million per team being paid out for the championship game. This is compared to the next largest bowl game payout of just $4.5 million (Statistic Brain, 2012). With an increase of nearly 400% between those bowls, it is easy to see why it is so crucial that the teams are properly ranked and placed in these games; this is the goal of the ranking system considered in this work.

CMS+ System History

The CMS+ Football Rankings System is an optimization-based methodology that is used in ranking college football teams. Since its initial publication in 2005 by collaborators Cassady, Maillart and Salman, the CMS+ system has been evolving steadily into what it is today (Cassady, 2005). At its core, the CMS+ system has always focused on 4 main data figures: what 2 teams played, what was the final score, was the winning team playing at home, on the road or at a neutral site and when in the season was the game played. As the system has developed, further factors have been considered including: whether the game went to overtime, how many overtime periods were played, what was the teams final rank in the Associated Press (AP) Poll and whether the winning team was their conferences' champion. Using these factors, the CMS+ system produces a matrix that assigns each pair of teams a value, originally referred to as degree of victory. As more ranking factors have been incorporated to compare teams that didn't play in a head-to-head game, this value has come to be called evidence of superiority.

The Ranking Problem

Currently, there are 125 teams in Division I football playing at the Football Bowl Subdivision (FBS) level. This means that there are 7,750 possible matchup combinations between teams. However, in a standard season, there will only be around 780 games played covering approximately 10% of the possible match-up combinations. This is what makes the ranking problem so difficult; teams are being compared that were never given the opportunity to play in a game.

The problem of ranking college football teams can be defined as a quadratic assignment problem. In this instance of the problem, each team is assigned one and only one ranking, and each ranking is assigned one and only one team. In addition, there is a distance value assigned to each pair of rankings based off of the standard normal cumulative distribution. The objective function then is to maximize the sum of these relative distances multiplied by the evidence of superiority values of the pair of teams.

Genetic Algorithm Optimizer

Once each pair of teams has been assigned an evidence of superiority value, which will be discussed in more detail later, the data is fed into an optimizer, which utilizes a genetic algorithm accompanied by a local search to produce the resulting ranking. The optimizer follows a three-step process in order to produce the best ranking as governed by the objective function, whose value will be referred to as fitness. The first step is the genetic algorithm, which includes mathematical representations of breeding and mutation.

To start off, the first generation of 100,000 total generations is produced at random. 100 feasible rankings are produced and sorted according to their fitness values. The ranking with the best fitness is noted and used for future comparison. For all subsequent generations, the top 75 rankings from the previous generation are kept, then 23 new rankings are produced from breeding and 2 new rankings are produced from mutation. The 100 rankings are then ordered again based off of fitness and the top ranking is again noted.

The breeding process itself goes through the following steps. First, 2 rankings are chosen at random to serve as parents for the new ranking. The 2 rankings are compared and any matches are carried down to the child ranking. Second, for each open position in the child ranking, one of the parents is chosen at random and, assuming the team in that position has not already been assigned, the team is assigned to that ranking in the child. The third and final step is to again, select a parent at random and order all unassigned teams based on that parent ranking. The unassigned teams are then assigned to all the gaps in the child ranking based on that order.

The mutation process is much shorter than the breeding process. One ranking is chosen at random to serve as the parent to the child ranking. From that parent, 2 positions in the ranking are also chosen at random. Then, all of the teams within those 2 positions are inverted and the result is the new child ranking. The reason that mutation is used at all in the algorithm is simply to ensure that the same solutions are not produced over and over. However, that is also why there are many more bred solutions than mutated solutions, because the solutions resulting from breeding are generally better solutions than those produced from mutation.

The second step in the optimization process is to use pair-wise switching to improve upon the best ranking that was produced by the genetic algorithm. Starting with that best ranking, the ranking positions are switched in the following order: 1 and 2, 1 and 3 all the way down to 1 and 125. Then continuing down the line: 2 and 3, 2 and 4 down to 2 and 125. This continues until it reaches the switch between 124 and 125. For each switch that is checked, if the overall fitness of the ranking is improved, then the switch is made and the switching process starts over. This will continue until the switching process goes through all 7,750 switches and fails to improve the fitness.

The third and final step has one focus and that is to address randomness. If you think back to the very first step of the optimization process, you will remember that the first generation was random. In order to combat this, the entire process is repeated 20 times. At the end of the 20th iteration, the ranking with best fitness value is output as the best solution.

