Criteria for Assessing a Ranking Method

Final Report: Undergraduate Research Assistantship
Summer 2003

Gordon Davis: dagojr@email.arizona.edu
Advisor: Dr. Russel Carlson

One of the many questions that surround the study of tournaments is how to rank players relative to one another. The most obvious application involves the ranking of sports teams, but this also has uses in other areas. For example, Google.com uses a modification of the Kendall-Wei ratings in order to determine the order of webpages when returning searches. A player can be many things, anything from a web page such as in Google's pagerank system, to a team such as in the RPI college basketball ratings, to an individual competitor such as in the Elo ratings used in chess. A tournament is a collection of these players and shows their results of playing against one another. In this paper, we will examine a set of criteria we would like a tournament to satisfy, justify their use, and examine existing ranking methods in order to determine how well they achieve these criteria.

In order to quickly and easily refer to a tournament, it is often put into the form of a matrix. Each row and column represents a particular player. Where the row and column intersect represents a game between these two players. If a player beats another, then a one is placed in the entry where the winning player’s row intersects the losing player’s column, and a zero is placed where the losing player’s row intersects the winning player’s column. If two teams to do not play each other, then a zero is placed is added where the player’s rows and columns intersect. Note that there is always a diagonal of zeros in every tournament matrix, since a player cannot play themselves. Here’s a quick example of how to read a tournament matrix:

0	0	1	0	1
0	0	0	1	0
0	1	0	1	0
1	0	0	0	1
0	1	0	0	0

Let’s examine the third player. If we move along the third row from left to right, we can see that there are ones in the second and fourth entries, signifying a win against the second and fourth player. If we move down the third column, we can see a one in the first entry, signifying a loss. Remember that zeros are placed in the entries not only for losses but also when teams do not play, so it is important to examine both a player’s row and column when determining their overall record.

There are two specific types of tournaments that we will be referring to throughout this paper. The first of these are called complete tournaments. These are tournaments in which every player goes up against every other player exactly once. The second of we will call partial tournaments. These are tournaments in which every player goes up against another player no more than once. Also, a player's ranking refers to their place relative to the other players in the tournament (i.e. first, fourth, etc.) while a player's rating is the number assigned to them according to a certain ranking method, signifying that player's strength. Because many of the ranking systems studied and referred to throughout this paper are for sports, the terms player and team will be used interchangeably.

As mentioned before, we would like for there to be a way for us to rank the strength of players based upon their performance in a tournament. Below are four possible criteria that we would like a ranking method to satisfy. It should be noted that this list is not all inclusive, as there are potentially other such criteria that a ranking method should satisfy. However, these were chosen as a starting point to analyze ranking methods and to determine how well they are able to achieve these criteria. They are listed below, along with a short justification for each.

For the first criteria, let's first look at a complete tournament (one in which all of the players play each other). Let's say the outcome is this tournament is shown below:

$$
\begin{matrix}
0 & 1 & 1 & 1 & 1 & 1 \\
0 & 0 & 1 & 1 & 1 & 1 \\
0 & 0 & 0 & 1 & 1 & 1 \\
0 & 0 & 0 & 0 & 1 & 1 \\
0 & 0 & 0 & 0 & 0 & 1 \\
0 & 0 & 0 & 0 & 0 & 0
\end{matrix}
$$

As we can see, the first team has defeated all of its opponents. Therefore, all indications in this tournament suggest that this team is stronger than all of the players it has gone up against. It would not make sense then, for example, if the second team to be ranked higher than the first, since the first team has defeated the second, along with all of the other opponents that the second team has played. If a player defeats all of its opponents in a complete tournament, they should be ranked first. For our purposes, we will call this the "unbeaten criteria".

The next criteria will apply to partial tournaments. Consider the following example:

0	1	1	0	0	0
0	0	1	1	0	0
0	0	0	1	0	1
1	0	0	0	1	1
1	1	0	0	0	0
1	1	0	0	0	0

The first and the second team have both defeated the third team in this tournament. Let's assume for a moment that the first team is ranked higher than the second team according to a particular ranking method. Now let's remove the third team from this tournament:

0	1	0	0	0
0	0	1	0	0
1	0	0	1	1
1	1	0	0	0
1	1	0	0	0

Because both the first and the second team defeated the third player, their ranking relative to each other should remain unchanged with the third player removed (the first team should still be ranked higher than the second). This criterion ensures that one player who defeats a particular opponent does not benefit more than another player who defeats the same opponent. So in a partial tournament, if player A and player B both defeat (or both lose to) player C, and player A is ranked higher then player B, then player B should not be ranked higher then player A if player C is removed. We will call this the "equal win criteria". We apply this criterion only to partial tournaments, because the argument behind this criterion breaks down if there are multiple games between opponents. For example, two wins by one player should not necessarily carry the same weight as one win by another. It should be noted that this criteria only applies to the relative ranking of these two players and does not apply to their ranking relative to all of the other players in the tournament. For instance, since the sixth player was defeated by the third player, the sixth player could move up ahead of the first player after the third player was removed and still satisfy the criteria.

The third criteria will involve any tournament and is not limited to just complete or partial tournaments. We would like a tournament to take both wins and losses into consideration when determining a player's ranking. To do this, we create a criterion which states that if all of the losses are changed to wins, and all of the wins are changed to losses, then the order in which the teams are ranked should be reversed. Here is another way of looking at it. Let’s look at the tournament below:

0	1	1	1	1	0
0	0	1	1	0	1
0	0	0	1	1	1
0	0	0	0	1	1
0	1	0	0	0	1
1	0	0	0	0	0

Let’s say the teams are ranked in such a way that the first team is ranked first, the second team is ranked second, and so on. Now let’s take the transpose of the following tournament:

0	0	0	0	0	1
1	0	0	0	1	0
1	1	0	0	0	0
1	1	1	0	0	0
1	0	1	1	0	0
0	1	1	1	1	0

Note that taking the transpose changes all losses to wins and wins to losses. If we were to apply the ranking method to this new tournament, we would like for the order to be exactly reversed (i.e. the sixth team is ranked first, the fifth team is ranked second, and so forth). Some ranking methods reward teams for wins against strong opponents but do not penalize teams for losses against weak opponents. This is a situation we would like to avoid. In other words, we would like wins and losses to be equally weighted, without one counting more than the other. Because we would like the order or the rankings of players to reverse when wins and losses are interchanged, we will call this the “symmetry criteria”.

The fourth criteria will involve any tournament as well. It states that if a loss of a player is changed to a win, a different player ranked lower before the loss is changed into a win should not be ranked higher than the player who gains the win after this game is changed. This intuitively makes sense, because a player who gains the win should see their ranking relative to other players increase. Consider the example below:

0	0	0	1	0	1
0	0	1	0	0	0
1	0	0	1	0	1
0	1	0	0	1	1
0	1	1	0	0	1
0	0	0	0	0	0

Let’s examine the third player for a moment. As we can see, the third player has lost a game to the second player. Let’s assume that the third player is ranked higher than the fourth player according to some ranking method. Now, we will change the third player’s loss to the second player into a win:

0	0	0	1	0	1
0	0	0	0	0	0
1	1	0	1	0	1
0	1	0	0	1	1
0	1	1	0	0	1
0	0	0	0	0	0

The third player has now increased their standing with a win. What we do not want is another player, like the fourth player, who was previously ranked lower then the third player to be ranked higher then the third player after the loss is changed to a win. In other words, we would like wins to only benefit a player, not to hurt them. This will be called the “monotone criteria”.

In addition to these criteria, we would like to have a ranking method which creates few ties between opponents. Since the goal of a ranking method is to determine the order of the strength of players, such a method which creates numerous ties is problematic. However, it should be noted that sometimes it is very difficult to break a tie. Consider the following tournament:

0	1	0
0	0	1
1	0	0

Every player has lost one game and won one game. There is no real way to differentiate between the players just based on their records against each other.

There are in fact ranking methods that will break this tie, but will involve other factors such as the order in which games are player or the final score of the game. It is debatable as to whether or not either should be used to rank players against one another. But if we simply look at the outcomes of the games and not the order in which they are played, it is impossible to break this tie. However, there are other cases where two players may have the same number of wins but it is possible to differentiate between the two, as we will see shortly.

