SportsNetRank: Network-based Sports Team Ranking

Konstantinos Pelechrinis
University of Pittsburgh
kpele@pitt.edu

Evangelos Papalexakis
Carnegie Mellon University
epapalex@cs.cmu.edu

Christos Faloutsos
Carnegie Mellon University
christos@cs.cmu.edu

ABSTRACT

Which team is the best in the league? How does my team fare with respect to the rest of the league? These are questions that every sports fan is interested in knowing the answers to. In other cases, such as in college sports, knowing the answer to these questions is crucial for shaping the picture of specific contests. In professional sports, sports networks provide power rankings regularly - typically every week or month depending on the season length of the league - based on their experts opinion. In this work we propose an alternative, objective and network-based way of ranking sports teams. In brief, our method is based on analyzing a directed network formed between the teams of the corresponding leagues that captures their win-lose relationships. Using data from the National Football League and the National Basketball Association, we show that even simple network theory metrics (e.g., Page Rank) can provide a ranking that has the same accuracy in predicting winners of upcoming match-ups as more complicated systems (e.g., Cortana). We further explore the impact of the network structure on the prediction accuracy and we show that the cycles in the network are significantly correlated with the performance. We finally propose an advanced ranking technique based on tensor decomposition.

1. INTRODUCTION

The sports market in the US is currently worth \$67B and is projected to reach \$73B by 2019[1]. One of the fastest growing parts of the industry is the media rights associated with the various leagues. Media networks connect the vast majority of the fans with their favorite sports and teams. They do not only provide live coverage of the games but they also provide analyses that improves the understanding, experience and current affairs of the game.

One of the central questions that fans are interested in is knowing which team is the best in the league and how everyone else ranks. Currently, television or online networks provide power rankings every week (or every month depending on the league). These power rankings are highly subjective, since they are based on each expert's opinions associated with the corresponding media channel. In this work we propose an alternative, objective ranking scheme that is based on network science methods, termed `SportsNetRank`. In a nutshell, `SportsNetRank` is based on a directed network between the league teams that captures the win-lose relationships between them. Ranking the nodes in this network based on the underlying connectivity patterns provides a universal, i.e., not possible to change based on an expert's opinion, team power ranking.

While the rankings provided by our method are stable and objective, the question about their "correctness" still remains. To evaluate the rankings from our approach, we use data from the last 7 seasons of the National Football League and the last season (2014-15) of the National Basketball Association to predict the game outcomes using a simple baseline approach, i.e., if team X ranks higher than team Y, X is projected to win Y. Our results indicate that the accuracy of this approach is as good as the current state-of-the-art prediction engines that use more complicated prediction algorithms, hence, providing strong positive evidence for the quality of our method's ranking. Furthermore, for NFL `SportsNetRank` provides a 7% improvement over a baseline approach where the win percentage is used for the ranking. However, in the case of NBA the baseline performs as good as `SportsNetRank`. This can potentially be attributed to the much larger NBA season, which allows for the win percentage to converge to the "actual" ranking.

While for professional sports, the power rankings are not crucial for the league itself, in cases such as college sports it can be very crucial. For instance, in college sports the power rankings are used to choose teams to play in the national championship. While ranking systems are already in place, `SportsNetRank` introduces another simple, yet powerful, dimension for team ratings, i.e., network dynamics. We would like to emphasize here that our work is not the first that utilized PageRank as the basis for team ranking (e.g., [10, 2]). Nevertheless, the contribution of our work is twofold; **(a)** we evaluate `SportsNetRank`'s performance as a function of structural properties of the win-lose relationships of the teams. In particular, we focus on the directed cycles which are known to create problems in PageRank computations and are also the connections that create inconsistencies in the win-lose relationships [7]. **(b)** We propose an advanced ranking technique - `SportsTensoRank` - based on tensor

[1]`http://www.forbes.com/sites/darrenheitner/2015/10/19/sports-industry-to-reach-73-5-billion-by-2019/#384f24511585`

decomposition techniques that we have used to analyze composite networks [12]. While in the current work we do not thoroughly evaluate `SportsTensoRank`, we believe that this is a very promising directions, which we aim at exploring in details as part of our ongoing and future work.