Evidence of Superiority

Evolution Over Time

Throughout the history of the CMS+ system, the genetic algorithm approach to optimizing the rankings has stayed the same. The majority of the changes to the system have been in the way that the evidence of superiority between teams has been computed.

The original evidence of superiority value, which will be referred to as the f-value from now on, was based solely on head-to-head game results. The calculation would begin with a set value based on where the team won the game. Games won at home would begin with a value of 0.35, games won on the road would begin with a value of 0.65 and neutral site wins would begin with a value of 0.5. This equates to road wins being worth 130% of neutral site wins and home wins being worth 70% of neutral site wins. Once the initial value is assigned, it is simply multiplied by a date multiplier (DM) determined by the day in the season the game was played. The convention used for determining the multiplier was that games won on day 1 would be worth 60% of games won on the last day of the season as shown in equation 1 below:

$$\mathrm{DM} = 0.6^{(\text{last day} - \text{game day})/(\text{last day} - 1)}$$

[Equation 1]

The next iteration of the CMS+ system was a collaboration between Cassady and Sullivan who modified the f-values in a way to begin addressing the large number of team connections that weren't represented by head-to-head play. (Sullivan, 2009) After all, if team A beats team B and team B beats team C, doesn't that provide some evidence that A is better than C? While it is assumed that that would be the case, that assumption is not enough to give the teams full head-to-head credit over teams they never played. Instead partial credit is given to the teams for these indirect victories. The new f-value would start to give team's half the head-to-head value over team's they presumably would have beat. The modified equation for computing f-values is shown in equation 2 where g-values represent the f-values that were achieved in head-to-head victories:

$$f_{i,i'} = g_{i,i'} + 0.5\left(\sum_{i''} g_{i'',i'}\right)$$

[Equation 2]

While this iteration helped to connect many more pairs of teams, there were still numerous pairs that were left unconnected. It was discovered that during any particular football season the maximum number of links between any 2 teams was 4. Knowing that, the transitivity equation can simply be extended to include more degrees of separation. Also, because each link meant that the teams were further apart in actuality, less credit needed to be rewarded with each successive link. This resulted in an exponential reduction of the credit awarded for each additional degree of separation as well as considering different values for θ, which had previously been locked into 0.5. The resulting computation for f-values is shown below in equation 3. There is an additional binary variable φ included which is set to 1 if the maximum degrees of separation for the pair of teams is greater than the current value for degrees of separation.

$$f_{i,i'} = g_{i,i'} + \varphi_2 \theta \left(\sum_{i''} g_{i,i''} g_{i'',i'} \right) + \varphi_3 \theta^2 \left(\sum_{i''} \sum_{i'''} g_{i,i''} g_{i'',i'''} g_{i''',i'} \right) + \varphi_4 \theta^3 \left(\sum_{i''} \sum_{i'''} \sum_{i''''} g_{i,i''} g_{i'',i'''} g_{i''',i''''} g_{i'''',i'} \right)$$

[Equation 3]

Current Iteration

With each change that was incorporated into the evidence of superiority metric the resulting rankings have always been improved. However, there have always been a select number of issues that could be improved upon. The first was that as more degrees of separation were incorporated into the ranking, the results became more and more dependent on a proper value for θ. To combat this issue, the metric has been modified so that the product of the g-values alone sets the value of the credit assigned for indirect victory. Since all g-values will be a value between 0 and 1, each additional g-value incorporated into the product will reduce the resulting f-value. In addition, more emphatic victories, which result in higher g-values, will increase the value of the product and thus increase the f-value.

The second change that was added was to address the small number of games that were played against non-FBS opponents. In past versions of CMS+, any games played against non-FBS opponents were not considered in the rankings. While this continues to be the case for any victories over non-FBS schools, a change was made to account for the occasional loss to a non-FBS school. The reason behind this is that it is assumed that FBS schools will win against a non-FBS school and should therefore not get any credit for padding their schedule with easy wins. The problem came with the losses, as top schools were able to remain high in the CMS+ rankings despite having a season ending upset to a small school. The case was such in the 2010 season when James Madison defeated Virginia Tech. This loss loomed over the team despite winning out during the remainder of the regular season and winning the ACC championship. However, because the CMS+ system didn't account for the loss to James Madison, their rank was skewed much higher in that system than that of the BCS. To combat this issue, the idea of a "dummy" team was incorporated to represent any non-FBS school. This team would only be present in the data set for victories over FBS schools. Then, because they would show up as an undefeated team with the potential to have many more wins than any of the other teams, the dummy team was locked into the bottom of the ranking. This would ensure that the team would not begin to rise up the ranking and that any FBS school that lost to them would be penalized to the maximum degree possible.