We should make a special note at this time about the equal win criteria after noting that there are situations like the above tournament where it is impossible to rank teams relative to each other. Let's illustrate why. Consider the following tournament:

0	1	0	0
0	0	1	1
1	0	0	1
0	0	0	0

Despite the second and third teams having the same number of wins, we can differentiate between these two teams. For example, only one of these players has defeated an opponent with two wins: the second team. Since we can differentiate between these two teams, it is possible to develop some ranking method that could rank these teams relative to each other. For instance, we could simply define a players rating to be the sum of the number of wins of all of the opponents they have defeated. The second player would have a rating of 2, and the third player a rating of 1. Many other ranking methods could be devised to rank these teams as well. Now let's remove the fourth player:

0	1	0
0	0	1
1	0	0

We now have the earlier case where it is impossible to differentiate between any of the teams, even though it was possible with the fourth player present. A ranking method just based on a player's wins and losses would not be able to differentiate between the second and third player once the fourth is removed. Therefore we should clarify that to satisfy the equal win criteria, a ranking method cannot change the ranking of two teams relative to each other if a team they both defeated or lost to is removed, although they may tie with each other once this team is removed. Now that we have all of our criteria, we will begin to examine some ranking methods.

The first method we'll look at is a very simple one. It consists of simply counting up the number of wins by a player, also called a player's score (not to be confused with the game's final score used in the margin of victory calculations). This ranking method satisfies the unbeaten criteria. Here is an explanation why. In a complete tournament, all of the players play the same number of games. If a player wins all of their games against every other player, then every other player must have at least one less win than the player who has won all of their games. Since the unbeaten team has the most wins, it will be ranked first. This also satisfies the equal win criteria. This is because the equal win criteria is only applied to players who have both defeated or both lost to the same player. Therefore, if this player is removed, both of these teams either have one less win or one less loss. Their rank relative to each must remain unchanged, since the difference in the number of wins between the two does not change as a result. The monotone criteria is also satisfied, since a win can only increase a player's score, and none of the other player's score can subsequently increase with the change of this one game. However, this method does not satisfy the symmetry criteria. Let's examine the tournament below:

0	1	1	0	0
0	0	0	0	1
0	0	0	1	0
1	0	0	0	0
1	0	0	0	0

Here we can see the first player has more wins then all of the other players shown here, and is therefore ranked first according the player's score. If this ranking satisfied the symmetry criteria, then the first team should be ranked last when we take the tournament's transpose. However, that is clearly not the case:

0	0	0	1	1
1	0	0	0	0
1	0	0	0	0
0	0	1	0	0
0	1	0	0	0

As we can see, the first player is still ranked first relative to all of the other players, having two wins. There is another ranking method, which is very similar to the wins ranking which does. If we take the wins minus the losses of a team, this will satisfy the symmetry criteria. Changing all of the wins to losses and losses to wins will simply create a sign change in a player's rating, and therefore the ranking of players will be reversed. But both of these methods can create numerous ties, as shown in the tournament below:

0	1	1	1	0	0	1	0
0	0	0	0	1	1	1	1
0	1	0	0	1	1	1	0
0	1	1	0	0	1	1	0
1	0	0	1	0	0	0	1
1	0	0	0	1	0	0	1
0	0	0	0	1	1	0	1
1	0	1	1	0	0	0	0

Both the wins and the wins minus losses ranking method would create ties between the first four teams, as well as the last four teams. However, it is possible to differentiate between these two teams. For example, the first player is the only of the first four players to have defeated three opponents who have four wins. Because of this, we need to find a ranking method that does not create ties like this between opponents who have the same score. Most rating systems are more complex then this in order to break ties such as in the situation shown above. We will examine some of these methods now.

The first we will look at is the RPI, one of the factors used in placing teams for the NCAA basketball tournament. The basic formula consists of adding 25% of a team's winning percentage, 50% of their opponent's winning percentage, and 25% of their opponent's opponent's winning percentage. In addition, games against the team being ranked are removed when calculating the opponent's winning percentage and the opponent's opponent's winning percentage. This does satisfy the unbeaten criteria. Let’s examine why this is so.

The first portion of the RPI formula is the team’s winning percentage. Obviously, a team that is undefeated will have a winning percentage of 1.000. All of the other teams in the tournament will have winning percentages less then this, since in a complete tournament, everyone plays each other, so the undefeated team must have beaten every other player once. The next part involves the opponent’s winning percentage. Remember that when calculating the opponent’s winning percentage, games against the team being ranked are removed. What this number will be is the following is the total number of games won by all of the other players divided by the total number of games played by all of these players. In every game, one player wins, and one player loses. Therefore, this number will always be .5. The same is true for the opponent’s opponent’s winning percentage. So in a complete tournament, a teams ranking is essentially:

$$.25 * WP + .375$$

WP stands for a teams winning percentage. Since an undefeated team will have the highest winning percentage, it will have the highest rating and be ranked first. Therefore, the unbeaten criteria is satisfied. However, the equal win criteria is not satisfied. Consider the following tournament:

0	0	0	0	0	0	0	1	0	0	0
1	0	1	0	0	0	0	0	0	0	0
0	0	0	0	1	1	1	0	0	0	0
0	0	1	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	1	0	0	0
0	0	0	0	0	0	1	0	0	0	0
0	1	0	0	0	0	0	0	0	0	0

The rating of the first player according to the RPI is .467, while the rating for the third player is .483. As we can see, both the first and the third player were defeated by the second player. Now let's remove this second player. Below is the tournament after this player is removed:

0	0	0	0	0	0	1	0	0	0
0	0	0	1	1	1	0	0	0	0
0	1	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	1	0	0	0
0	0	0	0	0	1	0	0	0	0
0	0	0	0	0	0	0	0	0	0

Now the rating of the first and third player, after the second was removed, are .500 and .438 respectively. The order of these two teams has switched as a result of this player being removed. The RPI doesn't satisfy the monotone criteria either, as shown thin the example below:

0	1	0	0	0
0	0	1	0	0
0	0	0	1	0
0	0	0	0	1
0	0	0	0	0

Here, the fourth player has a rating of .250, while the fifth player has a rating of .125. However, let's change the fourth player's loss to the third player into a win:

0	1	0	0	0
0	0	1	0	0
0	0	0	0	0
0	0	1	0	1
0	0	0	0	0

Now, the rating of the fourth player is .375, but the fifth player has now moved up past the fourth player to have a rating of .500. Although the RPI doesn't satisfy the two previous criteria, it does satisfy the symmetry criteria. Let's say a team in any tournament. Then, according to the RPI, their rating is:

.25 * WP + .5 * OWP + .25 * OOWP = RTG

Where WP is their winning percentage, OWP is their opponent's winning percentage, OOWP is there opponent's opponent's winning percentage, and RTG is their rating. Now, if we change all of their losses into wins and wins into losses, all of the percentages change into one minus what they were originally. So the new rating would be:

.25 * (1-WP) + .5 * (1 - OWP) + .25 * (1 - OOWP) = RTG'
1 - (.25 * WP + .5 * OWP + .25 * OOWP) = RTG'
1 - RTG = RTG'

So as we can see, using the RPI on the transpose of the tournament will simply change the sign of the player's ratings and add one to it. Therefore, the order of the teams will simply be reversed, and the symmetry criteria is satisfied. Let's now take a look at another system for ranking players, called the Colley matrix method. First we'll see if it satisfies the unbeaten criteria. Let's go over this ranking method. Consider the tournament shown below:

0	1	1	0
0	0	1	0
0	0	0	1
0	1	0	0

What the Colley matrix does is create a system of equations, which a variable corresponding to each player, and solves these equations. The value of the variable is then considered the rank of the team. For each player, an equation is written. It consists of multiplying the player's variable by two plus the total number of games they have played, and subtracts the variables of the players they have played. This equals the number of wins by that opponent divided by two. For example, the first equation would look like this:

$4x_1 - x_2 - x_3 = 1$

The variables x_1, x_2, x_3 and x_4 correspond to the first, second, third, and fourth teams respectively. Since the first player has played two games, it is multiplied by the number of games it has played plus two. The first player went against the second and third player, so these variables are subtracted. The first player did not play the fourth, so it is not included in the equation. Finally, this equation equals 1, since the first player won two games, divided by two. The whole system of equations, in matrix form, looks like this

4	-1	-1	0	1
-1	5	-1	-1	½
-1	-1	5	-1	½
0	-1	-1	4	½

The matrix is then solved using Gaussian elimination:

1	0	0	0	.396
0	1	0	0	.292
0	0	1	0	.292
0	0	0	1	.271

And we have the solutions to the variables, which correspond to each teams ranking. For example, the first player has a rating of .396, the value of x_1. Let's look at a complete tournament with 'n' number of players. An equation for any player 'i' would look like this:

$-x_1 - x_2 - ... - x_{i-1} + (2 + n) * x_i - x_{i+1} - ... - x_n = wins/2$

Let's say that the first team has defeated all of their other opponents. Then, their equation looks like this:

$(2 + n) * x_1 - x_2 - ... - x_n = n/2$

Since they have defeated all of their other opponents, they have n wins, and hence this equation equals n/2. For any other team, the following must also be true:

$-x_1 - x_2 - ... - x_{i-1} + (2 + n) * x_i - x_{i+1} - ... - x_n < n/2$

This is because any team that isn't the first team must have less then n wins, since the first team has defeated them. And now, we perform some algebra:

$-x_1 - x_2 - ... - x_{i-1} + (2 + n) * x_i - x_{i+1} - ... - x_n < (2 + n) * x_1 - x_2 - ... - x_n$
$-x_1 + (2 + n) * x_i < (2 + n) * x_1 - x_i$
$(3 + n) * x_i < (3 + n) * x_1$
Since $n > 0$, $3 + n > 0$, so
$x_i < x_1$

Since the first player will have a higher rating then any other player, the unbeaten criteria is satisfied. However, the equal win criteria is not. Consider the following tournament:

0	1	1	0	0
0	0	0	0	0
0	0	0	0	0
1	0	0	0	0
1	0	0	0	0

We now put it into the Colley matrix:

6	-1	-1	-1	-1	1
-1	3	0	0	0	0
-1	0	3	0	0	0
-1	0	0	3	0	½
-1	0	0	0	3	½

And solve:

1	0	0	0	0	.286
0	1	0	0	0	.095
0	0	1	0	0	.095
0	0	0	1	0	.262
0	0	0	0	1	.262

We can see that the first team is ranked first. Therefore, if the symmetry criteria is to be satisfied, the first team should be ranked last if we take the tournament's transpose. However, there is a problem once we do this:

0	0	0	1	1
1	0	0	0	0
1	0	0	0	0
0	0	0	0	0
0	0	0	0	0

This is essentially the same tournament we just ranked. If we simply rearranged the order in which we put the teams in the first tournament matrix, we could get the following tournament matrix as well. Since we are ranking the same tournament, the first player will still be ranked first, and the symmetry criteria is not satisfied.

However, there is a way in which we can modify the Colley matrix method, or any ranking method for that matter, in order to satisfy the symmetry criteria. This method of modification is used in the balanced Perron method, based upon the Kendall-Wei method. While the Kendall-Wei method itself does not satisfy the symmetry criteria, the balanced Perron method does. Here is how the modification works. First, the ratings for each player are obtained through the Kendall-Wei method. Then, the tournament's transpose is taken, and the Kendall-Wei rankings are calculated for each player. Then, the first number obtained is subtracted from the second, and that result is the player's ratings. Let's look for a moment as to why this works. We can look at a system which assigns numbers to each player in a tournament as a function. Given a tournament and the number of the player, we can find a rating for that player. In other words:

$$Rtg = f(A,x)$$

Where Rtg is the rating of a particular player according to a certain method, f is the function which determines a players rating, A is the tournament matrix, and x is the number of the player as they appear in the tournament. Now, we will define a new function:

$$g(A,x) = f(A,x) – f(A^T,x)$$

A^T is the tournament matrix's transpose, and g is the function for our new ranking method. So now:

$$Rtg = g(A,x) = f(A,x) – f(A^T,x)$$

If we take the transpose of A, that simply changes all of the losses to wins and wins to losses, as explained earlier. Let's take the transpose of A and see what that will now do to the player's rating:

$$Rtg' = g(A^T,x) = f(A^T,x) – f((A^T)^T,x) = f(A^T,x) – f(A,x) = –(f(A,x) – f(A^T,x))$$
$$Rtg' = –Rtg$$

As we can see, this simply results in a sign change from our original ratings, and therefore satisfies the symmetry criteria. We can do the same thing to the Colley matrix method in order to achieve symmetry as well. All we need to do is take the rating for the original tournament matrix, take the rating from the transpose of the tournament matrix, and subtract the former from the latter. This method can then be applied to essentially any ranking method in order to obtain symmetry. However, the effect that this would have on the other three criteria mentioned earlier or upon creating ties between players is unclear.

But this still does not solve a problem that both the RPI and the Colley matrix method both have. These methods, among others studied as well, will create ties in complete tournaments between teams that have the same number of wins, even if it is possible to differentiate between the teams. However, there are a few ranking methods which are able to break this tie. This includes the Kendall-Wei rankings, along with the balanced Perron rankings mentioned earlier. The ratings for the Kendall-Wei rankings are derived from the following vector:

$$\lim_{k\to\infty} T^k x \setminus |T^k x|$$

T is the tournament matrix, and x is a nonzero real vector. The values in this vector correspond to each team. For example, the first number in the vector is the rating of the first team in the tournament. In practice, we can use large numbers of k so that the values are close to the limit, and use those values as team's ranking. Let's use the following tournament as an example to see how this ranking method works:

0	1	1	0
0	0	1	1
0	0	0	1
1	0	0	0

We will let k = 100, and let x be a vector with all values equal to 1. The result is the following vector:

[.626 .552 .321 .488]

For the balanced Perron rankings, we would take the tournaments transpose, repeat the process we used for the Kendall-Wei on that matrix, and subtract the two vectors. The resulting vector is:

[.177 .230 -.230 -.177]

However, let's examine for a moment whether or not this satisfies the equal win criteria. Let's examine the following tournament:

0	1	1	0	0	0
0	0	1	1	0	0
0	0	0	1	1	1
1	0	0	0	1	1
1	1	0	0	0	1
1	1	0	0	0	0

The vector obtained using the Kendall-Wei method is:

[.359 .386 .488 .451 .432 .306]

While the vector from the balanced Perron method is:

[-.092 -.103 .103 .092 .126 -.126]

Now, let's examine the third and fourth teams. Note that the third player is ranked higher then the fourth. Both of these teams defeated the fifth and sixth players. If this ranking method satisfied the equal win criteria, then their ranking relative to each other should remain unchanged if these players are removed. Let's remove these two players from the tournament:

0	1	1	0
0	0	1	1
0	0	0	1
1	0	0	0

We now have the tournament we had in the earlier example. However, recall that both the Kendall-Wei and the balanced Perron rankings placed the third player higher then the fourth. So while these rankings are able to break the tie in complete tournaments, they are still unable to satisfy the equal win criteria. In fact, none of the tournaments that were studies were able to break the tie between these players and satisfy the equal win criteria in the example just presented. It’s uncertain whether this is a short coming of the tournaments studied, or if it may be impossible to break ties in these types of situations and still satisfy this criteria.

Another way in which ranking methods attempt to rank teams and break ties is by including margin of victory in the calculations. Margin of victory is simply the difference in the score of the game between two players. Using this in determining a player’s rank is controversial for several reasons, but we will just look at how it applies to satisfying our four criteria. Margin of victory seems to present the most problems with the unbeaten criteria. Let's examine the Heineman ratings, which factor margin of victory into the calculations. The Heineman ratings formula is as follows:

$$Rtg = 3 * (WP + OWP + OOWP) * MaxMV + MV$$

Where Rtg is the teams rating, WP is the winning percentage, OWP is the opponent's winning percentage, OOWP is the opponent's opponent's winning percentage, MaxMV is the maximum margin of victory obtained by any team, and MV is the teams' own margin of victory. Now let's consider the tournament below:

0	1	1	1	1	1	1	1
0	0	1	1	1	1	1	1
0	0	0	1	1	1	1	1
0	0	0	0	1	1	1	1
0	0	0	0	0	1	1	1
0	0	0	0	0	0	1	1
0	0	0	0	0	0	0	1
0	0	0	0	0	0	0	0

Let's assume that the first player has won all of their games by an average of 2 points, and that the maximum margin of victory for any player was 15 points. This player would have a rating of 92. Now let's assume the third player had an average margin of victory of 15 points. Their rating would be 92.143. Therefore, the unbeaten criteria is not satisfied, since the first team, which is undefeated, is not ranked first. If margin of victory had not been included in these calculations, and we simply added the winning percentage, the opponent's winning percentage, and the opponent's opponent's winning percentage, the unbeaten criteria would have been satisfied in this tournament. The problem with most of the rankings that use margin of victory in their calculations is that there's not a way to prevent lopsided victories from ranking teams high. As in the case above, a team consistently won all of their games by a small margin of victory, but the other team won their games by a considerably larger margin of victory, despite losing games as well. As a result, the team with the larger margin of victory then passes the undefeated team over in the rankings. If margin of victory is included in the calculations, it would need to be limited if the unbeaten criteria is to be satisfied. Now that we have looked in depth at a few specific example studied, here is a table showing all of the tournaments studied, and which criteria they do or do not satisfy:

Ranking method	Unbeaten Criteria	Equal win Criteria	Symmetry Criteria	Monotone Criteria
Wins	+	+	-	+
Wins minus losses	+	+	+	+
RPI	+	-	+	-
Colley Matrix method	+	-	-	0
Kendall-Wei method	+	-	-	0
Balanced Perron method	+	-	+	0
Greg Heineman' ratings	-	-	+	0
Elo ratings	-	-	+	-
Sauceda ratings	-	-	+	-
John Wobus' ratings	0	-	+	0
Holland's teampoint ratings	+	-	-	-
Bret Heppner's ratings	0	-	-	-
Elrod ratings	+	-	+	0
Todd Brown's ratings	-	-	+	0
Square Gear ratings	0	-	+	0
Cody Kellner ratings	0	-	+	0
Wilson ratings	0	-	+	0

A '+' indicates that the criteria was satisfied, a '–' indicates the criteria was not, and a '0' indicates that it could not be determined if the criteria is or is not satisfied.

One of the more striking parts of this table is the large number of tournaments which do not satisfy the equal win criteria. Note that the only tournaments that do satisfy this criteria are the wins and the wins minus losses ratings, both of which can create a large number of ties between opponents. All of the other ratings, however, fail to satisfy this criteria. This could mean that it may be impossible to break ties between opponents in certain tournaments and still satisfy this criteria. However, it could also mean that this is simply a difficult criteria to satisfy, and that other ranking systems could exist or could be developed which to satisfy the equal win criteria.

Also interesting is the large number of tournaments in which it was unable to be determined whether or not the monotone criteria was satisfied. This has a lot to do with the nature of the monotone criteria itself. Changing the result of one game also changes the ratings of several different teams, and therefore become very difficult to prove it satisfies this criteria. Also, some of these ranking methods use several iterations of a certain algorithm in order to find these ratings, consequentially making this task even more difficult. We would expect the equal win criteria to present the same sort of problems, since removing a team changes the entire dynamics of the tournament as well. However, in all of the ranking methods studied, counter examples could be found the equal win criteria and it was not necessary to provide a lengthy proof.

There is another issue not dealt with in this paper but related to the monotone criteria. Let's say that if a loss by player A is changed to a win, their rating increases by 25, and another player B's rating increases by 24. While this would satisfy the monotone criteria, we may want a rating system that does not reward player B this much relative to player A simply because player A gained a win. This is a problem that could potentially be investigated further.

Another note is that many of the counter examples used to show these ranking methods did not satisfy a certain criteria are unusual tournaments which do not resemble the types of tournaments they are designed rank. For instance, the counter example used to show the RPI did not satisfy the monotone criteria involved five teams, and the teams only played one or two games. However, the RPI is used to ranked college basketball teams, which involves over a hundred teams which play dozens of games. So while the RPI might not satisfy the monotone criteria in all tournaments, this does not necessarily mean that the RPI does not satisfy this criteria for the purposes of ranking college basketball teams and the type of schedule they play. Nonetheless, this example of where the RPI does not satisfy the monotone does raise suspicions as to whether or not the RPI satisfies the monotone criteria when it comes to ranking college basketball teams.

As stated before, these criteria do not make up an all inclusive list, and may need to be changed or adapted. Rather, these criteria instead serve as a starting point for analyzing ranking methods and assessing their strengths and weaknesses. These criteria could also possibly be used as a guideline to developing some other kind of ranking method that satisfies all of these criteria.