2. RELATED STUDIES

Various approaches for ranking teams have appeared through the years. The simplest one - that is actually used in professional sports as the basis for advancing to playoffs - is the winning percentage, i.e., simply the fraction of games won by a team. Other ranking techniques consider the strength of the opponents and are based on the assumption that not all the wins are equal (e.g., [14]). One of the most well-known sports ranking methods is Elo ranking [9], which assigns teams with an initial number of "rating credits" and based on the outcome of the games (compared to the expected one) these credits are exchanged between the teams. While the dominant network-based approach for ranking teams are based on PageRank, Park and Newman [13] proposed an interesting approach for cases where not every team plays against every other team in the league. They introduce the idea of "indirect wins", i.e., directed paths of length k, and apply a method that resembles Katz centrality to rank the teams. A comparison of various ranking systems can be found in [3, 4], while a comprehensive list of the corresponding literature is compiled by Wilson [17].

At a tangential line of research hybrid voting-ranking systems, such as the Borda count and the Condorcet method [15], have been utilized in the sports industry. Variations of these methods are used for handing the Heisman Trophy as well as the NBA's MVP. The Associated Press is also using a similar voting/ranking scheme to obtain the college teams ranking [1]. Finally, probabilistic models for modeling the intransitivity in game data for predicting the outcome of match-up have also been recently proposed [5, 6].

3. NETWORK-BASED POWER RANKING

In this section we will introduce the network structure we use for our ranking scheme as well as our approach.

> [`League Network` definition] The league network $\mathcal{G} = (\mathcal{V}, \mathcal{E})$, is a weighted directed network where an edge $e_{i,j} \in \mathcal{E}$ points from node $i \in \mathcal{V}$ to node $j \in \mathcal{V}$ iff team j has won against team i. The edge weight $w_{e_{i,j}}$ is equal to the point differential of the corresponding game.

The league network $\mathcal{G}$ captures "who-wins-against-whom" and by what margin. With $\boldsymbol{\beta}$ being a vector whose i-th element captures external (i.e., irrelevant from the network structure) factors affecting the strength of team i (e.g., total budget, number of pro-ball players in the roster, total offense, total defense etc.), the PageRank of $\mathcal{G}$ is given by:

$$\boldsymbol{\pi} = D(D - \alpha A)^{-1}\boldsymbol{\beta} \quad (1)$$

where A is the adjacency matrix of $\mathcal{G}$, α is a parameter (a typical value of which is 0.85) and D is a diagonal matrix where $d_{ii} = \max(1, k_{i,out})$, with $k_{i,out}$ being the out-degree of node i. While in most practical cases PageRank considers only the network structure, Eq. (1) is able to take into consideration - if needed - not only the network structure but external information that affect the "power" of team i through vector $\boldsymbol{\beta}$.

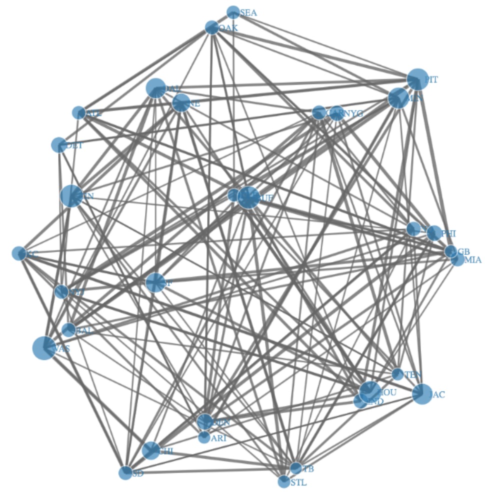

Figure 1: The directed NFL League Network for the regular season 2015-2016. More "powerful" nodes are represented with larger vertex size.

Figure 1 depicts the network obtained using all the games from the last NFL regular season. The edge width is proportional to the score differential in the corresponding game, while the vertex size is proportional to the node's PageRank score.

The rationale behind using PageRank is that compared to simple team standings that capture the fraction of wins for a team, PageRank further indirectly considers the strength of the opponent. In other words, it does not only matter to win but also to win an important team in order to improve in the rankings. In what follows we evaluate and delve into the details of the performance of `SportsNetRank`.

4. RESULTS

In this section we evaluate `SportsNetRank` through its ability to predict upcoming match-ups. In particular, let us assume that we want to predict the results of week τ. Using the games up to that week, i.e., between week 1 and week $\tau - 1$, we build the league network $\mathcal{G}_{\tau-1}$. This network will provide us with a ranking $\boldsymbol{\pi}(\tau - 1)$. The prediction algorithm simply goes over this ranking and predicts as the winner of the match-up i-VS-j team i, if $\pi_i(\tau - 1) > \pi_j(\tau - 1)$.

Table 1 depicts the accuracy results obtain on the NFL dataset. We also present the accuracy of a similar baseline algorithm where the ranking is based on the winning percentage of the team. As we can see `SportsNetRank` outperforms the baseline that uses simply the winning percentages. The overall accuracy of `SportsNetRank` is 59% with a standard error of 1.2%, while that of the winning percentages is only 52% with a standard error of 1.2%. We would like to emphasize here that a 59% accuracy is very close to the current state-of-the-art. In particular, during the last two years that Microsoft's Cortana [16] is used to predict NFL game outcomes, the corresponding accuracy is not statistically significant different compared to the `SportsNetRank`. More specifically for the season 2014-2015, Cortana exhibited a 66.4% accuracy with a standard error of 2.9% (compared with PageRank's 65.8% with a standard error of 3%), while for 2015-2016 Cortana's performance was 62.9% with standard error 3% (contrast with SportsNetRank's 59% within 3% standard error). Finally, the NFL league network is disconnected for the two first weeks of the season, and hence, SportsNetRank can be confidently used for ranking teams after week 3. The results presented in Table 1 correspond to $\boldsymbol{\beta} = \mathbf{1}$. We have performed experiments where β_i is set equal to the winning percentage of team i and the results were indistinguishable.

Year	SportsNetRank	Winning Percentage
2009	0.64	0.57
2010	0.56	0.5
2011	0.64	0.58
2012	0.65	0.56
2013	0.56	0.5
2014	0.66	0.56
2015	0.59	0.53

Table 1: Prediction accuracy with the SportsNetRank is better as compared to that of the simply winning percentages for each of the last 7 NFL seasons.

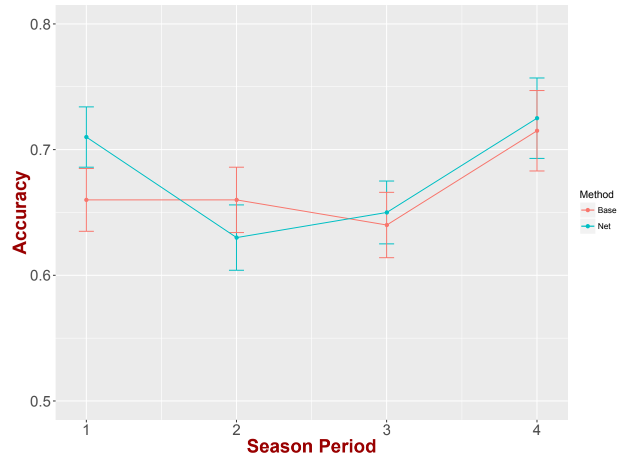

Figure 2: Even though overall SportsNetRank provides only marginal benefits over the winning percentage ranking, it is able to provide a good ranking early in the season.

For the case of the NBA league network for the season 2014-15, SportsNetRank provided an accuracy of 67%, while the baseline method came close to it with a 66% accuracy. Compared to the NFL case, NBA teams play many more games, and hence, despite a few "surprising" results, overall even the simple winning percentage provides a fairly good power ranking. In order to further explore this we compute the accuracy of SportsNetRank and the baseline method in 4 different parts of the season. In particular, we divide the season into 4 parts, each of which covers approximately a 40-day period, and we calculate the corresponding accuracy. The results are presented in Figure 2. As we can see the performance gap is much larger during the beginning of the season, when the winning percentage of the teams has not converge to the real one. Later in the season the gap is reduced and the two methods provide essentially indistinguishable performance. With respect to the absolute performance, during the last quarter of the season the prediction accuracy is the highest, possibly due to the fact that the ranking (either through SportsNetRank or through the winning percentage) has converged to one that reflects the actual power of the teams.

One of the problematic structures for PageRank are cycles in the network. In our case this corresponds to scenarios where team A has beaten team B, B has beaten team C and C has beaten team A. An acyclic network would provide a total order ranking, which in theory would have the highest accuracy. However, this is not the case in reality and cycles exists. Actually, given the fact that an NFL season consists of much fewer games than an NBA season (and not every teams plays with all the other teams), such cycles are expected to be less common. Hence, in what follows we examine the impact of the cycles on the performance of SportsNetRank.