The third improvement that was incorporated involved moving away from a sole dependence on the results of games. This involved increasing the resulting f-values of teams by giving teams credit for being the champions of their respective conferences and by giving teams credit over teams that they were ranked above in the AP Poll. The main motivation for incorporating conference champions into the ranking was to resolve the issue from past iterations where the losing teams in conference championship games were being unfairly punished in comparison to the other teams in their conference. While getting to the conference championship game and losing generally means that the team is the second best in its conference, the ranking system would only see the game as an additional loss in an extra game that the majority of teams didn't even have to play. To further complicate that issue, since games became worth more as the season progressed and championship games are the last games of the season, the teams that lost these games would be hit with the worst losses of the season; These coming in games that they actually qualified for based on their success over the course of the regular season.

By incorporating the AP Poll, the ranking would now be able to adjust to factors throughout the season that are difficult to quantify. Since the system focuses on data factors such as winner and losers, it doesn't take into account what experts consider "quality wins". Did the team dominate in all their games, or were there moments where they escaped by the skin of their teeth or won out of sheer luck? However, at the same time, the AP Poll can incorporate a certain amount of bias where powerhouse programs receive the benefit of the doubt and lesser known programs have to overly prove themselves to receive any sort of recognition. Because of this potential bias and the systems overarching goal of providing a ranking that is dependent on data alone, the weight assigned to the AP Poll is minimal.

Variable Weighting Factors

As additional factors have been incorporated into the rankings, the debate has always been raised about how much weight should be assigned to each factor. For example, is being a conference champion a better indicator of a team's superiority than it's relative rank in the AP Poll? Then if that is the case, how much of a better indicator is it? The answer is simply that there is no absolute right answer. The answer will almost always be the opinion of the user conducting the experiment and thus has lead to the development of variable weighting factors defined as α_1, α_2 and α_3. α_1 represents the percentage of the total points that are assigned to head-to-head victories. α_2 represents the percentage of the remaining points that are assigned to indirect victories. α_3 represents the percentage of the remaining points that are assigned to conference champions and all remaining points are assigned to the teams relative ranking in the AP Poll.

In addition to the alpha values, beta values were also incorporated to address previously static variables that were applied to the date when the game was played and whether the game was played at home, on the road or at a neutral site. For the same reason that the alpha values were incorporated, the beta values were added to address what is likely to be an opinion of the user. Some will argue that the better teams will always win no matter where the game is played. Others are firm believers in the impact that the crowd has on the game and that anything can happen when their team is at home. Then there are the arguments made for the second beta value concerning when the game was played. Some take the stance that the better teams start out and remain better throughout the entire season. This meaning that games early in the year should be worth the same amount as those played on the last day of the season. While still others will wait in the other corner and make the arguments that the truly great teams, particularly those with impactful coaches, will improve as the year goes on and will truly be their best at the end of the year. Regardless of which arguments hold true, the system now allows for these variables to be modified and see how they have an effect on the rankings. β_1 represents the percentage of the last day of the season that the first day will be worth; This was previously fixed at 60%. β_2 then represents the percentage of away games that home games are worth, with neutral site victories falling halfway in between β_2 and 1. This is altered slightly from our previous 70/100/130 percent split between home, neutral and away games.

The Resulting Equation

In order to combine the variable weightings formulation with the current iteration of the evidence of superiority, new variables needed to be established and some old ones modified in order to account for the new input factors. In addition to the new variables, the formulation was simplified so that setting input weights would be easier going forward; This is discussed more in the future work section. This is accounted for by making the maximum f-value possible a fixed 100 points.

As before, head-to-head victories would be represented by the variable g, with all possible g-values falling between 0 and 100*α_1. The resulting equations for g-values where team i beats team i' on day N in a season with M days for home, neutral and away victories are shown below in equations 4, 5 and 6:

Victories at Home:

$$g_{i,i'} = (100\alpha_1) \times \beta_1 \times \left(\beta_2 + \left((1-\beta_2) \times \left[\frac{N-1}{M-1}\right]\right)\right)$$

[Equation 4]

Victories at Neutral Site:

$$g_{i,i'} = (100\alpha_1) \times \left(1 - \left(0.5 - \frac{\beta_1}{2}\right)\right) \times \left(\beta_2 + \left((1-\beta_2) \times \left[\frac{N-1}{M-1}\right]\right)\right)$$

[Equation 5]

Victories Away from Home:

$$g_{i,i'} = (100\alpha_1) \times \left(\beta_2 + \left((1-\beta_2) \times \left[\frac{N-1}{M-1}\right]\right)\right)$$

[Equation 6]

Example – Team i beats i' at a neutral site on day 11 of a 101 day season.

$$\alpha_1 = 0.5;\ \beta_1 = 0.5;\ \beta_2 = 0.5;$$

(**Step 1**)

$$g_{i,i'} = \left(100(0.5)\right) \times \left(1 - \left(0.5 - \frac{(0.5)}{2}\right)\right) \times \left((0.5) + \left(\left(1-(0.5)\right) \times \left[\frac{11-1}{100-1}\right]\right)\right)$$

(**Step 2**)

$$g_{i,i'} = (50) \times (0.75) \times (0.55)$$

(**Step 3**)

$$g_{i,i'} = 20.625$$

Next, transitive victories would now be represented by variable $\bar{h}$ and the calculation method would remain the same. The only difference in the calculation of these $\bar{h}$-values is that they are the result of the product of new $\bar{g}$-values as opposed to the new g-values. The variable $\bar{g}$ is defined as:

$$\overline{g_{\imath,\imath'}} = \frac{g_{i,i'}}{100\alpha_1}$$

[Equation 7]

The variable $\bar{g}$was introduced as a result of the 100-point system that was implemented and the removal of the previous variable θ. Since the reduction percentage for transitive victories is now a result of the product of the old g-values between connected teams and the new g-values are greater than 1, the $\bar{g}$variable was incorporated in order to keep the transitive victory factor less than 1. Equation 8 shows the new method for computing $\bar{h}$.

$$\overline{h_{\imath,\imath'}} = \left(\sum_{i''} \overline{g_{\imath,\imath''}}\,\overline{g_{\imath'',\imath'}}\right) + \left(\sum_{i''}\sum_{i'''} \overline{g_{\imath,\imath''}\,g_{\imath'',\imath'''}\,g_{\imath''',\imath'}}\right) + \left(\sum_{i''}\sum_{i'''}\sum_{i''''} \overline{g_{\imath,\imath''}\,g_{\imath'',\imath'''}\,g_{\imath''',\imath''''}\,g_{\imath'''',\imath'}}\right)$$

[Equation 8]

While the $\bar{h}$-value is computed the same way as the previous transitive victory portion of the old g-values, because of the implementation of the new 100-point system the $\bar{h}$-value needed to be modified before it was added into the final f-value. This was done through the addition of two additional variables, h_{max} and h. The value for h_{max} requires no additional computation and is simply defined as the maximum of all the $\bar{h}$-values from that season. Then using the value for h_{max} and the defined variable α_2, all of the $\bar{h}$-values can be transformed into generic h-values as shown below in equation 9:

$$h_{i,i'} = 100\alpha_2(1-\alpha_1)\left(\frac{\overline{h_{i,i'}}}{\overline{\overline{h_{max}}}}\right)$$

[Equation 9]

It is worth revisiting our initial declaration of α_2 where it was defined as the percentage of the points remaining after the points for α_1 have been distributed. This is why the term $(1-\alpha_1)$ is included in the calculation for the h-value. Similar terms are incorporated later in our calculations for the points assigned to conference champions and those teams ranking higher in the AP Poll.

Next, the variable j was designated to represent the points that were awarded to a team over the other teams in their conference for being the conference champion. Since the value for j is based strictly on a binary argument, the computation for j is much more straightforward and shown in equation 10:

$$j_{i,i'} = \begin{cases} 100\alpha_3(1-\alpha_1)(1-\alpha_2) \\ 0 \end{cases}$$

[Equation 10]

The final variable to be added is to account for team's that are ranked higher in the AP Poll. This variable was defined as k and is based on a binary argument similar to that of the variable j. The computation for the value of k is shown in equation 11:

$$k_{i,i'} = \begin{cases} 100(1-\alpha_1)(1-\alpha_2)(1-\ \alpha_3) \\ 0 \end{cases}$$

[Equation 11]

To cap off the equation, the values of the four variables only need to be summed in order to determine the resulting value of f as shown in equation 12:

$$f_{i,i'} = g_{i,i'} + h_{i,i'} + j_{i,i'} + k_{i,i'}$$

[Equation 12]

Experiments

In order to test the new ranking formulation, a series of experiments needed to be devised. Since the new system allowed for the changing of input variables with very little restriction, experiments needed to cover a wide range of input values to measure the various effects of the different input parameters. The parameters were narrowed down in such a way that resulted in 75 different experiment combinations for one season. Figure 1 outlines the different experiment parameter combinations that were performed.

Experiment Combinations:

$$\alpha_1 = \{0.5, 0.75, 1\}$$

$$\alpha_2 = \{0, 0.5, 1\}$$

$$\alpha_3 = \{0, 0.5, 1\}$$

$$\beta_1 = \{0.5, 0.75, 1\}$$

$$\beta_2 = \{0.5, 0.75, 1\}$$

[Figure 1]

*Note: If β_1 or β_2 were equal to 1, the other beta value was set equal to 1.

Results

The previously described experiment combinations were performed on the 2012, 2011 and 2003 seasons for analysis. Each of these seasons contained circumstances that made them interesting to consider. 2012 had two undefeated teams, Notre Dame and

Ohio State, but Ohio State was banned from postseason play for that year. In addition, the regular season produced four one-loss teams competing for that second spot, Kansas State, Oregon, Florida and the eventual champion Alabama. 2011 was the year of the all-SEC championship game where LSU and Alabama squared off for the championship game despite having already played each other during the regular season. During that season there were also two other teams, Oklahoma State and Stanford, which only suffered one loss during the regular season, the same as Alabama. 2003 had a similar situation where both Oklahoma and USC had 12-1 records and were competing to face off against undefeated LSU in the championship game.

In analyzing the results, two approaches were taken. The first is to analyze the fitness value that is output by the rankings. Fitness is defined as the sum of the product of all the f-values of teams and their relative distance from each other in the ranking. This is shown in Equation 13 where D represents the relative distance between the two teams in the ranking.

$$\text{Fitness} = \sum_{i} \sum_{i'} f_{i,i'} D_{i,i'}$$

[Equation 13]

The next step is to take this fitness value and compare it to the average fitness of 1000 random rankings. Using this average fitness and the standard deviation of the random rankings, calculations are performed to find how many standard deviations above the mean the chosen ranking is. Then the z-value is calculated using this number of standard deviations to find the probability of producing a ranking with a higher fitness than the ranking that is being analyzed.

One limitation to this form of analysis is that it can only provide good values for α_1, α_2 and α_3 because of the effect of the beta values on the fitness. Since higher beta values produce higher f-values throughout the course of the season, their sum will naturally result in a higher fitness value. This results in the rankings being skewed towards those where $\beta_1 = 1$ and $\beta_2 = 1$.

The next approach used to analyze results is to take the rankings and compare them to the BCS rankings and attempt to minimize the number of anomalies, particularly higher up in the ranking. An anomaly is defined in this context as an instance where the ratio between the CMS+ rank and the BCS rank of a team is less than ½ or greater than 2.

Anomaly

$$\frac{\text{CMS+Rank}}{\text{BCS Rank}} < \frac{1}{2} \text{ or } \frac{\text{CMS+Rank}}{\text{BCS Rank}} > 2$$

By defining an anomaly in this way, teams that are considerably out of place in the ranking as compared to the BCS ranking are identified. Now, this is not to say that rankings with a large number of anomalies are considered bad. This would require an acknowledgement that the current BCS system is perfect, which is not the case. If it were perfect, there would be no point in producing additional ranking systems at all. Instead, the purpose of the anomaly style of analysis is to get rankings that are comparable enough to the BCS to be examined as potentially good rankings.

2003 Top Rankings with Fitness Approach

#	α_1	α_2	α_3	β_1	β_2	σ's above the mean
1	0.75	0.5	0.5	1	1	11.41
2	0.75	1	0	1	1	10.90
3	0.75	0.5	1	1	1	10.71
4	0.75	0	1	1	1	10.60
5	0.5	1	0	1	1	10.55

BCS	Ranking #1	Ranking #2	Ranking #3	Ranking #4	Ranking #5
Oklahoma	Oklahoma	Oklahoma	Oklahoma	Miami (OH)	Oklahoma
LSU	USC	USC	USC	Boise State	USC
USC	LSU	LSU	Miami (OH)	USC	Florida State
Michigan	Michigan	Florida State	LSU	LSU	Miami (FL)
Ohio State	Miami (OH)	Miami (FL)	Boise State	Oklahoma	LSU
Texas	Florida State	Ohio State	Florida State	Michigan	Ohio State
Florida State	Miami (FL)	Miami (OH)	Michigan	Florida State	Michigan
Tennessee	Georgia	Boise State	Miami (FL)	Miami (FL)	Georgia
Miami (FL)	Ohio State	Michigan	TCU	TCU	Texas
Kansas State	Boise State	Georgia	Utah	Southern Miss.	Kansas State

2003 Top Rankings Minimizing Anomalies

BCS	(0.5,0.5,0.5,1,1)	(0.75,0.5,0,0.75,0.75)	(0.75,0,0.5,0.5,0.5)
Oklahoma	Oklahoma	Oklahoma	LSU
LSU	USC	USC	Oklahoma
USC	LSU	LSU	USC
Michigan	Michigan	Michigan	Michigan
Ohio State	Florida State	Texas	Texas
Texas	Miami (FL)	Georgia [12]	Georgia [12]

Florida State	Georgia [12]	Tennessee	Tennessee
Tennessee	Ohio State	Florida State	Miami (OH) [11]
Miami (FL)	Texas	Miami (OH) [11]	Florida State
Kansas State	Miami (OH) [11]	Miami (FL)	Ohio State

2011 Top Rankings with Fitness Approach

α_1	α_2	α_3	β_1	β_2	σ's above the mean
0.75	1	0	1	1	11.15
0.75	0.5	1	1	1	11.14
0.5	0.5	1	1	1	10.85
0.75	0	1	1	1	10.64
1	0	0	1	1	10.58

BCS	Ranking #1	Ranking #2	Ranking #3	Ranking #4	Ranking #5
LSU	LSU	LSU	LSU	LSU	LSU
Alabama	Oklahoma St.	Oklahoma St.	Oklahoma St.	Oklahoma St.	Houston
Oklahoma St.	Stanford	Houston	Oregon	Wisconsin	Oklahoma St.
Stanford	Boise State	Stanford	Virginia Tech	Virginia Tech	Boise State
Oregon	Houston	Virginia Tech	Wisconsin	Oregon	Stanford
Arkansas	Virginia Tech	Boise State	Stanford	Southern Miss.	Virginia Tech
Boise State	Alabama	Oregon	Houston	Houston	Alabama
Kansas State	Oregon	Alabama	Southern Miss.	Stanford	Oregon
South Carolina	USC*	Wisconsin	Boise State	Boise State	USC*
Wisconsin	Wisconsin	USC*	TCU	Arkansas State	Michigan

2011 Top Rankings Minimizing Anomalies

BCS	(0.5,0,0,0.5,0.5)	(0.75,0.5,0,0.5,0.5)	(0.75,0,0.5,0.5,0.5)
LSU	LSU	LSU	LSU
Alabama	Alabama	Oklahoma State	Oklahoma State
Oklahoma St.	Oklahoma State	Stanford	Alabama
Stanford	Stanford	Alabama	Stanford
Oregon	USC*	USC*	Boise State

Arkansas	Oregon	Boise State	USC*
Boise State	Arkansas	Oregon	Oregon
Kansas State	Boise State	Wisconsin	Arkansas
South Carolina	Wisconsin	Arkansas	Wisconsin
Wisconsin	South Carolina	Kansas State	Michigan State [17]

*USC suspended from postseason bowl play

2012 Top Rankings with Fitness Approach

α_1	α_2	α_3	β_1	β_2	σ's above the mean
0.75	0	1	1	1	13.76
0.75	0.5	0	1	1	12.18
0.5	0.5	1	1	1	11.263
1	0	0	1	1	11.262
0.5	1	0	1	1	10.99

