

# Ranking the Teams in European Football Leagues With Agony

Stefan Neumann<sup>1</sup>, Julian Ritter<sup>2</sup>, and Kailash Budhathoki<sup>3</sup>

<sup>1</sup> University of Vienna, Faculty of Computer Science, Vienna, Austria  
stefan.neumann@univie.ac.at

<sup>2</sup> Laboratoire d'informatique, École polytechnique, France  
julian.ritter@polytechnique.edu

<sup>3</sup> Max Planck Institute for Informatics and Saarland University, Saarbrücken  
kbudhath@mpi-inf.mpg.de

**Abstract.** Ranking football (soccer) teams is usually done using league tables as provided by the organizers of the league. These league tables are designed to yield a total ordering of the teams in order to assign unique ranks to the teams. Hence, the number of levels in the league table equals the number of teams. However, in sports analytics one would be interested in categorizing the teams into a small number of *substantially* different levels of *playing quality*. In this paper, our goal is to solve this issue for European football leagues. Our approach is based on a generalized version of agony which was introduced by Gupte et al. (WWW'11). Our experiments yield natural rankings of the teams in four major European football leagues into a small number of interpretable quality levels.

**Keywords:** football, leagues, ranking, hierarchy, agony

## 1 Introduction

When thinking about football (soccer) leagues, one of the first things that comes to mind is the league table. The position of a team in the league table determines the fate of managers, the happiness of fans and how much money the club can spend in the next season. Since the league table must determine which teams win the league, which qualify for European competitions and which are relegated, it contains a lot of tie breakers so that each team is assigned a unique rank.

However, if we wish to assess the performances of teams from a data analytics point of view, the league table has multiples downsides: First, we would like to evaluate the number of *substantially* different levels of playing quality in the league. The league table offers as many levels as there are teams in the league, but one would expect a much smaller number of quality levels (for example, the top clubs who compete for the trophy, the clubs fighting against relegation and the ones somewhere in between). Second, we would like to understand against which kind of opposition teams perform well. For instance, some teams might earn a lot of points by very consistently beating bad teams but at the same time losing most games against the very good teams. The league table would cover this up since the number of points awarded against good/bad opposition are exactly the same.

## 1.1 Our Contributions

In this paper, we take a step towards resolving the aforementioned issues. That is, we show how to find a small number of interpretable levels of playing quality in the top 4 major European leagues. We also point out how some of our results indicate that certain teams perform well (or badly) against very good teams.

To obtain our results, we define graphs based on the match results in the football leagues we consider. We then apply algorithms for finding hierarchies in graphs. These algorithms will provide us with a ranking of the teams into a small number of levels which are highly interpretable.

In Section 3 we briefly discuss for how many matchdays such rankings can be obtained without any *conflicting results*. For example, a conflict refers to results with team  $A$  beating team  $B$ , team  $B$  beating team  $C$  and team  $C$  beating team  $A$ . For such conflicting results it is inherently difficult to find good rankings. We will formalize this in Section 3.

We address this issue with conflicts by introducing a generalized version of the agony problem [7] in Section 4.

To be a bit more specific, we will be looking at two types of graphs: (1) Directed graphs, where the vertices are the teams of the league and where we insert an edge from team  $A$  to team  $B$  if  $A$  beats  $B$ , and (2) undirected graphs, where we insert edges if  $A$  and  $B$  played a draw. The graph algorithms will then partition the vertices into levels such that (1) most directed edges point from “good” teams (at high levels) towards “bad” teams (at lower levels), and (2) most undirected edges are between clubs on the same level.

Section 5 presents our results. We provide hierarchies of the teams in four major European football leagues in the 2017/2018 season.

## 2 Related Work

*Ranking Football Teams.* Several approaches for assessing the quality of football teams have been proposed.

Hvattum and Arntzen [11] and Leitner et al. [13] considered a modification of the ELO rating system. ELO was initially developed for chess [5] and [11,13] modified it for usage in the football domain. Some implementations of ELO ratings are also available online [2,1].

Constantinou and Fenton [4] introduced pi-ratings and they demonstrated that pi-ratings perform better than the ELO rankings of [11,13] when predicting the outcomes of future matches. Pi-ratings have also been used in state of the art methods for predicting the outcomes of football matches [3,10].

ELO and pi-ratings assign quality scores to teams and these scores can be used to assess the strength of the teams. However, unlike the results presented in this paper, these scores are continuous-valued and do not group the teams into a small number of substantially different playing levels.

*Hierarchies in Graphs.* Finding hierarchies in directed graphs is an old and fundamental problem. In this line of work, one is given a directed graph  $G = (V, E)$  and one wants to assign levels to the vertices in  $V$  such that certain properties are satisfied. A drawing of the graph respecting these levels is called a *hierarchy*.

For example, if  $G$  is a directed acyclic graph (DAG), we can draw  $G$  such that all of its edges point downwards. Then the drawing immediately provides the levels of the vertices in  $V$ . However, such an approach fails as soon as  $G$  contains a cycle. See Section 3 for more details.

One classic way to obtain hierarchies from graphs which contain cycles is to solve the *feedback arc set problem*. This problem asks to find a subgraph  $H = (V, E')$  of  $G$  such that  $H$  is a DAG and the number of *non-DAG edges*  $E \setminus E'$  is minimized. Unfortunately, this problem is NP-hard to solve [12] and known polynomial time algorithms only achieve poly-logarithmic approximations [6]. Hanauer [8,9] conducted an experimental study on algorithms for the feedback arc set problem.

Gupte et al. [7] introduced *agony* which can be viewed as a tractable version of the feedback arc set problem. Agony does not minimize the number of non-DAG edges (i.e., the edges pointing upwards) but it weights each of these non-DAG edges by the number levels it points upwards; we provide a technical definition of agony in Section 4. [7] showed that agony can be used to identify hierarchies in social networks and to find hierarchies in American football leagues; they also provided a polynomial time algorithm. Later, Tatti [15,16,17] provided faster and more general algorithms for computing agony.

However, neither the feedback arc set problem nor agony as studied in [7,15,16,17] allow for finding hierarchies in which certain vertices in the graph are supposed to be on the same level. Thus, while these algorithms can be used for sport leagues *without* draws (such as American football or basketball), they do not apply to sports like football where draws frequently occur.

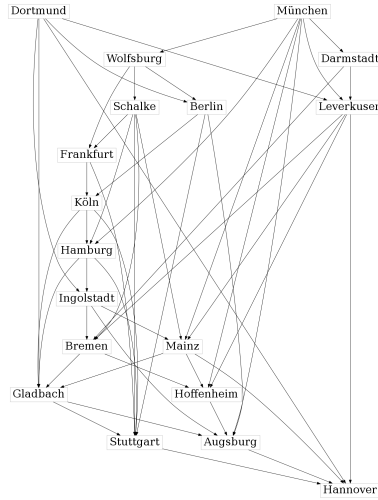
### 3 A Partial Order for a Partial Season

In this section, we discuss for how many matchdays we can rank the teams of a football league based on the results of their games *without contradicting results*.

To start out, let us first formalize what we mean by contradicting results. Consider the set of teams  $T$  and define a binary relation  $>$  on the set of teams  $T$  as follows. For teams  $A, B \in T$ , let  $A > B$  if  $A$  won against  $B$ . Now, we have *contradicting results* if there are teams  $A_1, \dots, A_k$  such that  $A_1 > A_2 > \dots > A_k > A_1$ . For example, such a contradiction occurs if Chelsea beat Liverpool, who in turn beat Manchester United, who in turn beat Chelsea.

Note that if the results contain a contradiction, we cannot rank the teams in the league properly. In the previous example, it is not clear how we should rank Chelsea, Liverpool and Manchester United because either of them could be placed on level 1.

Hence, for the rest of this section, let us only consider the first  $\ell$  matchdays of a league where the results contain no contradicting results. In such a case the result graph  $G$  is a directed acyclic graph (DAG). More formally,  $G = (T, E)$  is the graph with the teams in  $T$  as vertices and containing an edge from team  $A$  to team  $B$  if  $A > B$ .



**Fig. 1.** The first seven matchdays of the German Bundesliga season 2015/16.

If  $G$  is a DAG, then it is quite simple to obtain a ranking and a drawing for it. We can place the teams without incoming edges at level 1 (i.e., those teams which did not lose). Next, at level 2, we place all vertices which only have incoming edges from vertices already placed. Then we continue in this fashion for all levels  $i \geq 3$ . Note that all edges must be pointing downwards.

To obtain interesting rankings for a given league, we consider the largest number  $L$  such that the results of the first  $L$  match days have no contradicting results. Figure 1 presents the result graph  $G$  for the German Bundesliga season 2015/16. We picked this season for the reason that it had  $L = 7$ , yielding an elaborate diagram of the first seven matchdays. Subsequent Bundesliga seasons each had  $L = 4$ .

Unfortunately, contradicting results are quite common in football leagues or tournaments. Skinner and Freeman [14, Section 3] study this problem for the FIFA World Cup results from 1938–2006 and in case of random game outcomes.

## 4 Agony for Sport Leagues with Draws

In this section, we show how we can obtain meaningful hierarchies even in the presence of contradicting results as per Section 3. To do this, we will consider a generalized version of agony which we will be using throughout the rest of the paper. Agony was first studied by Gupte et al. [7] (see Section 2).

**Definition.** Let  $G = (V, E_{\text{dir}}, w)$  be a directed graph with edge weights  $w(u, v)$  and let  $H = (V, E_{\text{undir}})$  be an *undirected* unweighted graph. Both graphs are defined over the same set of vertices. The ranking that we will compute is related to  $G$  and  $H$  in the following way. Edges from  $u$  to  $v$  in the directed graph  $G$  indicate that  $u$  should be above  $v$  (this corresponds to team  $u$  winning against team  $v$ ). The edges in the undirected graph

$H$ , on the other hand, indicate that  $u$  and  $v$  should be on the same level in the hierarchy (this corresponds to teams  $u$  and  $v$  drawing). We will now make this more formal.

A *ranking*  $r$  is a function  $r : V \rightarrow \mathbb{N}$  which maps a vertex  $v$  to a level  $r(v)$ . In the following, when we draw the vertices in  $V$  according to some ranking function  $r$ , we will start with vertices in level 1 at the top, followed by the vertices on level 2 below that, and so on. Thus, an edge from  $u$  at level 5 to a node  $v$  on level 2 is pointing *upwards*; vice versa, an edge from  $u$  at level 1 to  $v$  at level 9 is pointing *downwards*.

Now consider a fixed ranking function  $r$ . We will define two different quantities: First, the agony caused by  $r$  on  $G$  and, second, the agony of  $r$  on  $H$ . Intuitively, edges pointing upwards w.r.t.  $r$  will cause agony and we will want to find a ranking function which has few of such edges.

The *agony of  $r$  on  $G$*  is defined as

$$A_{\text{dir}}(G, r) = \sum_{(u,v) \in E_{\text{dir}}} \max\{r(u) - r(v) + 1, 0\} \cdot w(u, v). \quad (1)$$

To develop a better understanding of the sum in Equation 1, consider an edge directed from  $u$  to  $v$  and suppose that all weights  $w(u, v)$  are 1. Now, if  $u$  has a higher rank than  $v$  (e.g.,  $r(u) = 5$  and  $r(v) = 1$ ), then the edge is pointing upwards in the hierarchy. Intuitively, this should introduce some agony and, indeed, we have  $r(u) - r(v) + 1 > 0$ . On the other hand, if  $u$  has a smaller rank than  $v$  (e.g.,  $r(u) = 1$ ,  $r(v) = 2$ ), the quantity  $r(u) - r(v) + 1$  is non-positive and no agony will be caused by this edge. This aligns well with our intuition since the edge is pointing downwards in the drawing of the vertices. Finally, if two vertices are on the same level, their agony will be exactly 1 due to the  $+1$  term. Note that in Equation 1 we also multiply this level difference with the corresponding edge weight.

We note that  $A_{\text{dir}}(G, r)$  is exactly the definition of agony used in [7] and generalized to weighted graphs.

The *agony of  $r$  on  $H$*  is given by

$$A_{\text{undir}}(H, r) = \sum_{(u,v) \in E_{\text{undir}}} |r(u) - r(v)|. \quad (2)$$

Note that  $A_{\text{undir}}(G, r)$  does not introduce any agony for edges  $(u, v)$  which have both endpoints on the same level (because then  $|r(u) - r(v)| = 0$ ). However, for edges  $(u, v)$  with endpoints on different levels, it will count the level difference.

Now our goal is to find a ranking function  $r^*$  which minimizes the sum of the agony on both graphs, i.e., we look for the minimizer of

$$A(G, H) = \min_{r \in V^{\mathbb{N}}} A_{\text{dir}}(G, r) + A_{\text{undir}}(H, r). \quad (3)$$

We note that when setting the undirected graph  $H$  to an empty graph (i.e.,  $E_{\text{undir}} = \emptyset$ ), then Equation 3 becomes exactly the same as in [7].

**Integer Linear Program (ILP).** Based on the definitions of agony, we now state the ILP that we solve to compute the agony  $A(G, H)$ . Solving the ILP also provides the ranking function  $r^*$  which is minimizing the agony.

The ILP is stated in Equation 4 and it is using the following variables:  $r(u)$  is the level assigned to vertex  $u$ ;  $x(u, v)$  is the agony caused by the directed edge  $(u, v) \in E_{\text{dir}}$ ;  $y(u, v)$  corresponds to the agony caused by the undirected edge  $(u, v) \in E_{\text{undir}}$ .

$$\begin{aligned}
& \text{minimize} && \sum_{(u,v) \in E_{\text{dir}}} x(u, v) + \sum_{(u,v) \in E_{\text{undir}}} y(u, v) && (4) \\
& \text{subject to} && x(u, v) \geq (r(u) - r(v) + 1) \cdot w(u, v) && \text{for all } (u, v) \in E_{\text{dir}} \\
& && x(u, v) \geq 0 && \text{for all } (u, v) \in E_{\text{dir}} \\
& && y(u, v) \geq r(u) - r(v) && \text{for all } (u, v) \in E_{\text{undir}} \\
& && y(u, v) \geq r(v) - r(u) && \text{for all } (u, v) \in E_{\text{undir}} \\
& && r(u) \geq 1 && \text{for all } u \in V \\
& && r(u), x(u, v), y(u, v) \in \mathbb{Z}
\end{aligned}$$

We now show that the ILP in Equation 4 correctly computes the agony.

**Lemma 1.** *The ILP stated in Equation 4 correctly computes the agony as defined in Equation 3.*

*Proof (Sketch).* Let  $r(u)$ ,  $x(u, v)$ ,  $y(u, v)$  be a solution for the ILP.

First, we show that  $x(u, v) = \max\{r(u) - r(v) + 1, 0\} \cdot w(u, v)$  for each  $(u, v) \in E_{\text{dir}}$ . Indeed, the constraints imply that  $x(u, v) \geq 0$  and  $x(u, v) \geq (r(u) - r(v) + 1) \cdot w(u, v)$ . Since  $x(u, v)$  is not present in any other constraint and since we are minimizing the objective function, we obtain the desired equality. (The same constraints were also derived in [7].)

Second, we show that  $y(u, v) = |r(u) - r(v)|$  for each edge  $(u, v) \in E_{\text{undir}}$ . Note that either both  $r(u) - r(v)$  and  $r(v) - r(u)$  are zero or exactly one of them is positive. This implies that  $y(u, v)$  is non-negative and also that  $y(u, v) \geq |r(u) - r(v)|$ . As before, the equality follows since  $y(u, v)$  does not occur in any other constraint and because we are minimizing in the objective function.

It follows that the ILP and agony compute exactly the same.  $\square$

#### 4.1 Agony for Football Leagues

Next we discuss how an algorithm for the agony problem from the previous subsection can be used to create the ranking for a football league.

Based on the results of the football league, we create two graphs  $G$  and  $H$  and each of these graphs has one vertex for each team. The graph  $G$  is directed with edge weights  $w(u, v)$  and the graph  $H$  is undirected and unweighted.

First, assume that all teams play each other exactly once. Suppose team  $A$  beats team  $B$  with a goal difference of  $z$ . For example, Borussia Dortmund beats Schalke with 5:0. In this case, we insert an edge from team  $A$  to team  $B$  with weight  $z = 5$  into the directed graph  $G$ . If team  $A$  and team  $B$  play draw, we insert an undirected edge into the undirected graph  $H$ . Thus,  $G$  will correspond to games with a winner and  $H$  will correspond to games that ended in draws.

Now, to handle multiple games between the same teams, we just add up the goal differences of all of their games. For example, if Borussia Dortmund and Schalke play three games ending 5:1, 0:2, and 4:4, then the goal difference on the edge will be 2. I.e., we will consider the “direct comparison” of the teams.

## 5 Evaluation

We implemented our algorithm in Python, and used the PuLP library for solving the ILP. The ILP can be solved in a few seconds because the problem instances are small.

To obtain better and more robust rankings, we set the weights in the directed graph  $G$  slightly differently than discussed above. If the goal difference of teams  $u$  and  $v$  is  $d$ , then we set the weight of the edge  $(u, v)$  to  $w(u, v) = 1 + \log(d)$ , where the logarithm has base 2.

There were three reasons for this choice of  $w(u, v)$ : (1) When we ran the experiments with weight function  $w(u, v) = d$ , the ranking was not very robust. For example, if a team lost only a *single* game with a relatively large goal difference this often led to the team losing one level in the hierarchy. Hence, a single bad performance could affect the ranking of a team excessively much. (2) The choice of  $w(u, v) = 1 + \log(d)$  has several nice properties. For narrow wins/losses with goal difference  $d = 1, 2$ ,  $w(u, v) = d$ . Thus, for narrow games the choice of  $w(u, v)$  has no influence on the ranking. For higher wins/losses, however,  $d \geq 3$  and  $w(u, v) < d$ ; thus, larger losses are penalized less and the rankings become more robust than discussed in point (1). (3) We also tried out different weight functions such as  $w(u, v) = d^\beta$  for different parameters  $\beta \in (0, 5]$  but the function we used gave the most natural rankings.

### 5.1 Case Study on European Football Leagues

We consider four major European football leagues, namely Premier League (England), La Liga (Spain), Serie A (Italy), and Bundesliga (Germany) for evaluating the ILP. To this end, we used the data from [www.football-data.co.uk](http://www.football-data.co.uk). The dataset for a football league contains statistics such as match stats, full time results, and total goals from all the games played in that league.

The results of the ILP on all four leagues are shown in Figures 2 and 3. For comparison, we also provide the league tables for the season 2017/2018 alongside the resulting hierarchy produced by the ILP.

**Bundesliga.** In Bundesliga, the ILP discovers four levels. Bayern Munich is the only team in the top level (LEVEL 1). This is also corroborated by the massive points difference (Pts) as well as the goal difference (GD) between Bayern Munich and the second team in Bundesliga. The teams from the second till the sixth position in the league table are in the second level (LEVEL 2). Given that they have similar statistics (the number of wins, draws, losses, and the goal difference), the result of ILP is in agreement with our intuition. LEVEL 4 consists of the two bottom-last teams Koln and Hamburg, as well as Freiburg (rank 15 in the table). Freiburg is in the fourth level since they lost the direct comparison against five teams on LEVEL 3 (some of them with large margins). The remaining midfield teams are in the third level (LEVEL 3).

Rank	Teams
1	Bayern Munich
2	Dortmund, Hoffenheim, Schalke 04, Leverkusen, Leipzig
3	Mainz, Hannover, Wolfsburg, M'gladbach, Augsburg, Hertha, Stuttgart, Ein Frankfurt, Werder Bremen
4	Hamburg, Freiburg, FC Köln

Rank	Teams
1	Manchester City, Manchester United
2	Liverpool, Arsenal, Tottenham
3	Leicester, Newcastle, West Ham, Burnley, Bournemouth, Chelsea, Crystal Palace, Watford, Everton
4	Southampton, Stoke, Huddersfield, West Brom, Swansea, Brighton

Team	P	W	D	L	GD	Pts
Bay. Munich	34	27	3	4	64	84
Schalke 04	34	18	9	7	16	63
Hoffenheim	34	15	10	9	18	55
Dortmund	34	15	10	9	17	55
Leverkusen	34	15	10	9	14	55
Leipzig	34	15	8	11	4	53
Stuttgart	34	15	6	13	0	51
Ein Frankfurt	34	14	7	13	0	49
M'gladbach	34	13	8	13	-5	47
Hertha	34	10	13	11	-3	43
Werd Bremen	34	10	12	12	-3	42
Augsburg	34	10	11	13	-3	41
Hannover	34	10	9	15	-10	39
Mainz	34	9	9	16	-14	36
Freiburg	34	8	12	14	-24	36
Wolfsburg	34	6	15	13	-12	33
Hamburg	34	8	7	19	-24	31
FC Köln	34	5	7	22	-35	32

Team	P	W	D	L	GD	Pts
Man City	38	32	4	2	79	100
Man United	38	25	6	7	40	81
Tottenham	38	23	8	7	38	77
Liverpool	38	21	12	5	46	75
Chelsea	38	21	7	10	24	70
Arsenal	38	19	6	13	23	63
Burnley	38	14	12	12	-3	54
Everton	38	13	10	15	-14	49
Leicester	38	12	11	15	-4	47
Newcastle	38	12	8	18	-8	44
Crystal Palace	38	11	11	16	-10	44
Bournemouth	38	11	11	16	-16	44
West Ham	38	10	12	16	-20	42
Watford	38	11	8	19	-20	41
Brighton	38	9	13	16	-20	40
Huddersfield	38	9	10	19	-30	37
Southampton	38	7	15	16	-19	36
Swansea	38	8	9	21	-28	33
Stoke	38	7	12	19	-33	33
West Brom	38	6	13	19	-25	31

**Fig. 2.** Results on (top) Bundesliga, and (bottom) English Premier League. Accompanying the results on the right hand side are the corresponding league tables for the year 2017/2018.



**English Premier League (EPL).** The ILP yields an interesting result on the EPL. Although second-place Manchester United has much less points than the premier league winner Manchester City, they are both in the first level. This may seem counter-intuitive at first. There's more to it, however. Manchester United lost the home game, but won the away game against Manchester City. In addition to that, the results of Manchester United are very similar to those of Manchester City against the teams on LEVEL 2 (4W, 1D, 1L of United compared to 5W, 0D, 1L of City).

Interestingly, even though Chelsea is fifth in the league table, it is placed in LEVEL 3. Although it might seem odd at a first glance, it makes sense at a closer look. The goal differences in direct comparison of Chelsea against the teams in the top three levels are:

- (LEVEL 1) Manchester City:  $-2$ , Manchester United:  $0$
- (LEVEL 2) Liverpool:  $+1$ , Arsenal:  $0$ , Tottenham:  $-1$
- (LEVEL 3) Leicester:  $+1$ , Newcastle:  $-1$ , West Ham:  $-1$ , Burnley:  $0$ , Bournemouth:  $-2$ , Crystal Palace:  $0$ , Watford:  $-1$ , Everton:  $+2$

In the direct comparison, Chelsea did not win against a single team on LEVEL 1, only one team on LEVEL 2, and only two (out of eight) teams on LEVEL 3. This also reflects their inconsistency throughout last season. As for the teams in LEVEL 4, they are also in the bottom of the league table, and share similar overall statistics.

**Serie A.** As in the previous two leagues, the ILP discovers four levels in Serie A. Although Roma is above Inter in the league table, only Inter makes it to the top level. The goal differences in direct comparison of Inter against the teams in the top two levels are as follows:

- (LEVEL 1) Juventus:  $-1$ , Napoli:  $0$
- (LEVEL 2) Roma:  $+2$ , Lazio:  $+1$ , Milan:  $+1$ , Atalanta:  $+2$ , Fiorentina:  $+2$ , Torino  $-1$ , Sampdoria:  $+6$

This clearly suggests that Inter had a very good record against the teams in the top two levels (except teams from Turin), and hence is placed in the top level.

**La Liga.** Contrary to the other leagues, the ILP discovers five levels in La Liga. Barcelona won La Liga last season with only one loss, and they were 14 points ahead of second-placed Atletico Madrid. It is sensible that they are the only team in LEVEL 1. Although Atletico Madrid is above Real Madrid in the league table (with mere 3 points difference), Real Madrid is the only team on LEVEL 2. If we look at last season's La Liga matches, we find that although Real Madrid has only marginally better performance against the top three teams in the league table compared to Atletico Madrid, Real Madrid did very well against the rest of the teams. The inclusion of Getafe on LEVEL 3 is due to its better head-to-head performances facing the teams on the same level.

Rank	Teams
1	Juventus, Inter, Napoli
2	Roma, Torino, Lazio, Atalanta, Sampdoria, Milan, Fiorentina
3	Udinese, Crotone, Chievo, Spal, Bologna, Sassuolo, Genoa, Cagliari
4	Verona, Benevento

Rank	Teams
1	Barcelona
2	Real Madrid
3	Atlético Madrid, Sociedad, Villarreal, Valencia
4	Celta, Eibar, Atlético Bilbao, Girona, Sevilla, Alaves, Leganes, Espanyol, Levante, Betis
5	Malaga, La Coruna, Las Palmas

Team	P	W	D	L	GD	Pts
Juventus	38	30	5	3	62	95
Napoli	38	28	7	3	48	91
Roma	38	23	8	7	33	77
Inter	38	20	12	6	36	72
Lazio	38	21	9	8	40	72
Milan	38	18	10	10	14	64
Atalanta	38	16	12	10	18	60
Fiorentina	38	16	9	13	8	57
Torino	38	13	15	10	8	54
Sampdoria	38	16	6	16	-4	54
Sassuolo	38	11	10	17	-30	43
Genoa	38	11	8	19	-10	41
Chievo	38	10	10	18	-23	40
Udinese	38	12	4	22	-15	40
Bologna	34	11	6	21	-12	39
Cagliari	34	11	6	21	-28	39
Spal	34	8	14	16	-20	38
Crotone	34	9	8	21	-26	35
Verona	34	7	4	27	-48	25
Benevento	34	6	3	29	-51	21

Team	P	W	D	L	GD	Pts
Barcelona	38	28	9	1	70	93
At Madrid	38	23	10	5	36	79
Real Madrid	38	22	10	6	50	76
Valencia	38	22	7	9	27	73
Villarreal	38	18	7	13	7	61
Betis	38	18	6	14	-1	60
Sevilla	38	17	7	14	-9	58
Getafe	38	15	10	13	9	55
Eibar	38	14	9	15	-6	51
Girona	38	14	9	15	-9	51
Espanyol	38	12	13	13	-6	49
Celta	38	13	10	15	-1	49
Sociedad	38	14	7	17	7	49
Alaves	38	15	2	21	-10	47
Levante	38	11	13	14	-14	46
Ath Bilbao	38	10	13	15	-8	43
Leganes	38	12	7	19	-17	43
La Coruna	38	6	11	21	-38	29
Las Palmas	38	5	7	26	-50	22
Malaga	38	5	5	28	-37	20

**Fig. 3.** Results on (top) Serie A, and (bottom) La Liga. Accompanying the results on the right hand side are the corresponding league tables for the year 2017/2018.

## 6 Conclusion

We generalised hierarchies in weighted networks to sport leagues with draws. As a proof of concept, we considered four major European football leagues. The results show that the proposed algorithm yields a natural ranking of the teams in a football league into a small number of interpretable quality levels. Future work includes applying our generalised formulation of agony to other application areas, such as ranking the products of a retail store, and developing an efficient algorithms to discover hierarchies in those domains.

## Acknowledgements

Stefan Neumann gratefully acknowledges the financial support from the Doctoral Programme “Vienna Graduate School on Computational Optimization” which is funded by the Austrian Science Fund (FWF, project no. W1260-N35). Kailash Budhathoki is supported by the International Max Planck Research School for Computer Science, Germany.

## References

1. Football club elo ratings. <http://clubelo.com>, accessed: 2018-08-03
2. World football elo ratings. <https://www.eloratings.net>, accessed: 2018-08-03
3. Constantinou, A.C.: Dolores: a model that predicts football match outcomes from all over the world. *Machine Learning* (2018). <https://doi.org/10.1007/s10994-018-5703-7>
4. Constantinou, A.C., Fenton, N.E.: Determining the level of ability of football teams by dynamic ratings based on the relative discrepancies in scores between adversaries. *Journal of Quantitative Analysis in Sports* **9**(1), 37–50 (2013)
5. Elo, A.E.: *The rating of chessplayers, past and present*. Arco Pub. (1978)
6. Even, G., Naor, J., Schieber, B., Sudan, M.: Approximating minimum feedback sets and multi-cuts in directed graphs. In: *IPCO*. pp. 14–28 (1995)
7. Gupte, M., Shankar, P., