A Solution To The Unequal Strength Of Schedule Problem

Roy Bethel

rbethel@mitre.org

August, 2005

Contents

1 Introduction And Overview

For purposes of determining standings in sports (or possibly other) league play, the normal way to rank teams is winning percentage. For a *balanced* schedule, where each team plays every other team an equal number of games over the season, this is the equitable, or fair, method and is universally accepted as the preferred method. A problem arises for an *unbalanced* schedule, where each team does not play every other team an equal number of games over the season even though the total number of games each team plays is usually the same. Each team's schedule, in general, will not have the same strength.

A solution is required which ranks teams equitably for unbalanced and inherently unequal schedules. The desired solution is *earned*, or *retrodictive*, with the objective being the fair ranking of teams based only on past won-lost performance, rather than *predictive*, with the objective being prediction of won-lost outcomes for future games. For balanced schedules, the solution must be identical to winning percentage. The solution must depend only on each team's schedule and won-lost outcomes of games.

A solution meeting these conditions is presented. The solution is developed both heuristically with winning percentage as a baseline and mathematically from principles of applied statistical inference. The primary result is team strengths, which are assumed numerical attributes reflecting teams' ability to win games and are computed for all teams from game won-lost outcomes. The strengths are used to rank teams and may also be used to compute a projected winning percentage for each team as if it had played a hypothetical balanced schedule over the season. The projected winning percentage of a team either improves or worsens over its actual winning percentage depending on strength of schedule of that team and, thus, provides a quantification of strength of schedule. Results are computed from the 1999 NFL regular season schedule as an example.

2 Background And Roadmap

Most of the current literature and interest in sports ranking methods centers around the Bowl Championship Series (BCS) for college football [5] [17]. The current BCS computer based methods are [1], [4], [7], [10], [12], [14], and [20]. However, the methods are general and not limited to college football. Good sources of information are [16] and [18]. From a review of the literature, sports ranking methods fall into two categories. The first is *predictive*. The second is *earned*, or *retrodictive*. There is a difference in objective between them, which leads to a difference in assumptions, approach, and solution.

The purpose of a *predictive* method is prediction of won-lost outcomes for *future* games. Rankings, or computed measures of strength, of teams are a means to this end. There is a simple self-evident metric for judging performance of a predictive method: its success in actually predicting future won-lost outcomes. Therefore, the assumptions, approach, and solution of a predictive method do not have to be justified in terms of fairness or any other criterion except the ability to predict future won-lost outcomes of games. All other considerations are subsumed in this ability. Any available data such as home advantage, weather conditions, injuries to key personnel, etc. may be used in a predictive method in addition to past won-lost performance.

On the other hand, the purpose of an *earned*, or *retrodictive*, method is the fair ranking of teams for purposes of league standings based only on past game won-lost performance. Possible implications are awarding of championship trophies or selection and seeding of teams for post-season play. By far, the most common earned method is winning percentage or closely related methods such as the point system in the NHL to account for ties. For a balanced schedule, winning percentage is the equitable method and is universally accepted as the preferred method. The only issue for earned methods is unbalanced schedules, where strengths of schedule over the season are, in general, unequal. Re-stated, the objective of an earned method is to find an equitable metric of strength per team which is equivalent to winning percentage for a balanced schedule and plays the role of winning percentage for an unbalanced schedule.

There is no performance metric or measure of success for an earned method other than its perceived fairness and acceptance. An earned method *can* be used for predictive purposes, but to do so misses the point of its intended purpose. As an example, consider the ideal situation of a balanced schedule over the season with winning percentage used both to determine league standings and to select teams for post-season play. Since winning percentage is *a priori* agreed upon as a metric and is accepted as fair, the ranking results and post-season selections are accepted as fair regardless of any other factors such as won-lost outcomes of games in post-season play.

Based on these considerations, the challenge for an earned method is to be accepted, or perceived, as fair for a general, or unbalanced, schedule over the season. As contrasted to a predictive method, the assumptions, approach, and solution of an earned method must be fully explained and justified with all details. A discussion on this topic is in [15].

3 Design Rationale

A top down design rationale is used to justify and develop the presented method. There are two fundamental assumptions. First, the presented method is firmly rooted in the earned method camp. All discussion of earned methods in Section 2 applies. Second, winning percentage is embraced both as the preferred method for a balanced schedule and as a baseline for a general, or unbalanced, schedule. In particular with a balanced schedule as a special case, the presented method must be equivalent to winning percentage for purposes of ranking teams. Because of acceptance of winning percentage as a method, all assumptions implicit in computation of winning percentage are also accepted.

Any earned method must ultimately reduce to a computed univariate metric of strength per team for purposes of ranking teams. Winning percentage is a univariate metric of team strength, which is sufficient for a balanced schedule. However, for an unbalanced schedule, winning percentage as a metric of team strength has difficulties due to potentially unequal strengths of schedule since winning percentage obviously depends on schedule. Therefore, a univariate numerical *strength* attribute is assumed for each team to perform this role. The assumption is that strengths do not *depend* on schedule but rather *determine* winning percentages for a given schedule and are *observable* through winning percentages for a given schedule.

The last assumption is the key to the presented method and, in fact, is the foundation. Based on this assumption, there must exist invertible transformations between strengths and winning percentages for a given schedule. Specifically, winning percentages must be mathematically predictable by strengths, and vice-versa, for a given schedule. This property must be true for any arbitrary schedule, actual or hypothesized. In the computational procedure, strengths are computed from the actual winning percentages for the actual schedule. Then, *projected* winning percentages for a hypothetical balanced schedule are computed from these strengths. Since winning percentage as a ranking method is accepted as fair for a balanced schedule, the resultant ranking from these projected winning percentages is accepted as fair. This viewpoint, of course, places a burden on acceptance of the transformations as fair. In particular, computation of strengths must be maximally compatible and consistent with winning percentage while still equitably taking into account differences in strength of schedule.

For a balanced schedule, winning percentages and strengths must be monotonic with each other. This is a required property of the presented method. Therefore, either projected winning percentages for a hypothetical balanced schedule or computed strengths may be equivalently used to rank teams. If the schedule is actually balanced, actual winning percentages are projected winning percentages and, thus, are also equivalent to computed strengths and projected winning percentages for purposes of ranking teams. This property satisfies the requirement of equivalence of the presented method to winning percentage for an actual balanced schedule. The equivalence of computed strengths and projected winning percentages to actual winning percentages is, in general, not true for an actual unbalanced schedule, which is the primary point of this paper. The projected winning percentage of a team either improves or worsens over its actual winning percentage depending on strength of schedule of that team. Thus, strength of schedule may be quantified by this difference.

Consistent with computation of winning percentage, only game won-lost outcomes are used as input data. Specifically, margins of victory have no influence. Tie games are also not included since there is no won-lost outcome. The use of only won-lost outcomes precludes use of any subjective factors such as home advantage, injuries, etc., which may be relevant and allowable for a predictive method. In addition, strengths are computed as an aggregate over the entire season. That is, all games count equally.

4 Approach

Notation is introduced based on discussion in Section 3. A *league* consists of T *teams*. Each team is designated by t $(t{=}1,\ldots,T)$. The *schedule* for the entire league over the *season* consists of G *games*. Each game g $(g{=}1,\ldots,G)$ is played between two teams t and t'. There are two allowed outcomes for each game g: $1)$ team t wins, team t' loses $2)$ team t' wins, team t loses. The *winner* of game g is w_g, and the *loser* of game g is l_g. In the case of a tie, that game is discounted since it contains no won-lost information. Actual won-lost outcomes w_g,l_g for all G games are the data. A univariate numerical *strength* attribute s_t is assumed for each team t, which is the only measure of team t's ability to win games. Strengths s_t $(t{=}1,\ldots,T)$ are computed from actual game won-lost outcomes w_g,l_g $(g{=}1,\ldots,G)$. The ranking of teams t depends on their relative strengths s_t.

An approach is required which follows the design rationale in Section 3. Let W_t be the *actual number of wins* for team t and is computed as

$$W_t = \operatorname*{count}_{g=1,\ldots,G}(w_g = t) = \sum_{\substack{g=1 \\ w_g=t}}^{G} 1 \qquad t = 1,\ldots,T \tag{1}$$

Actual number of wins W_t in (1) is the count, or summation, over all G games of the number of games g where winner w_g is team t. Actual number of wins W_t is redundant with actual number of losses L_t and actual winning percentage $W_t/(W_t{+}L_t)$ so that only actual number of wins W_t need be considered.

Restructure actual number of wins W_t in (1) as a summation over games g involving team t.

$$W_t = \sum_{\substack{g=1 \\ w_g=t,l_g=t' \\ w_g=t',l_g=t}}^{G} \mathrm{f}_{tt'}(w_g,l_g) \qquad t = 1,\ldots,T \tag{2}$$

$$\mathrm{f}_{tt'}(w_g, l_g) = \begin{cases} 0 & w_g = t', l_g = t \\ 1 & w_g = t, l_g = t' \end{cases} \tag{3}$$

Teams t' in (2) and (3) are opponents of team t in games g involving team t. Specifically, team t' is loser l_g in game g where winner w_g is team t, and team t' is winner w_g in game g where loser l_g is team t. Function $\mathrm{f}_{tt'}$ in (2) and (3) may be viewed as a *game outcome function* [15], which is restricted to 0,1 for winning percentage computation.

Let $\widehat{W_t}$ $(t{=}1,\ldots,T)$ be the *predicted number of wins* for team t as a function of all strengths $s_{t'}$ $(t'{=}1,\ldots,T)$ given team t's schedule. From Section 3, winning percentages must be mathematically predictable by strengths s_t, and vice-versa, for a given schedule. Equivalently, actual number of wins W_t in (1) and (2) must equal predicted number of wins $\widehat{W_t}$ for all T teams, given each team t's schedule.

$$W_t = \widehat{W_t} \qquad t = 1,\ldots,T \tag{4}$$

A mathematical expression for predicted number of wins $\widehat{W_t}$ in (4) as a function of strengths t' given team t's schedule must be determined. Because of the requirement of compatibility with winning percentage, the same functional form as actual number of wins W_t in (2) is assumed for predicted number of wins $\widehat{W_t}$ (summation over games g involving team t).

$$\widehat{W_t} = \sum_{\substack{g=1 \\ w_g=t,l_g=t' \\ w_g=t',l_g=t}}^{G} \widehat{\mathrm{f}}_{tt'}(s_t, s_{t'}) \qquad t = 1,\ldots,T \tag{5}$$

Teams t' in (5) are opponents of team t in games g involving team t similar to (2) and (3). Since predicted number of wins $\widehat{W_t}$ can depend only on strengths $s_{t'}$ $(t'{=}1,\ldots,T)$ given team t's schedule, functions $\widehat{\mathrm{f}}_{tt'}$ in (5) collectively also can depend only on strengths $s_{t'}$ $(t'{=}1,\ldots,T)$ given team t's schedule. For an individual game g between teams t and t' in (5), function $\widehat{\mathrm{f}}_{tt'}(s_t, s_{t'})$ can depend only on their strengths s_t and $s_{t'}$ since no other teams are involved.

Function $\widehat{\mathrm{f}}_{tt'}$ in (5) is the analog of function $\mathrm{f}_{tt'}$ in (2) and (3) and may be viewed as a *game outcome prediction function* [15]. Its purpose is to provide a measure of the won-lost outcome for game g between two teams t and t' but depending only on their strengths s_t and $s_{t'}$. Ideally, if the won-lost outcome of a game could always be deterministically predicted by the relative strengths of the two teams, function $\widehat{\mathrm{f}}_{tt'}$ in (5) would be 0 or 1 similar to function $\mathrm{f}_{tt'}$ in (3).

Clearly, the won-lost outcome of a game cannot be deterministically predicted. The stronger team does not win all games, only "most" games. The consequence is that, in general, strengths s_t $(t{=}1,\ldots,T)$ satisfying (4) cannot be found following the deterministic model above for function $\widehat{\mathrm{f}}_{tt'}$ (0 or 1) in (5).

In a game between two teams, the best that can be done is determination of the probabilities of each team winning. The relative strength between the two teams manifests itself in these probabilities. Even though the winner of an individual game is random and, therefore, not predictable, the percentage of wins by each team *is* predictable by these probabilities. This is an application of the principle of *frequency substitution* in the theory of probability [3]. "Predicted" is used here in the context that function $\widehat{\mathrm{f}}_{tt'}$ in (5) depends only on strengths s_t and $s_{t'}$ for games that have already been played as contrasted with function $\mathrm{f}_{tt'}$ in (3), which uses actual won-lost outcomes w_g, l_g of those games. Thus, games are "retroactively predicted", or *retrodicted*.

Let $p_{tt'}$ be the *game won-lost outcome probability*, which is the probability that team t wins and team t' loses in a game between team t and team t'. From discussion above, the average, or predicted, number of wins for team t against team t' is game won-lost outcome probability $p_{tt'}$ times the number of games between team t and team t'. As a generalization, the average, or predicted, number of wins for team t over its schedule is the sum of game won-lost outcome probabilities $p_{tt'}$ over all teams t', which are opponents of team t in all its games. This last equality follows the functional form of predicted number of wins $\widehat{W_t}$ in (5).

Based on arguments above, function $\widehat{\mathrm{f}}_{tt'}$ in (5) is assumed to be game won-lost outcome probability $p_{tt'}$. For a predictive method, there are no restrictions on determination of game won-lost outcome probability $p_{tt'}$. Assumptions such as *stationarity* (invariance over time) and *transitivity* (existence of monotonicity over $p_{tt'}$'s) need not be made. The goal of a predictive method is accurate determination of game won-lost outcome probability $p_{tt'}$ by any means using all information, following arguments in Section 3. For an earned method, function $\widehat{\mathrm{f}}_{tt'}$ and, thus, game won-lost outcome probability $p_{tt'}$ can depend only on strength s_t and strength $s_{t'}$, also following arguments above and in Section 3.

Substitute game won-lost outcome probability $p_{tt'}$ for function $\widehat{\mathrm{f}}_{tt'}$ into predicted number of wins $\widehat{W_t}$ in (5).

$$\widehat{W_t} = \sum_{\substack{g=1 \\ w_g=t, l_g=t' \\ w_g=t', l_g=t}}^{G} p_{tt'}(s_t, s_{t'}) \qquad t = 1, \ldots, T \tag{6}$$

The remaining task is determination of game won-lost outcome probability $p_{tt'}(s_t, s_{t'})$ in (6). Game won-lost outcome probability $p_{tt'}(s_t, s_{t'})$ in (6) has three required mathematical properties: 1) $0 \le p_{tt'} \le 1$ 2) $p_{tt'} + p_{t't} = 1$ 3) $p_{tt'}$ monotonically increases with strength s_t and monotonically decreases with strength $s_{t'}$. The first two properties are from the theory of probability [3] since function $\widehat{\mathrm{f}}_{tt'}$ in (5) has been placed into a probabilistic context. In the first property, limits 0,1 on game won-lost outcome probability $p_{tt'}$ correspond to the deterministic case above on game won-lost outcomes. The second property results from $p_{tt'}$ and $p_{t't}$ being the probabilities of the only two allowed outcomes of a game. The general solution for game won-lost outcome probability $p_{tt'}(s_t, s_{t'})$ to satisfy these required properties is

$$p_{tt'}(s_t, s_{t'}) = \frac{\mathrm{h}(s_t)}{\mathrm{h}(s_t) + \mathrm{h}(s_{t'})} \tag{7}$$

Function $\mathrm{h}(x)$ in (7) is any non-negative monotonically increasing function. There is no reason to consider any other function for $\mathrm{h}(x)$ except $\mathrm{h}(x) = x$. Otherwise, strength s_t would effectively be re-defined as $\mathrm{h}(s_t)$ since game won-lost outcome probability $p_{tt'}$ in (7) is observable only through function $\mathrm{h}(x)$. Then, game won-lost outcome probability $p_{tt'}$ in (7) becomes

$$p_{tt'}(s_t, s_{t'}) = \frac{s_t}{s_t + s_{t'}} \tag{8}$$

From (8), a required property of strength s_t is non-negativeness. Two additional properties of game won-lost outcome probability $p_{tt'}$ from (8) are 1) $p_{tt'} \propto s_t/s_{t'}$ 2) $p_{tt'}/p_{t't}{=}s_t/s_{t'}$. In the first property, game won-lost outcome probability $p_{tt'}$ depends only on the ratio of strengths s_t and $s_{t'}$ and monotonically increases with this ratio. In the second property, the ratio of the relative likelihoods, or *odds*, of teams t and t' winning is the same as the ratio of their strengths s_t and $s_{t'}$.

Substitute game won-lost outcome probability $p_{tt'}$ from (8) into predicted number of wins $\widehat{W_t}$ in (6).

$$\begin{aligned}\widehat{W_t} &= \sum_{\substack{g=1\\ w_g=t,l_g=t'\\ w_g=t',l_g=t}}^{G} p_{tt'} = \sum_{\substack{g=1\\ w_g=t,l_g=t'\\ w_g=t',l_g=t}}^{G} \frac{s_t}{s_t+s_{t'}} \qquad t=1,\ldots,T \\ &= \sum_{\substack{g=1\\ w_g=t}}^{G} p_{tl_g} + \sum_{\substack{g=1\\ l_g=t}}^{G} p_{tw_g} = \sum_{\substack{g=1\\ w_g=t}}^{G} \frac{s_t}{s_t+s_{l_g}} + \sum_{\substack{g=1\\ l_g=t}}^{G} \frac{s_t}{s_t+s_{w_g}} \end{aligned} \tag{9}$$

Each summation in (9) over all games g involving team t can be broken up into two summations. The first summation is over games g where winner w_g is team t and its opponent t' is loser l_g. The second summation is over games g where loser l_g is team t and its opponent t' is winner w_g.

The approach is summarized. In (4), the requirement is that the actual number of wins for every team t $(t{=}1,\ldots,T)$, which is W_t in (1), must equal the predicted number of wins for team t, which is $\widehat{W_t}$ in (9) and depends on strengths $s_{t'}$ $(t'{=}1,\ldots,T)$. This requirement is equivalent to actual and predicted winning percentages being equal and is the basis of the approach. Desired strengths s_t $(t{=}1,\ldots,T)$ must be found to satisfy the requirement in (4) and thus mathematically predict actual winning percentages as required. These strengths are used to rank teams and may be used to compute projected winning percentages for a hypothetical balanced schedule.

5 Maximum Likelihood Estimation Approach

This problem can be viewed as a statistical parameter estimation problem. Probability $p_{tt'}$ in (8) is a probabilistic model for a game won-lost outcome with strengths s_t as known parameters. In fact, strengths s_t are desired unknown parameters. The data to estimate strengths s_t $(t{=}1,\ldots,T)$ are actual won-lost outcomes w_g,l_g for all G games $(g{=}1,\ldots,G)$.

Rewrite game won-lost outcome probability $p_{tt'}$ in (8) as probability $p_{w_gl_g}$ of a hypothesized won-lost outcome for each game g between teams t and t' for all G games. From Section 4, there are two hypotheses for each game g: 1) $w_g{=}t,l_g{=}t'$ 2) $w_g{=}t',l_g{=}t$.

$$p_{w_gl_g} = \frac{s_{w_g}}{s_{w_g}+s_{l_g}} \qquad g=1,\ldots,G \tag{10}$$

With an independent games assumption, total probability P for all G games is the product of game won-lost outcome probability $p_{w_g l_g}$ in (10) of each game g over all G games.

$$P = \prod_{g=1}^{G} p_{w_g l_g} = \prod_{g=1}^{G} \frac{s_{w_g}}{s_{w_g} + s_{l_g}} \tag{11}$$

Total probability P in (11) is the probability of any particular hypothesized game won-lost outcome set of winners w_g and losers l_g for all G games out of 2^G possible sets. Strengths s_t in (11) are fixed known parameters.

Maximum likelihood estimation is a widely used mathematical parameter estimation technique in the theory of *applied statistical inference* [3]. Rewrite total probability P in (11) as the same equation but with an applied statistical inference interpretation as contrasted with the probabilistic interpretation in (11).

$$\Lambda = \prod_{g=1}^{G} p_{w_g l_g} = \prod_{g=1}^{G} \frac{s_{w_g}}{s_{w_g} + s_{l_g}} \tag{12}$$

Winners w_g and losers l_g in (12) are the fixed actual game won-lost outcome data. Strengths s_t in (12) are desired unknown parameters. Then, Λ in (12) is the *likelihood function* (LF) for any particular hypothesized set of strengths s_t. The applied statistical inference interpretation in (12) is the situation at hand. The set of strengths s_t $(t{=}1,\ldots,T)$ that maximizes LF Λ in (12) comprises the *maximum likelihood estimates* [3] of strengths s_t and is the desired solution.

Log likelihood function (LLF) λ is mathematically easier to work with than LF Λ and is equivalent for maximization. From LF Λ in (12), LLF λ is

$$\begin{aligned} \lambda = \ln(\Lambda) &= \sum_{g=1}^{G} \ln\left(p_{w_g l_g}\right) \\ &= \sum_{g=1}^{G} \ln\left(\frac{s_{w_g}}{s_{w_g} + s_{l_g}}\right) \end{aligned} \tag{13}$$

The necessary conditions for maximization of LLF λ and LF Λ are that each partial derivative of LLF λ in (13) with respect to strength s_t $(\partial\lambda/\partial s_t)$ for all strengths s_t $(t{=}1,\ldots,T)$ must equal zero. Rewrite LLF λ in (13) as LLF λ_t with an orientation for team t.

$$\begin{aligned} \lambda = \lambda_t &= \sum_{\substack{g=1 \\ w_g=t}}^{G} \ln\left(p_{t l_g}\right) + \sum_{\substack{g=1 \\ l_g=t}}^{G} \ln\left(p_{w_g t}\right) + \ \ldots \qquad t = 1,\ldots,T \\ &= \sum_{\substack{g=1 \\ w_g=t}}^{G} \ln\left(\frac{s_t}{s_t + s_{l_g}}\right) + \sum_{\substack{g=1 \\ l_g=t}}^{G} \ln\left(\frac{s_{w_g}}{s_{w_g} + s_t}\right) + \ \ldots \end{aligned} \tag{14}$$

Each summation in (13) over all G games is broken up in (14) into two summations plus leftover games. The first summation is over games g where winner w_g is team t and its opponent is loser l_g. The second summation is over games g where loser l_g is team t and its opponent is winner w_g. The rest of the games do not involve team t and are not explicitly expressed in (14). Partial derivative $\partial\lambda/\partial s_t$ follows (14).

$$\frac{\partial\lambda}{\partial s_t} = \frac{\partial\lambda_t}{\partial s_t} = \sum_{\substack{g=1\\w_g=t}}^{G} \frac{1}{p_{tl_g}}\frac{\partial p_{tl_g}}{\partial s_t} + \sum_{\substack{g=1\\l_g=t}}^{G} \frac{1}{p_{w_g t}}\frac{\partial p_{w_g t}}{\partial s_t}$$

$$= \sum_{\substack{g=1\\w_g=t}}^{G} \frac{s_{l_g}+s_t}{s_t}\frac{s_{l_g}}{\left(s_{l_g}+s_t\right)^2} + \sum_{\substack{g=1\\l_g=t}}^{G} \frac{s_{w_g}+s_t}{s_{w_g}}\frac{-s_{w_g}}{\left(s_{w_g}+s_t\right)^2}$$

$$= \frac{1}{s_t}\left(\sum_{\substack{g=1\\w_g=t}}^{G} \frac{s_{l_g}}{s_{l_g}+s_t} - \sum_{\substack{g=1\\l_g=t}}^{G} \frac{s_t}{s_{w_g}+s_t}\right) \qquad t=1,\ldots,T \tag{15}$$

Continue from (15).

$$\frac{\partial\lambda}{\partial s_t} = \frac{1}{s_t}\left(\sum_{\substack{g=1\\w_g=t}}^{G} 1 - \left(\sum_{\substack{g=1\\w_g=t}}^{G} \frac{s_t}{s_{l_g}+s_t} + \sum_{\substack{g=1\\l_g=t}}^{G} \frac{s_t}{s_{w_g}+s_t}\right)\right) \qquad t=1,\ldots,T \tag{16}$$

The first summation in (16) is the actual number of wins for team t, which is W_t in (1). The sum of the last two summations in (16) is the predicted number of wins for team t, which is $\widehat{W_t}$ in (9). Substitute actual number of wins W_t from (1) and predicted number of wins $\widehat{W_t}$ from (9) into (16) and manipulate.

$$\frac{\partial\lambda}{\partial \ln(s_t)} = s_t\frac{\partial\lambda}{\partial s_t} = W_t - \widehat{W_t} \qquad t=1,\ldots,T \tag{17}$$

An interesting result from (17) is that partial derivative $\partial\lambda/\partial\ln(s_t)$ is the difference between actual number of wins W_t in (1) and predicted number of wins $\widehat{W_t}$ in (9), which depends on strengths $s_{t'}$. With partial derivative $\partial\lambda/\partial s_t$ equal zero as required for maximization of LLF λ, (17) reduces to

$$W_t = \widehat{W_t} \qquad t=1,\ldots,T \tag{18}$$

The maximum likelihood estimates for desired strengths s_t $(t{=}1,\ldots,T)$ are strengths s_t that satisfy (18) for all T teams. Eqns. (18) in this section and (4) in Section 4 are identical. Thus, the approaches in this section and in Section 4 are in every respect equivalent. They lead to the same solution for desired strengths s_t $(t{=}1,\ldots,T)$.

The theory of maximum likelihood estimation in this section is a formal mathematical approach. The approach in Section 4 is heuristic in the sense that it is based on arguments intended to achieve compatibility with winning percentage both exactly for a balanced schedule and in concept for an unbalanced schedule. A contribution of this paper is the mathematical derivation of equivalence of these two approaches.

6 Solution

The solution is the transformation from actual winning percentages to strengths for the actual schedule as required in Section 3. Specifically, the solution is to find strengths s_t $(t{=}1,\ldots,T)$ so that (4), or identically (18), is satisfied for every team t $(t{=}1,\ldots,T)$. Actual number of wins W_t in (4) is from (1) and depends on actual game won-lost outcomes w_g, l_g $(g{=}1,\ldots,G)$. Predicted number of wins $\widehat{W_t}$ in (4) is from (9) and depends on strengths s_t $(t{=}1,\ldots,T)$. Ranking of teams t is determined by strengths s_t from this solution.

The data for the solution of strengths s_t in (4) and (9) are actual number of wins W_t from (1) for all T teams. However, the total number of wins over all T teams is constrained to equal total number of games G. Thus, the number of independent data points is $T-1$. All game won-lost outcome probabilities $p_{tt'}$ in (8), and, therefore, all predicted numbers of wins $\widehat{W_t}$ in (9) depend only on the relative and not the absolute values of strengths s_t. That is, if all strengths s_t are multiplied by the same constant, the same $p_{tt'}$'s and $\widehat{W_t}$'s would result. This same property is seen for LF Λ in (12) and LLF λ in (13) in that both are indifferent to a scale factor in strengths s_t. Thus, the number of free parameters is also $T-1$. Because of these conditions, a scaling requirement of $\Sigma_{t=1}^{T}\log(s_t){=}0$ is imposed so that a unique solution will result. Any base logarithm will suffice.

A closed form analytic solution for strengths s_t is impossible. Therefore, a numerical solution is required. Substitute predicted number of wins $\widehat{W_t}$ from (9) into (4).

$$\sum_{\substack{g=1\\ w_g=t,l_g=t'\\ w_g=t',l_g=t}}^{G} \frac{s_t}{s_t+s_{t'}} = \sum_{\substack{g=1\\ w_g=t}}^{G} \frac{s_t}{s_t+s_{l_g}} + \sum_{\substack{g=1\\ l_g=t}}^{G} \frac{s_t}{s_t+s_{w_g}} = W_t \qquad t=1,\ldots,T \tag{19}$$

Solve for strength s_t in the numerators in (19).

$$s_t = \frac{W_t}{\displaystyle\sum_{\substack{g=1\\ w_g=t,l_g=t'\\ w_g=t',l_g=t}}^{G} \frac{1}{s_t+s_{t'}}} = \frac{W_t}{\displaystyle\sum_{\substack{g=1\\ w_g=t}}^{G} \frac{1}{s_t+s_{l_g}} + \sum_{\substack{g=1\\ l_g=t}}^{G} \frac{1}{s_t+s_{w_g}}} \qquad t=1,\ldots,T \tag{20}$$

If existing strengths s_t $(t{=}1,\ldots,T)$ on the right hand sides of (20) are correct and satisfy (4), then computation of strengths s_t from (20) using existing strengths s_t yields the same existing correct strengths s_t due to the derivation of (20) starting from (4). If existing strengths s_t do not satisfy (4), then computation of strengths s_t from (20) using existing strengths s_t yields strengths s_t "closer" to the correct strengths s_t, which satisfy (4).

Therefore, the solution for the correct strengths s_t $(t{=}1,\ldots,T)$ can be found using an iterative procedure based on (20). All strengths s_t are initialized to 1 so that all s_t's are equal and satisfy the scaling requirement of $\Sigma_{t=1}^{T}\log(s_t){=}0$. At every iteration, each strength s_t is updated from (20) for all teams t $(t{=}1,\ldots,T)$ using existing strengths s_t either from the previous iteration or from initialization. At the end of each iteration, all strengths s_t are re-scaled to satisfy the scaling requirement of $\Sigma_{t=1}^{T}\log(s_t){=}0$. The iterations continue either for a fixed number or until a test for convergence of strengths s_t is satisfied. The number of iterations should be sufficiently large so that actual number of wins W_t from (1) and predicted number of wins $\widehat{W_t}$ from (9) are "close enough" for all teams t. In the limit, the iterations converge to an exact solution so that (4) is identically satisfied for all teams t (t=1,...,T). The same solution for strengths s_t that satisfies (4) also maximizes LLF λ in (13), as derived in Section 5.

Projected winning percentage $\widehat{WP}_t$ over a hypothetical balanced schedule for team t is perhaps more intuitive and useful than strength s_t. Insight is gained by comparing projected winning percentages to actual winning percentages. Projected winning percentage $\widehat{WP}_t$ for team t is the average game won-lost outcome probability $p_{tt'}$ over all $T-1$ other teams using the solution for strengths s_t.

$$\widehat{WP}_t = \frac{1}{T-1}\sum_{\substack{t'=1\\t'\neq t}}^{T} p_{tt'} = \frac{1}{T-1}\sum_{\substack{t'=1\\t'\neq t}}^{T} \frac{s_t}{s_t+s_{t'}} \qquad t=1,\ldots,T \tag{21}$$

Projected losing percentage $\widehat{LP}_t$ over a hypothetical balanced schedule for team t is similar and equals $1-\widehat{WP}_t$. The *projected number of wins* and *projected number of losses* for team t over a hypothetical balanced schedule are the number of games that team t played times projected winning and losing percentages $\widehat{WP}_t$ and $\widehat{LP}_t$ respectively. As with the actual number of wins and losses, the sum of projected number of wins over all T teams and the sum of projected number of losses over all T teams are both constrained to equal total number of games G. Projected winning percentage $\widehat{WP}_t$ in (21) is equal to the actual winning percentage for team t with an actual balanced schedule by inspection since each is over the same schedule.

An assumption in Section 3 is that winning percentages are monotonic with strengths for a balanced schedule. Specifically, if strength s_t is greater than strength $s_{t''}$, then projected winning percentage $\widehat{WP}_t$ is greater than projected winning percentage $\widehat{WP}_{t''}$, both from (21), and vice-versa. This property is critical to the presented method and must be demonstrated to be true.

Consider the following form for projected winning percentages $\widehat{WP}_t$ and $\widehat{WP}_{t''}$ over a hypothetical balanced schedule, which is identical to (21).

$$\widehat{WP}_t = \frac{1}{T-1}\left(\sum_{t'=1}^{T} p_{tt'} - \frac{1}{2}\right) = \frac{1}{T-1}\left(\sum_{t'=1}^{T} \frac{s_t}{s_t+s_{t'}} - \frac{1}{2}\right) \tag{22}$$

$$\widehat{WP}_{t''} = \frac{1}{T-1}\left(\sum_{t'=1}^{T} p_{t''t'} - \frac{1}{2}\right) = \frac{1}{T-1}\left(\sum_{t'=1}^{T} \frac{s_{t''}}{s_{t''}+s_{t'}} - \frac{1}{2}\right) \tag{23}$$

The identity of (22) and (23) to (21) is due to the inclusion of the game of teams t and t'' playing a virtual version of themselves in the summation in (21). The summations in (22) and (23) are over all teams t' (t'=1,...,T), which allows a demonstration of the desired monotonicity property. In particular, all game won-lost outcome probabilities $p_{tt'}$ and $p_{t''t'}$ "match up" for all teams t' in the summations over all T teams in (22) and (23) respectively.

Every probability $p_{tt'}$ in (22) is greater than every probability $p_{t''t'}$ in (23) if and only if strength s_t in (22) is greater than strength $s_{t''}$ in (23). If every probability $p_{tt'}$ in (22) is greater than every probability $p_{t''t'}$ in (23), then projected winning percentage $\widehat{WP}_t$ in (22) is greater than projected winning percentage $\widehat{WP}_{t''}$ in (23). Therefore, projected winning percentage $\widehat{WP}_t$ in (22) is greater than projected winning percentage $\widehat{WP}_{t''}$ in (23) if and only if strength s_t in (22) is greater than strength $s_{t''}$ in (23). In general, this monotonicity property is true only for a balanced schedule.

7 1999 NFL Schedule Example

The presented method is applied to the 1999 NFL regular season schedule as an example to demonstrate performance. First, the computed solution from Section 6 must be shown to be the correct solution. This is the purpose of Table 1. 200 iterations are performed using the iterative procedure in Section 6 to obtain a solution for strengths s_t. The *games difference* for team t is the absolute difference between the actual number of wins for team t, which is W_t in (1), and the predicted number of wins for team t, which is $\widehat{W_t}$ in (9). The second and third columns in Table 1 are the maximum and root mean square (RMS) *games difference* over all T teams per iteration. As expected, the *games differences* decrease over iterations and in the limit are driven to zero so that (4) ($W_t = \widehat{W_t}$ for all teams t) is satisfied as required. The last column in Table 1 is LLF λ in (13). As expected, LLF λ increases over iterations and in the limit converges to a maximum to satisfy the required maximum likelihood estimation criterion for a solution in Section 5.

The final solution is in Table 2. The teams are ranked by strength s_t. The base 2 logarithm of strength is also shown as well as actual won-lost records. The logarithms sum to zero as required in Section 6. The third-to-last column in Table 2 is the projected winning percentage over a hypothetical balanced schedule, which is $\widehat{WP}_t$ in (21), and is monotonic with strengths s_t as expected. The last two columns in Table 2 are the projected number of games won and lost over a hypothetical balanced schedule. As previously discussed, strength of schedule can be quantified by the difference between actual and projected won-lost records.

There is a significant difference between the ranking in Table 2 and the ranking due to the actual won-lost records. In particular, note the descendancy of St. Louis and the ascendancy of AFC East teams. The reason is explained by Table 3. The NFC West as a whole is weak. Note the non-division records. Since St. Louis played half its games in the NFC West, it is penalized for a weak schedule. This also applies to other NFC West teams. Similarly, the AFC East as a whole is strong. Also, note the non-division records. All AFC East teams are rewarded for a strong schedule since half their games are played in their division. Other differences in ranking are also due to differences in strength of schedule.

All required properties for the presented method are demonstrated by results in this section. Results for NFL regular seasons from 2000 to 2004 are in [2]. The 1999 NFL regular season results were selected for inclusion as perhaps the clearest example of differences in ranking due to readily observable differences in strength of schedule. Application of this method to other NFL regular season schedules yield smaller differences in ranking.

Iteration	Maximum Games Difference	RMS Games Difference	Log Likelihood Function
0	6.00000000	2.94026551	–171.90050077887
1	2.86084152	1.30187194	–146.23256207134
2	2.02986439	0.81722005	–140.62808838759
3	1.55118484	0.58442339	–138.41268060466
4	1.23046327	0.44412407	–137.28057316125
5	0.99965725	0.34989026	–136.62717916255
10	0.42579888	0.13944649	–135.57454378150
15	0.20906936	0.06849857	–135.39117434943
20	0.10907164	0.03618026	–135.34699847244
25	0.06154157	0.01959876	–135.33487674185
30	0.03493526	0.01074087	–135.33135449257
35	0.01952615	0.00594647	–135.33029716301
40	0.01085598	0.00333630	–135.32997077090
45	0.00603862	0.00190701	–135.32986672350
50	0.00337273	0.00111704	–135.32983219124
60	0.00113012	0.00041885	–135.32981571830
70	0.00047268	0.00017555	–135.32981328023
80	0.00025836	0.00007900	–135.32981284204
90	0.00013246	0.00003674	–135.32981275321
100	0.00006567	0.00001731	–135.32981273412
120	0.00001538	0.00000388	–135.32981272898
140	0.00000350	0.00000087	–135.32981272872
160	0.00000079	0.00000019	–135.32981272871
180	0.00000018	0.00000004	–135.32981272871
200	0.00000004	0.00000001	–135.32981272871

Table 1: Convergence Performance

8 Issues and Future Work

The major issue for the presented and any earned method is *connectivity*. Two teams are connected if either a direct or indirect path exists between them. A direct path exists if the two teams play each other. An indirect path exists if the two teams play a common opponent. This is the 1st level of indirection (direct path is 0th level). The definition for an indirect path is extended to multiple levels of indirection. That is, the two teams play teams that play teams in common, and so on. Two teams are unconnected if no such path exists between them. Computed strengths of two unconnected teams are meaningless for purposes of relative ranking. The ideal case of connectivity is a balanced schedule. Future work would be investigation of connectivity, including quantification for any pair of teams other than level of indirection and the quantitative effect of low connectivity on reliability of ranking. This problem is discussed in [8].

The data in the presented method is restricted to only game won-lost outcomes. This is pri-

			Log 2	Actual		Projected		
Rank	Team	Strength	Strength	Won	Lost	Win Pct	Won	Lost
1	Ind	6.9927	+2.8058	13	3	0.8454	13.53	2.47
2	Jack	5.0117	+2.3253	14	2	0.7987	12.78	3.22
3	Buff	3.8538	+1.9463	11	5	0.7566	12.11	3.89
4	Tenn	3.7348	+1.9010	13	3	0.7513	12.02	3.98
5	Miam	2.5624	+1.3575	9	7	0.6830	10.93	5.07
6	TBay	2.2356	+1.1606	11	5	0.6565	10.50	5.50
7	NYJ	2.1129	+1.0792	8	8	0.6453	10.32	5.68
8	St.L	2.0762	+1.0539	13	3	0.6418	10.27	5.73
9	Wash	1.7842	+0.8353	10	6	0.6111	9.78	6.22
10	KanC	1.7660	+0.8205	9	7	0.6090	9.74	6.26
11	Minn	1.7598	+0.8154	10	6	0.6083	9.73	6.27
12	Oakl	1.6575	+0.7290	8	8	0.5960	9.54	6.46
13	Seat	1.6179	+0.6941	9	7	0.5910	9.46	6.54
14	NEng	1.5941	+0.6727	8	8	0.5879	9.41	6.59
15	Det	1.2561	+0.3290	8	8	0.5383	8.61	7.39
16	SanD	1.2465	+0.3178	8	8	0.5366	8.59	7.41
17	GBay	1.0705	+0.0983	8	8	0.5048	8.08	7.92
18	Denv	1.0331	+0.0469	6	10	0.4973	7.96	8.04
19	Dall	1.0171	+0.0244	8	8	0.4941	7.90	8.10
20	NYG	0.9479	–0.0773	7	9	0.4794	7.67	8.33
21	Chi	0.7546	–0.4063	6	10	0.4325	6.92	9.08
22	Balt	0.7478	–0.4193	8	8	0.4306	6.89	9.11
23	Ariz	0.5909	–0.7591	6	10	0.3838	6.14	9.86
24	Phil	0.5693	–0.8127	5	11	0.3766	6.02	9.98
25	Car	0.4306	–1.2157	8	8	0.3242	5.19	10.81
26	Pitt	0.3533	–1.5010	6	10	0.2894	4.63	11.37
27	Atl	0.2434	–2.0383	5	11	0.2297	3.67	12.33
28	Cinc	0.2023	–2.3053	4	12	0.2029	3.25	12.75
29	SF49	0.1591	–2.6522	4	12	0.1712	2.74	13.26
30	NOrl	0.1062	–3.2347	3	13	0.1253	2.00	14.00
31	Clev	0.0830	–3.5912	2	14	0.1016	1.63	14.37

Table 2: Final Solution

AFC	Total		Division		Non Division		NFC	Total		Division		Non Division	
East	W	L	W	L	W	L	East	W	L	W	L	W	L
Ind	13	3	5	3	8	0	Wash	10	6	5	3	5	3
Buff	11	5	6	2	5	3	Dall	8	8	5	3	3	5
Miam	9	7	3	5	6	2	NYG	7	9	3	5	4	4
NEng	8	8	2	6	6	2	Ariz	6	10	5	3	1	7
NYJ	8	8	4	4	4	4	Phil	5	11	2	6	3	5
Central							Central						
Jack	14	2	8	2	6	0	TBay	11	5	5	3	6	2
Tenn	13	3	9	1	4	2	Minn	10	6	4	4	6	2
Balt	8	8	6	4	2	4	Det	8	8	4	4	4	4
Pitt	6	10	3	7	3	3	GBay	8	8	4	4	4	4
Cinc	4	12	3	7	1	5	Chi	6	10	3	5	3	5
Clev	2	14	1	9	1	5							
West							West						
KanC	9	7	4	4	5	3	St.L	13	3	8	0	5	3
Seat	9	7	4	4	5	3	Car	8	8	4	4	4	4
Oakl	8	8	3	5	5	3	Atl	5	11	4	4	1	7
SanD	8	8	5	3	3	5	SF49	4	12	2	6	2	6
Denv	6	10	4	4	2	6	NOrl	3	13	2	6	1	7

Table 3: Division Won-Lost Records

marily for compatibility with winning percentage as discussed in Section 3 and agreement with arguments in [1], [4], [7], and [20] that other factors such as margin of victory should have no influence. The strongest counter-argument against only game won-lost outcomes is that ties should not be excluded. Future work would be extension to general game won-lost outcomes. The general outcome of game g between team t and team t' would be *degree of win* $w_{t,g}$ for team t and *degree of win* $w_{t',g}$ for team t'. Degrees of win $w_{t,g}$ and $w_{t',g}$ must satisfy two properties: 1) $0 \leq w_{t,g}, w_{t',g} \leq 1$ 2) $w_{t,g} + w_{t',g} = 1$. Actual number of wins W_t in (2) would be generalized as a summation over degree of win $w_{t,g}$ for all games g involving team t. That is, degree of win $w_{t,g}$ is a generalized version of function $f_{tt'}$ in (2) and (3). The presented method is a special case with degree of win $w_{t,g}$ equal 1 for winner w_g and 0 for loser l_g which, of course, preserves all winning percentage related properties. This is not true for the general case. For a tie, degree of win $w_{t,g}$ is 0.5 for both teams, and, thus, ties are fully represented. This extension would allow a more general *game outcome function* [10] [13] [15].

Two special cases of actual winning percentage merit comment: a winless team and an undefeated team. By inspection from (4), an actual winless team is deemed certain to lose all games, and an actual undefeated team is deemed certain to win all games. That is, game won-lost outcome probability $p_{tt'}$ in (8) is always 0 for a winless team (zero strength) and is always 1 for an undefeated team (infinite strength). These effects are a direct result of the presented method being totally data driven and the requirement from Section 3 that predicted winning percentages must equal actual winning percentages. The primary consequence is that game won-lost outcome probability $p_{tt'}$ in (8) of a game between either two winless teams or two undefeated teams is indeterminate, which prevents multiple winless or undefeated teams from being relatively ranked. The general game won-lost outcome extension above prevents these conditions with the restriction on degree of win $w_{t,g}$ of $0<w_{t,g}<1$ (i.e., limits 0,1 excluded). Another possible prevention of these conditions [7] is the imposition of an additional win and loss on all teams, which plays the role of a Bayesian prior. Neither of these preventions preserves winning percentage related properties. Note that multiple winless or undefeated teams are impossible for a balanced schedule.

From arguments in Sections 2 and 3, home and away factors per game should not have any influence in the ranking results. There are two primary reasons. First, home and away factors per game are not accounted for in winning percentage. Second, a fair schedule would include an equal number of home and away games even if the schedule is unbalanced with respect to teams played. Thus, any home or away advantage would "cancel out" over the season. However, a quantitative analysis of a home or away advantage may be desired, and/or a team's schedule may not necessarily include an equal number of home and away games. In future work to allow for this, univariate strength attribute s_t would be expanded to a bivariate strength attribute. That is, home strength attribute s_t^h and away strength attribute s_t^a would be assumed for each team t. Each game g between teams t and t' would be either: 1) a home game for team t and an away game for team t' or 2) a home game for team t' and an away game for team t. The appropriate strength for each team would be used. As an example, game won-lost outcome probability $p_{tt'}$ in (8) for the former would be

$$p_{tt'} = \frac{s_t^h}{s_t^h + s_{t'}^a} \tag{24}$$

The approach and solution in Sections 4, 5, and 6 would be re-visited to determine both home strength s_t^h and away strength s_t^a for all T teams. The ideal hypothetical balanced schedule over the season would also include an equal number of home and away games for each opponent. For an actual schedule with an equal number of home and away games, an anticipated result would be the projected winning percentage with the ideal hypothetical balanced schedule above the same as in Section 6. For an actual schedule with an unequal number of home and away games, an unfair home or away advantage or disadvantage would be accounted for by a projected winning percentage with the ideal hypothetical balanced schedule above. Analysis of home or away advantage would be computation of projected winning percentage over a hypothetical balanced schedule but with an unequal number of home and away games, such as no home or no away games.

9 Historical Precedence, Contributions, and Summary

The presented method is used in the BCS. An overview is in [20]. Eqns. (8) and (11) are given, and references [6], [8], and [9] are cited. The references contain the maximum likelihood estimation approach in Section 5 along with other material. They start with (8) as an assumed probabilistic model for a game won-loss outcome and proceed to the maximum likelihood estimation solution. The same iterative solution as in Section 6 is in [8]. A mathematical proof of existence of a solution (i.e., a unique maximum) is in [8] and [9]. A mathematical proof of convergence for the iterative solution in Section 6 is in [8]. An alternate solution is in [9]. This paper relies on those results. The computation of the solution is demonstrated numerically in Table 1 in this paper and in the corresponding tables in [2].

The contributions of this paper are as follows. First and most important, the relationship of the presented method to winning percentage is not addressed in the references above. That is the primary focus in this paper. The viewpoint here is that the relationship of any earned method to winning percentage is essential for reasons discussed in Section 2. For purposes of ranking, the presented method is exactly equivalent to winning percentage for a balanced schedule and attempts to be maximally consistent with winning percentage for an unbalanced schedule while equitably taking into account differences in strength of schedule. The former is a mathematical property that is argued to be required for an earned method. The latter is a goal for the solution for an unbalanced schedule based on arguments in Sections 2 and 3. This goal is accomplished by computation of team strengths that are independent of schedule. The approach in Section 4 is based on these arguments. A succinct summary of the presented method is that the *mathematically modeled* probability of each team winning a game over that team's schedule is forced to equal the *empirically observed* probability, or winning percentage, of that team winning a game over that team's schedule.

The second contribution is related to the first. This problem may be viewed as a purely statistical parameter estimation problem using a maximum likelihood estimation approach with (8) as an assumed probabilistic model, as in the references above and in Section 5. However, the viewpoint in this paper is that the probabilistic model in (8) cannot be simply assumed with no justification based on the discussion in Section 2. Section 4 derives and justifies (8) based on the rationale in Section 3, which acknowledges the importance of winning percentage for earned methods as described in Section 2. That is, tie-in to winning percentage must be established. The mathematical derivation in this paper of equivalence of the maximum likelihood estimation approach in the references above and in Section 5 to the heuristic approach in Section 4 based on winning percentage arguments may be regarded as additional validation of the presented method and is a contribution of this paper.

The third contribution of this paper concerns the level of mathematical detail. The references above do not have the same level of detail as this paper in the derivation of the maximum likelihood estimation approach and solution. In particular, details of the derivation of the required conditions for the maximum likelihood estimation solution (i.e., (18) in this paper) are missing. This is in addition to all details of justification of the presented method from a winning percentage compatibility perspective. [15] argues for the openness of any earned ranking method, which includes all mathematical and implementation details.

References

[1] Anderson, J. and Hester, C., "Anderson & Hester - Seattle Times College Football Computer Rankings", `http://www.andersonsports.com/football/ACF_frnk.html`

[2] Bethel, R. E., "NFL Rankings Results"

[3] Bickel, P. and Doksum, K., *Mathematical Statistics*, Prentice Hall, Englewood Cliffs, NJ, 1977

[4] Billingsley, R., "In Search Of A National Champion", `http://www.cfrc.com/searchof.htm`, July 21, 2001

[5] Bowl Championship Series Home, `http://www.bcsfootball.org/`

[6] Bradley, R. A. and Terry, M. E., "Rank Analysis of Incomplete Block Designs: I. The Method of Paired Comparisons", Biometrika, Vol. 39, 1952, pp. 324-45

[7] Colley, W. N., "Colley's Bias Free College Football Ranking: The Colley Method Explained", `http://www.colleyrankings.com/`

[8] Ford, L. R. Jr., "Solution of a Ranking Problem from Binary Comparisons", American Mathematical Monthly, Vol. 64, No. 8, part 2, Oct. 1957, p. 28-33.

[9] Keener, J. P., "The Perron-Frobenius Theorem and the Ranking of Football Teams", SIAM Review, Vol. 35, No. 1, March 1993, pp. 80-93

[10] Massey, K., "Massey Ratings Description", `http://www.mratings.com/theory/massey.htm`

[11] Massey, K., "Mathematics Invades College Sports", `http://www.mratings.com/theory/uttalk_files/outline.htm`

[12] New York Times, "New York Times BCS College Top 50", `http://www.nytimes.com/`

[13] Rothman, D., "FACT College Football Standings", `http://www.cae.wisc.edu/~dwilson/rsfc/rate/rothman.html`

[14] Sagarin, J., "Jeff Sagarin's College Football Ratings (USA Today)", `http://www.kiva.net/~jsagarin/sports/cfsend.htm`

[15] Sorensen, S. P., "An Overview of Some Methods For Ranking Sports Teams", `http://www.phys.utk.edu/sorensen/ranking/`

[16] Sorensen, S. P., "Ranking Sports Teams", `http://www.phys.utk.edu/sorensen/ranking/`

[17] Wilson, D., "American College Football – Ranking", `http://www.cae.wisc.edu/~dwilson/rsfc/rate/`

[18] Wilson, D., "Bibliography on College Football Ranking Systems",
http://www.cae.wisc.edu/∼dwilson/rsfc/rate/biblio.html

[19] Wilson, D., "Wilson Performance Ratings",
http://www.cae.wisc.edu/∼dwilson/

[20] Wolfe, P. R., "Wolfe Ratings Description",
http://www.bol.ucla.edu/∼prwolfe/cfootball/descrip.htm#three