ARCENG
INTERNATIONAL JOURNAL OF ARCHITECTURE AND ENGINEERING

ISSN
2822-6895

Article Arrival Date 10.09.2024
Article Type Research Article
Article Published Date 20.12.2024

RANKING TEAMS IN AMERICAN FOOTBALL THROUGH TOTAL ENUMERATION

MEHMET MİMAN[1]

[1]Harran Üniversitesi, Mühendislik Fakültesi, Endüstri Mühendisliği Bölümü, Şanlıurfa, Türkiye,

Ahmet Sabri ÖĞÜTLÜ[2]

[2]Harran Üniversitesi, Mühendislik Fakültesi, Endüstri Mühendisliği Bölümü, Şanlıurfa, Türkiye,

Abstract

The American National Football League (NFL) stands as a significant sporting event with profound economic and social implications. Due to the physically demanding nature of the sport and the associated risks to player health, a full round-robin play format is not feasible within the league. Therefore, achieving fair and merit-based rankings of teams is imperative. This paper proposes a total enumeration approach as a precise solution to the ranking formulation put forth by Cassady et al., despite the NP-hard nature of the quadratic assignment problem inherent in the mathematical model. Total enumeration not only furnishes an exact solution for leagues with a small number of teams but also serves as a benchmark for evaluating the efficacy of heuristic methods for larger-scale problems. The implementation of the total enumeration algorithm in VBA facilitates performance analysis, with its efficiency assessed for a league comprising nine teams through statistical modeling using the JMP software package. Future research avenues include the incorporation of pruning strategies inspired by the quadratic assignment problem literature and the exploration of heuristic methods for practical league applications. Additionally, the development of a hypothetical, probabilistic game offers an experimental framework to assess team rankings, enhancing our understanding of competitive dynamics within the NFL.

Keywords: American football league, ranking teams, total enumeration

Introduction

Historical Context of the American Football League

The American National Football League (NFL) holds a significant position within American culture, dating back to the late 19th century. Originating from sports like rugby and soccer, the NFL was formally established in 1920 as the American Professional Football Association, later rebranding as the NFL in 1922. Over the years, it has evolved into a cornerstone of professional sports, captivating audiences with its rich history, competitive spirit, and iconic teams.

Ranking Football Teams

In the competitive landscape of the NFL, determining the hierarchy among teams is a continual pursuit with practical and theoretical implications. The endeavor to rank teams accurately extends beyond academic interest, influencing fan engagement, commercial partnerships, and broadcasting rights. Reliable rankings serve as a key measure of team performance, shaping perceptions and commercial activities within the sports industry.

The Quadratic Assignment Problem

At the heart of ranking football teams lies the challenge of mathematical optimization, exemplified by the quadratic assignment problem (QAP). This problem involves allocating facilities to locations based on associated costs or distances. In the context of ranking football teams, the QAP offers a structured framework for evaluating and organizing teams according to performance metrics.

Characteristics and Methodologies in Existing Literature

Upon the foundation of the diverse array of methodological approaches explored in scholarly discussions on ranking football teams, our exploration of ranking American football teams through total enumeration is informed by prior research while striving to innovate and provide comprehensive solutions. Through meticulous analysis and methodological advancements, the aim is to elucidate the intricate dynamics of team rankings in the NFL, contributing to both scholarly discourse and practical insights in the realm of sports management. Benefit can be derived by total enumeration in ranking American football teams from prior research on ranking methodologies[1][2][3]. While objectivity is offered by statistical models like the PageRank algorithm by considering player interactions as graph edges [4][5], players are evaluated based on various performance metrics and clustered into distinct groups by multivariate methods. However, the need for diverse analysis methods and benchmarks to ensure comprehensive and accurate results is highlighted by challenges in scholarly ranking, as seen in academic paper evaluations. Predictive capabilities can be enhanced by optimizing existing algorithms, like FIFA's, through incorporation of formal modeling principles and factors like home-field advantage and goal differentials. A nuanced understanding of team rankings in the NFL can be achieved and significant contributions can be made to sports management practices by amalgamating insights from these diverse approaches.

Solution Approach

This paper adopts the quadratic assessment model proposed by Cassady et. al. (2005) [6] in order to rank football teams in the American National Football League. In their study they pointed out their modeling approach as follows. In addition to existence of competition among teams in the league, ranking them with a fair merit is worthwhile and most of the time necessary. Except full round robin competition, they indicate the inherent bias in ranking systems due to coverage of individual opinions included in ranking to some extent. Ranking system incorporate this opinion directly such as opinion polls or indirectly such as mathematical models that reflect the modeler's bias. Thus, the mathematical model proposed by them requires decision maker to quantify their biases and uses mathematical programming that applies these biases fairly and merit across all competitors in the league.

That is to say that even a total enumeration technique employed in the optimal solution is just as the optimal solution to the described mathematical model not, yet the best solution ever exists.

Mathematical Model

The mathematical model formulation that ranks *n* teams in a competitive team can be displayed in Figure 1.

Maximize $$z=\sum_{i=1}^{n}\sum_{i'=1}^{n}\sum_{j=1}^{n}\sum_{j'=1}^{n} f_{ii'}d_{jj'}x_{ij}x_{i'j'} \qquad (0)$$

subject to

$$\sum_{j=1}^{n} x_{ij}=1, \quad i=1,2,\ldots,n \qquad (1)$$

$$\sum_{i=1}^{n} x_{ij}=1, \quad j=1,2,\ldots,n \qquad (2)$$

$$x_{ij} \text{ binary}, \quad i=1,2,\ldots,n$$
$$j=1,2,\ldots,n \qquad (3)$$

Figure 1. Quadratic Assessment Model for Ranking Teams

(0): Total weighted victory points to be maximized

(1): Each team is placed to the only one ranking position

(2): Each ranking position is occupied by only one team

(3): Decision variable that indicates whether the team *i* is placed to the ranking posiştion *j* (1) or not (0)

Intuitively, the objective function works to move winners up and losers down in the rankings. Decision maker can provide his preference of ranking, i.e. bias, in the objective function too.

The first index *i* refers to the team number while the second index *j* refers to the ranking position. $f_{ii'}$ refers to the degree of victory of team *i* over team *i*' . It is designed such that If team *i* defeats team *i*_, then the degree of victory for team *i* over team *i*_ is some positive value specified by the decision maker. If team *i*_ defeats team *i*, or these two teams do not play, then the degree of victory for team *i* over team *i*_ is zero. This degree of victory of winner team can be defined in any way wanted as it is one of the way from which bias of the mathematical program comes. Cassady et. al. (2005) uses it in the special example they mentioned such that (Eq. 1):

$$f_{ii'} = 0.9576^{17-w} f_{ii'}^{4} \qquad \text{(Eq. 1)}$$

Note that *w* refers to the week during which the game was played. It seems that this is how their mathematical program tries to capture the opinion pools contribution to BCS rankings providing the latest week's wins greater victory margins.

In (Eq. 1), $f_{ii'}^{4}$ is a parameter that captures the wins of 15 points margins to have equal level of strength (Eq. 2).

$$f_{ii'}^{4} = \min\left(f_{ii'}^{3}, 15\right) \qquad \text{(Eq.2)}$$

where $f_{ii'}^{3}$ takes the factor of the place, i.e. the homeland victory is not as valuable as the win in the opponent field (Eq.3).

$$f_{ii'}^{3} = \begin{cases} 0.8 f_{ii'}^{2} & \text{if team i played and defeated team } i' \text{ at the home field of team } i \\ 1.2 f_{ii'}^{2} & \text{if team i played and defeated team } i' \text{ at the home field of team } i' \\ f_{ii'}^{2} & \text{if team i played and defeated team } i' \text{ at the home field of team } i' \end{cases} \qquad \text{(Eq.3)}$$

$f_{ii'}^{2}$ reflects provides a means to assess the less victory for the wins in overtimes (Eq. 4).

$$f_{ii'}^{2} = \begin{cases} f_{ii'}^{1} & \text{if team i played and defeated team } i' \text{ in regulation} \\ 1 & \text{if team i played and defeated team } i' \text{ in overtime} \\ 0 & \text{otherwise} \end{cases} \qquad \text{(Eq.4)}$$

The victory of game is assessed by the $f_{ii'}^{1}$ (Eq. 5):

$$f_{ii'}^{1} = \begin{cases} \text{margin of victory} & \text{if team } i \text{ played and defeated team } i' \\ 0 & \text{otherwise} \end{cases} \qquad \text{(Eq.5)}$$

In Figure1. $d_{ii'}$ is a parameter that refers to the relative distance between every pair of ranking positions. The relative distance between ranking position *j* and ranking position *j_* is positive if position *j* is better than position *j_*, and negative otherwise. This is another place in the model decision maker can provide the bias in some way. For example, the relative distance for the higher-ranking positions (1-2) is much more important than relative distance for the lower level of ranks (118-119) in the BCS rankings. They use (Eq.6) to reflect this intuitive reasoning:

$$d_{j,j'} = \begin{cases} \tau^{(i-1)/(n-1)} \left(j' - j \right)^{\lambda} & \text{if } j' > j \\ -\tau^{(i-1)/(n-1)} \left(j' - j \right)^{\lambda} & \text{if } j > j' \end{cases} \quad \text{where} \quad \begin{matrix} 0 < \tau \leq 1 \\ 0 < \lambda \leq 1 \end{matrix} \qquad \text{(Eq.6)}$$

Total Enumeration

In this study, total enumeration strategy (i.e. evaluation of all possible solutions) and its efficiency (in terms of solution time) are to be investigated as the exact solution approach to the ranking teams problem formulized as the quadratic assignment model. Although Cassady et. al. (2005) investigate the meta-genetic algorithm where the solution of genetic algorithm can be exposed to pair-wise switchability for improvements for practical size problems, total enumeration solution is important to evaluate the efficiency of heuristics solutions where the exact solution to the formulized mathematical model is to be known by sure even for small sized problems.

Pseudo code for the total enumeration technique to be investigated is displayed in Figure 2.

```
zmax=0
For i = 1 to 4
    For j = 1 to 4
        For k =1 to 4
            For h = 1 to 4
            If (i ≠ j) & (j ≠ k) & (i ≠ k) & (h ≠ i) & (h ≠ j) & (h ≠ k)
                Compute z
                If z > zmax
                    Set Ranking as i-j-k-h
                End if
            End if
            Next h
        Next k
    Next j
Next i
```

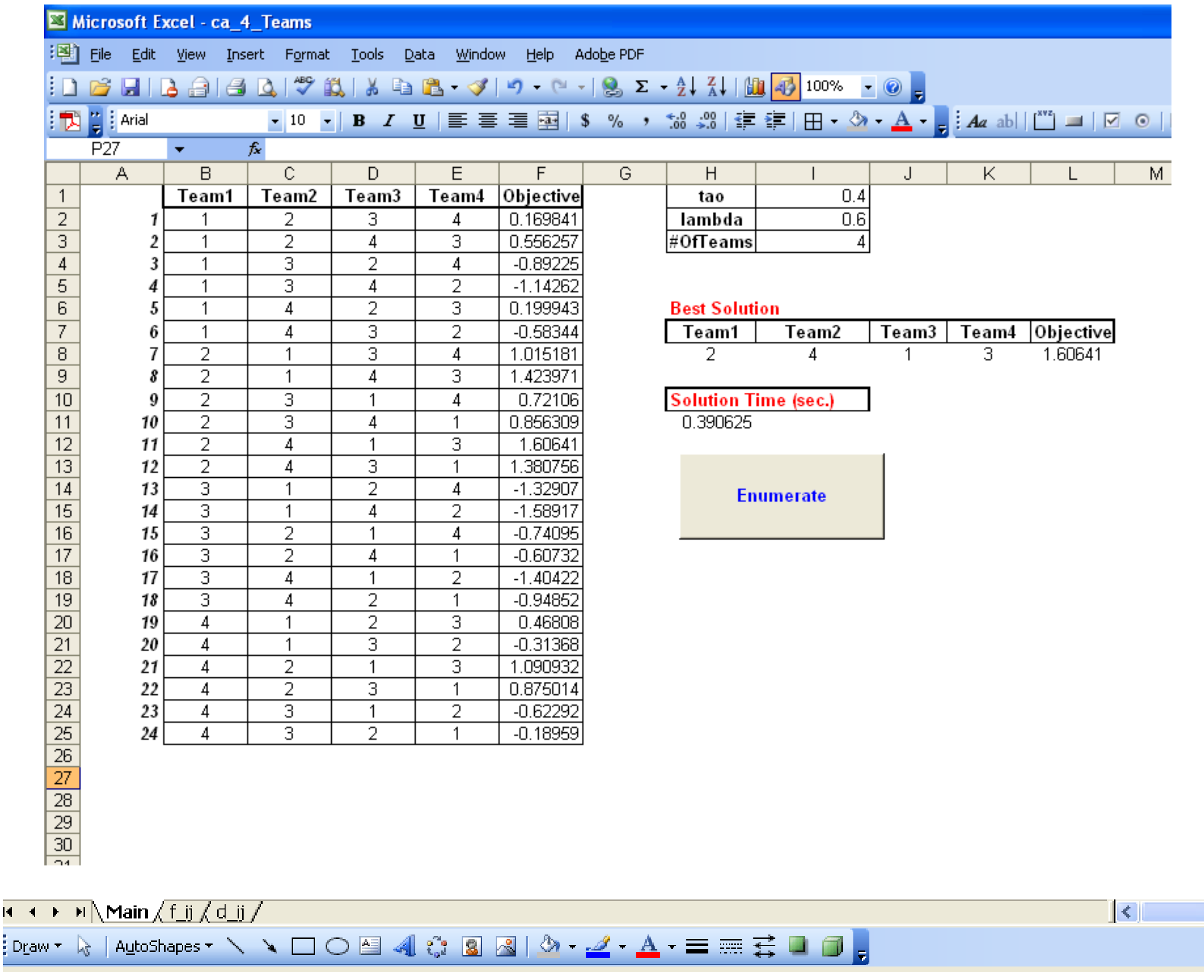
Figure 2. Pseudo Code for the Total Enumeration

Note that the algorithm has a recursive formula, hence depends on the number of teams. Therefore, the pseudo code is displayed for the n = 4.

Illustrative Example

Since the focus of this paper is the exact solution, i.e. total enumeration approach, to ranking American football teams problem and its performance in terms of solution time, i.e. the performance of the enumerations, the parameters of $f_{ii'}$ is selected randomly between uniform (0,1) for the purpose of illustration. (It is worthwhile to note that the decision maker can provide any kind of rules to judge it and our main consideration is the CPU time for the "specified problem".)

A VBA application is developed in Excel. For the purpose of the illustration for n = 4, all the objective values (4! = 24 possible ranking cases) are displayable in the application. The snapshot of this application is provided in Figure 3.



	Team1	Team2	Team3	Team4	Objective
1	1	2	3	4	0.169841
2	1	2	4	3	0.556257
3	1	3	2	4	-0.89225
4	1	3	4	2	-1.14262
5	1	4	2	3	0.199943
6	1	4	3	2	-0.58344
7	2	1	3	4	1.015181
8	2	1	4	3	1.423971
9	2	3	1	4	0.72106
10	2	3	4	1	0.856309
11	2	4	1	3	1.60641
12	2	4	3	1	1.380756
13	3	1	2	4	-1.32907
14	3	1	4	2	-1.58917
15	3	2	1	4	-0.74095
16	3	2	4	1	-0.60732
17	3	4	1	2	-1.40422
18	3	4	2	1	-0.94852
19	4	1	2	3	0.46808
20	4	1	3	2	-0.31368
21	4	2	1	3	1.090932
22	4	2	3	1	0.875014
23	4	3	1	2	-0.62292
24	4	3	2	1	-0.18959

tao	0.4
lambda	0.6
#OfTeams	4

Best Solution

Team1	Team2	Team3	Team4	Objective
2	4	1	3	1.60641

Solution Time (sec.)
0.390625

Enumerate

Figure 3.VBA Application Interface for the Illustrative Example

Figure 3. is the VBA Application Interface for the total enumeration for ranking (n = 4) teams. The application consists of 3 spread sheets. The Spreadsheet “f_ij” is denoted to specify the input parameters for marginal victories, $f_{ii'}$. (As of now, it is generated from Uniform Distribution (0,1). The specified values can be provided from the calculation based on the (Eq.1-5). The spreadsheet “d_ij” is denoted to the relative distance parameters based on Eq.6. One can specify the values of τ and λ in the “Main” spreadsheet. (However, still one can change the associated logic programmed, depending on their preference of the relative distances among rankings).

Solution Time

In this study, the performance of total enumeration is the particular interest area of the researchers. Hence, a statistical model of CPU time as a function of the number of teams in the leagues is to be develop in order to determine at what point, does the solution time become unmanageable.

As mentioned before, since the algorithm is recursive the VBA code is *n*-dependent and separate models are developed for the cases of n = 4, 5, 6, 7, 8, 9 and 10 to observe the CPU time requirements and fit a statistical model that can explain the variations on the response well. (Note that the other models can display only an optimal solution that maximizes the objective function z rather than displaying the all the cases as illustrated in Figure 3. In the results the team, who is ranked first is displayed under the cell of "Team1", who is ranked second displayed under the cell of "Team2" and so forth.

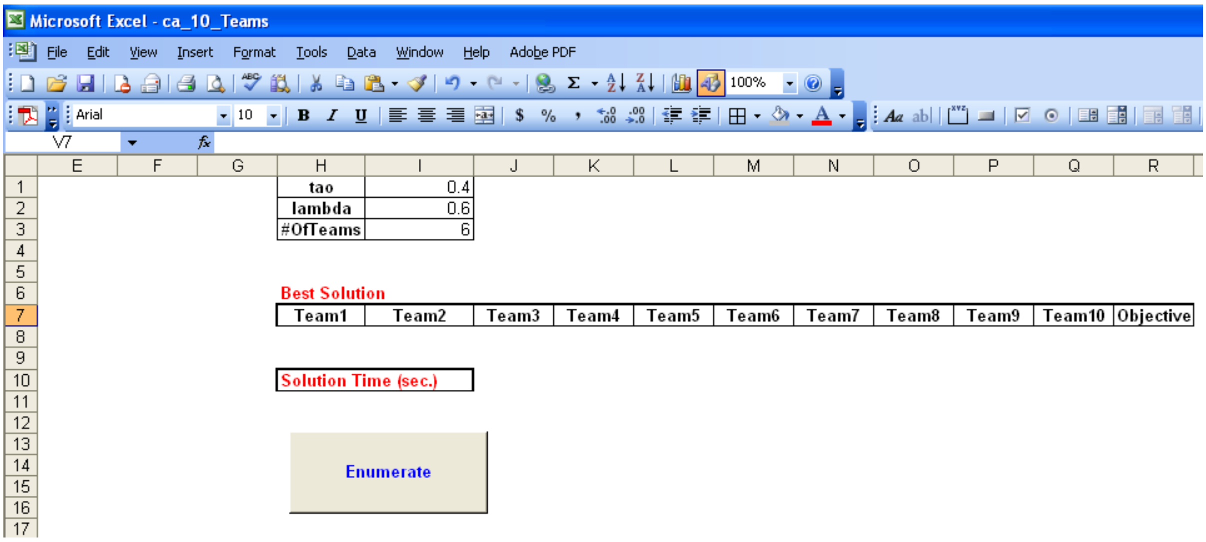


Figure 4. Snapshot from the display screen for other applications: (n = 10)

Table 1. provides the CPU times recorded for quadratic assignment problems of interest, whose details provided above, for the number of players n = 4, 5, 6, 7, 8, 9, 10 and executed in a PC of Pentium® 4 with CPU of 3.40Ghz and 1.99 GB of RAM.

Table 1. CPU Times vs. # of Teams

# of Teams (*n*)	CPU Time (sec.)
0	0.00
4	0.19
5	0.47
6	2.09

ISSN
2822-6895

7	13.63
8	188.27
9	3405.73
10	79380.45(*)

(*) The last record is obtained as the collapsed time through the start and end of experimentation.

Figure 5 provides visual illustration of CPU time versus # of Teams.

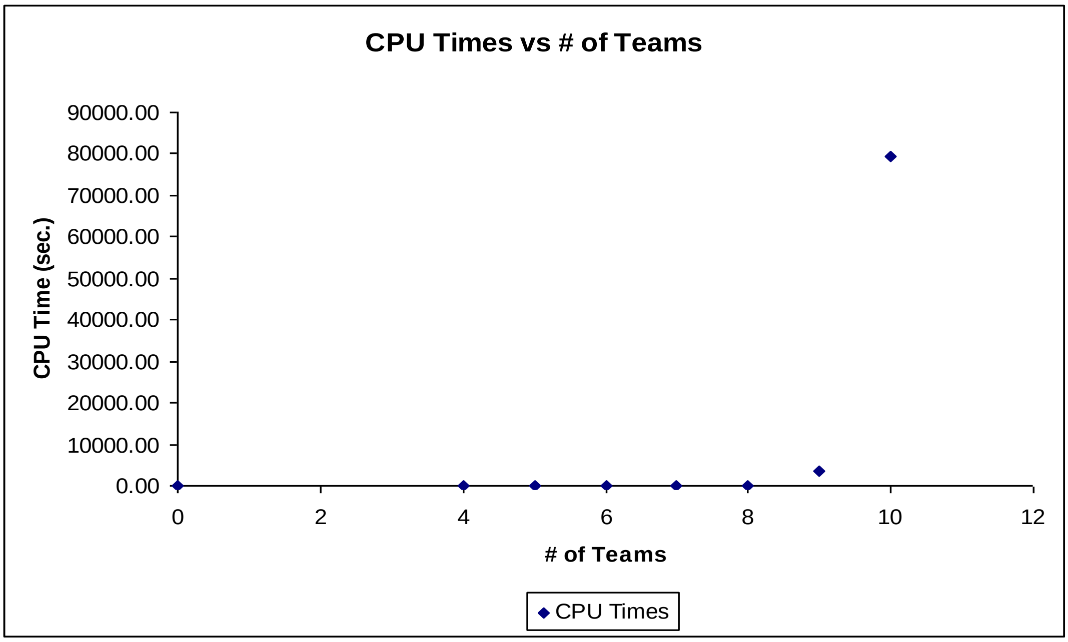


Figure 5. CPU Time vs. *n* (# of Teams)

As intuitively expected, the relationship between number of teams and required CPU time for total enumeration looks exponential (Figure 5). In fact ,an exponential model defined in Eq.7. (A variety of Non Linear Statistical models are attempted to fit data through JMP. Two of them are provided below including their statistical measures of goodness of fits (Table 2-3), such as SSE (MSE) as well as fit plots (Figure 6-7). Once can also fit an exponential distribution following the procedures of transformations described in Eq. 8, for instance,

$$Y = \beta_1 e^{\beta_2(X)} \rightarrow \ln(Y) = \ln(\beta_1) + \beta_2(X) \qquad \text{(Eq. 8)}$$

Log transferred CPU times vs. #Teams can be applied to find a fit. The best fit observed and JMP provides) is

$$Y = 8.263437 + (1.6530627)10^{-9} e^{(3.1502538576)n} \quad \text{(Eq.7)}$$

Table 2. Nonlinear Fit Model 1 by JMP

SSE	DFE	MSE	RMSE
1821.1432206	6	303.52387	17.421936

Parameter	Estimate	ApproxStdErr
theta1	1.6955746e-9	8.6389e-11
theta2	1942952561.9	3145469.72
theta3	1421284736.6	0

Solved By:

Analytic NR

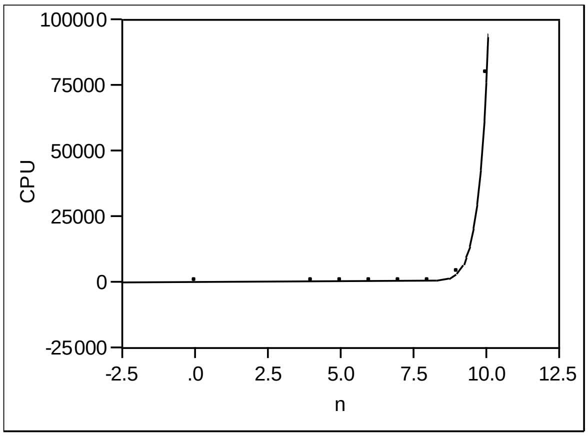


Figure 6. Nonlinear fit plot 1 by JMP

Table 3. Nonlinear Fit Model 2 by JMP

SSE	DFE	MSE	RMSE
1423.7419772	5	284.7484	16.874489

Parameter	Estimate	ApproxStdErr
theta1	8.2634371564	6.99412111
theta2	1.6530627e-9	8.918e-11
theta3	3.1502538576	0.00539213

114

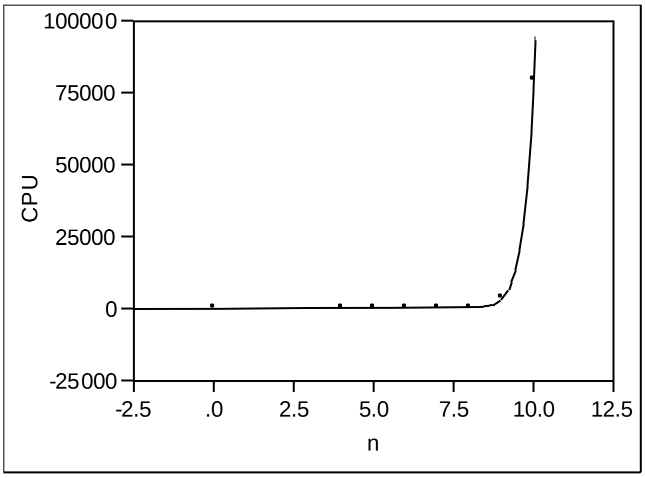


Figure 7. Nonlinear fit plot 2 by JMP

In fact, for fans, who wanted to celebrate their victory of teams, 1 hour waiting even is not good so n=9 sounds the biggest number of team, although the proper answer may depend on the how patient people are.

However, if we says 1 year = 365 days and 1 day is equal to 86400 seconds, then, the maximum number of teams manageable to ranks with total enumeration within 63 years is 13 (which takes actually approximately 33 years to solve) based on Eq. 9.

$$n = Y^{-1} = \frac{\ln\left(\frac{10^9\left(Y-8.263437\right)}{1.6530627}\right)}{3.1502538576} \qquad \text{(Eq.9)}$$

Conclusion

American National Football League is one of the most significant sportive events not only for fans but also for economy it generates directly and indirectly. Since it consists of severe actions that are potentially harmful for the health of players, a full round robin play does not exist in the league. In the absence of full round robin play, it is important to have a fair and merit ranking teams in the league from the perspective of both economic and social approach.

This paper proposes a total enumeration approach as the exact solution to the ranking formulation proposed by Cassady et. al. although the demonstrated mathematical model is a quadratic assignment problem known to be NP-hard problem in terms of solution time it requires. However, not only total enumeration provides the exact solution to the given mathematical model for a league consisting of small numbers, but does it also provide a basis for the evaluation of performance of heuristics for big problems' whose efficiency can be tested for the league consisting of small number of teams with the exact solution.

The total enumeration algorithm proposed in this paper is coded in VBA and its performance efficiency in terms of solution time is determined to be a league of 9 teams through a statistical model fit by the statistical software package JMP.

This study can be extended in several ways. First, some pruning strategies for the proposed enumerative solution approach can be incorporated using ideas from the quadratic assignment problem literature. Later heuristic methods can be investigated as solution approaches for the practical leagues.

Since the true ranking of the teams is unknown for real leagues, one can develop a hypothetical, probabilistic game for which it is possible to experimentally control the true ranking of the teams and evaluate the quality of rankings resulting from a competitive league that plays this game using full round-robin play.

References

1 Markus, Krämer. (2022). Using the PageRank Algorithm to Rank Football Players in a Game. doi: 10.33774/coe-2022-86jt3

2 A., Iyer., Shadi, Ghiasi. (2022). Using the PageRank Algorithm to Rank Football Players in a Game. Journal of Student Research, doi: 10.47611/jsrhs.v11i3.3864

3 Aniela, Pilar, Campos, de, Melo., Luiz, Henrique, Arruda, Lima., Ana, Beatriz, Dutton, da, Silva, Braúna., Hafy, Mourad, Jacoub, de, Cuba, Kouzak., Fábio, Borges, Costa., João, Felipe, Batista, da, Silva, Pinto., Fernando, de, Castro, Júnior. (2022). Football player ranking generated by multivariate methods. Research, Society and Development, doi: 10.33448/rsd-v11i11.33901

4 Yu, Zhang., Min, Wang., Morteza, Saberi., Elizabeth, Chang. (2022). Analysing academic paper ranking algorithms using test data and benchmarks: an investigation. Scientometrics, doi: 10.1007/s11192-022-04429-z

5 Alberto, Lopes. (2022). FIFA ranking: Evaluation and path forward. Journal of sports analytics, doi: 10.3233/jsa-200619

6. Cassady, C.R., Maillart, L.M., Salman, S. 2005. Ranking sports teams: a customizable quadratic assignment approach, *Interfaces*, v35, n6, Nov.-Dec. 2005, p 497-510.