In particular, we first calculate the minimum number of edges that one needs to remove from the network in order for it to be acyclic. This corresponds to the minimum feedback arc set problem [8]. Consequently, we compute the correlation between the fraction of edges in the minimum arc set and the difference in the performance between SportsNetRank and the baseline approach. In particular, we compute the minimum arc set for the network after every game week for NFL and after every game day for NBA. Obtaining the correlation between the minimum arc set and the difference in the accuracy of SportsNetRank and the baseline gives as **-0.19 (p-value = 0.015) for NBA** and **-0.21 (p-value = 0.039) for NFL**. The negative (and significant) correlation essentially means that SportsNetRank is expected to outperform the winning percentage ranking as the network includes fewer directed cycles.

Summarizing our results, we can say that SportsNetRank is able to capture the interactions between teams better than the simple winning percentage ranking used in determining the teams that advance to the playoffs. In particular, in the case of NFL SportsNetRank outperforms the baseline by approximately 7%. In the case of NBA, the two approaches perform similar to each other, however, SportsNetRank is able to converge to the appropriate ranking earlier in the season.

5. SportsTensoRank : EXTENDING SportsNetRank

The presence of cycles in the directed league network will always appear and cause problems, regardless of what method is used for ranking. However, it is possible to use additional information in order to break the "gridlock" in these situations. In particular, we propose to utilize tensor theory in order to account for multiple attributes simultaneously. More specifically, we have the following definition:

> [League Tensor definition] The league tensor $\underline{\mathbf{T}}(i,j,t)$, captures the temporal evolution of the win-loss relationships between the league teams. In particular, the entry $\underline{\mathbf{T}}(i,j,t) = \delta > 0$ if team j won against team i during game t by δ points differential.

One way to identify latent patterns in the data modeled by the tensor - which can then be used to rank entities in the dataset - is the PARAFAC decomposition [11]. In particular, PARAFAC decomposes $\underline{\mathbf{T}}$ to a sum of F components (see Figure 3), such that:

$$\underline{\mathbf{T}} \approx \sum_{f=1}^{F} \mathbf{a}_f \circ \mathbf{b}_f \circ \mathbf{c}_f, \qquad (2)$$

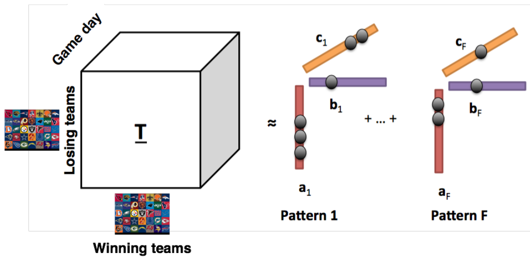

Figure 3: The PARAFAC decomposition of the league tensor can capture latent patterns in the win-lose interactions of the teams that can facilitate power ranking.

where $\mathbf{a}_f \circ \mathbf{b}_f \circ \mathbf{c}_f(i,j,k) = \mathbf{a}_f(i)\mathbf{b}_f(j)\mathbf{c}_f(k)$. In other words, each component, i.e., triplet of vectors, of the decomposition is a rank one tensor. Each vector in the triplet corresponds to one of the three modes of the tensor: **a** corresponds to the losing teams, **b** corresponds to the winning teams, and **c** corresponds to the time (game week/day). Each of these F components can be considered as a cluster, and the corresponding vector elements as soft clustering coefficients. We can then process the components that correspond to the teams (i.e., **a** and **b**) in order to obtain a final ranking of the teams.

The above decomposition is the central building block of SportsTensoRank. We can further enhance the ranking by utilizing non-network elements for the teams. In particular, every team is associated with a variety of performance indices (e.g., for the case of NFL these can include offensive/defensing yards per game, turnovers per game etc.). This information can be represented by a matrix $\boldsymbol{Y}$, where each row corresponds to a team and the columns correspond to the external attributes. Matrix $\mathbf{Y}$ is said to be coupled with league tensor $\underline{\mathbf{T}}$ at the team dimension, since they share this dimension in common. If we want to differentiate the performance of a team during the wins and the loses, we can create two matrices $\boldsymbol{Y}_w$ and $\boldsymbol{Y}_l$ that correspond to the performance metrics of the teams only during the winning and losing games respectively. Both of these matrices are coupled with the league tensor ($\boldsymbol{Y}_w$ is coupled with the winning team dimension while $\boldsymbol{Y}_l$ is coupled with the losing team dimension). In general an n-dimensional tensor can be coupled with at most n matrices $\mathbf{Y_i}$, $i \in \{1,\ 2, ...,\ n\}$. The coupled matrix-tensor factorization is given as the solution to the following optimization problem:

$$\min_{\mathbf{A_i},\mathbf{D_i}} \|\underline{\mathbf{T}} - \sum_k \mathbf{a_1}_k \circ \mathbf{a_2}_k \circ ... \circ \mathbf{a_n}_k\|_F^2 + \sum_{i=1}^{n} \|\mathbf{Y_i} - \mathbf{A_i}\mathbf{D_i}^T\|_F^2 \tag{3}$$

where $\mathbf{a_i}_k$ is the k^{th} column of $\mathbf{A_i}$. The idea behind the coupled matrix-tensor decomposition is that we decompose $\underline{\mathbf{T}}$ and $\mathbf{Y_i}$ to latent factors that are coupled in the shared dimension. This joint factorization essentially provides a low dimensional embedding of the data in a common contextual subspace that can enable a variety of tasks including ranking.

6. CONCLUSIONS

In this work we examined the performance of PageRank on ranking sports team with respect to structural characteristics of a directed network built based on the win-lose relationships of a league's team. The correctness of the ranking is inferred based on its ability to predict upcoming match-ups. We focus in particular on directed cycles and we showcase that they can significantly affect the performance of the ranking scheme. To overcome these problems we propose `SportsTensoRank` that is based on tensor theory and decomposition techniques. `SportsTensoRank` considers both the league's network as well as other performance indices simultaneously in order to provide a robust ranking. In the future, we plan on thoroughly evaluating the performance of `SportsTensoRank` in a variety of sports including NFL, NBA, NHL and MLB.

Acknowledgments: This material is based upon work supported by the NSF under Grant No. IIS-1247489 and the ARO Young Investigator Award W911NF-15-1-0599 (67192-NS-YIP). Research was sponsored by the Defense Threat Reduction Agency and was accomplished under contract No. HDTRA1-10-1-0120. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the NSF, or other funding parties. The U.S. Government is authorized to reproduce and distribute reprints for Government purposes notwithstanding any copyright notation here on.

7. REFERENCES

[1] Borda voting explained. *http://rangevoting.org/EMorg/Borda.htm*.
[2] New algorithm ranks sports teams like google's pagerank. *http://phys.org/news/2009-12-algorithm-sports-teams-google-pagerank.html*.
[3] D. Barrow, I. Drayer, P. Elliott, and G. Gaut. Ranking rankings: an empirical comparison of the predictive power of sports ranking methods. *Journal of Quantitative Analysis in Sports*, 9(2):187–202, 2013.
[4] T. Chartier, E. Kreutzer, A. Langville, and K. Pedings. Sensitivity and stability of ranking vectors. *SIAM Journal on Scientific Computing*, 33(3):1077–1102, 2011.
[5] S. Chen and T. Joachims. Modeling intransitivity in matchup and comparison data. In *Proceedings of the Ninth ACM International Conference on Web Search and Data Mining*, pages 227–236. ACM, 2016.
[6] S. Chen and T. Joachims. Predicting matchups and preferences in context. In *SIGKDD*. ACM, 2016.
[7] M. Chiang. *Networked Life: 20 Questions and Answers.* Cambridge University Press, 2012.
[8] P. Eades, X. Lin, and W. F. Smyth. A fast and effective heuristic for the feedback arc set problem. *Information Processing Letters*, 47(6):319–323, 1993.
[9] A. E. Elo. *The rating of chessplayers, past and present.* Arco Pub., 1978.
[10] A. Y. Govan and C. D. Meyer. Ranking national football league teams using google's pagerank. 2006.
[11] R. Harshman. Foundations of the parafac procedure: Models and conditions for an" explanatory" multimodal factor analysis. 1970.
[12] E. E. Papalexakis, K. Pelechrinis, and C. Faloutsos. Location based social network analysis using tensors and signal processing tools. *IEEE CAMSAP*, 2015.
[13] J. Park and M. E. Newman. A network-based ranking system for us college football. *Journal of Statistical Mechanics: Theory and Experiment*, 2005(10):P10014, 2005.
[14] D. Pickle and B. Howard. Computer to aid in basketball championship selection. In *NCAA News (page 4)*, 1981.
[15] S. Stahl and P. Johnson. *Understanding Modern Mathematics.* Jones & Bartlett Learning, 2006.
[16] N. Tower. Bing predicts falls short in nfl week 17; wild card weekend preview, http://tinyurl.com/zbhpoff 2016.
[17] D. Wilson. Bibliography on college football ranking systems. *http://homepages.cae.wisc.edu/~dwilson/rsfc/rate/biblio.html*.