BCS	Ranking #1	Ranking #2	Ranking #3	Ranking #4	Ranking #5
Notre Dame	Alabama	Notre Dame	Alabama	Notre Dame	Notre Dame
Alabama	Northern Ill.	Ohio State*	Stanford	Ohio State*	Ohio State*
Florida	Notre Dame	Alabama	Notre Dame	Oregon	Florida
Oregon	Ohio State*	Florida	Ohio State*	Florida	Stanford
Kansas State	Stanford	Oregon	Northern Ill.	Northern Ill.	Oregon
Stanford	Florida State	Georgia	Florida	Alabama	Alabama
Georgia	Kansas State	Stanford	Kansas State	Kent State	Georgia
LSU	Utah State	Kansas State	Florida State	Georgia	Kansas State
Texas A&M	Florida	LSU	Oregon	Stanford	Nebraska
South Carolina	Tulsa	Northern Ill.	Georgia	Kansas State	LSU

2012 Top Rankings Minimizing Anomalies

BCS	(0.5,0,0,1,1)	(0.5,0.5,0,0.5,0.5)	(0.75,0,0,0.5,0.5)
Notre Dame	Notre Dame	Notre Dame	Notre Dame
Alabama	Ohio State*	Ohio State*	Ohio State*
Florida	Alabama	Alabama	Alabama
Oregon	Florida	Florida	Florida

Kansas State	Oregon	Oregon	Oregon
Stanford	Georgia	Georgia	Georgia
Georgia	Kansas State	Stanford	Kansas State
LSU	Stanford	Kansas State	Stanford
Texas A&M	LSU	Texas A&M	LSU
South Carolina	Texas A&M	LSU	Texas A&M

*Ohio State suspended from postseason bowl play

Rankings Analysis

In analyzing the rankings, it quickly becomes clear that there is no "right" answer to what the best combination of weights is. There are several combinations of factor weights that can give you credible rankings, particularly high up in the rankings. What this does show is that regardless of how the weights are established, the better teams tend to come out near the top of the ranking. This is due to the intertwined nature of the weights available to choose from. If you win, not only do you get more head-to-head points, but also you get more and more transitive victory points because you beat more opponents. Then when you win, you move up in the AP Poll giving you even more points. Then if you win even more, you can claim your conference crown, which provides an additional bonus to your ranking resume.

Future Work

With the new flexibility provided by the variable weighting process, future work can be done to lead to the overarching goal of letting various users produce their own sets of rankings very quickly. Users will be able to decide what they deem the most important factor in evaluating teams at the end of the season and then see what effect that has on the rankings. The first step towards this goal is to improve the efficiency of the ranking application and initial efforts are being focused on the number of replications. Recall that the current application generates 100,000 rankings and then repeats the process 20 times. Analysis is being done to see if there is a point during the application at which ranking fitness is no longer improving and the best ranking has already been achieved. If comparable results can be produced with only 20,000 generations and 5 repetitions, the process time can be reduced exponentially. This will greatly improve the ability of the user to analyze different weighting combinations.

Since the idea of variable weights and multiple factors is still in it's infancy, new potential factors are always being considered. As new factors are presented, they can be tested to see how much of an effect they would really play on the final rankings. This could range from organized metrics like strength of schedule to less traditional metrics such as time of kickoff. As overall system efficiency is improved, the possibilities are quickly expanded and the opportunities for future work become more immense.

Conclusion

By incorporating additional ranking factors and allowing their weights to be variable, the CMS+ system has received many new benefits. Not only do these changes provide improvement and increased balance to the rankings, but they also open multiple doors for the direction the system wants to go. In the future, any user will be able to produce a viable ranking using only the factors that they think are important. This will not only provide a large increase in the sample size of potential rankings, but it provides a larger buy-in from potential users to view the CMS+ as a viable college football ranking alternative.

References

[1] "The BCS Is..." *BCSFootball.org*. 1 Oct. 2013. Web.

[2] Bery, Soven. "NFL: Is the Most Popular Sports League in the USA Really Too Big To Fail?" *Bleacher Report*. 17 July 2013. Web.

[3] Cassady, C. Richard, Lisa M. Maillart, and Sinan Salman. "Ranking Sports Teams: A Customizable Quadratic Assignment Approach." *Interfaces* 35.6 (2005): 497-510. Print.

[4] "College Bowl Game Payouts." *Statistic Brain*. 1 July 2012. Web.

[5] Sullivan, Kelly M., and C. Richard Cassady. "The CMS+ System for Ranking College Football Teams." Proceedings of 2009 Industrial Engineering Research Conference, Doral Golf Resort and Spa, Miami.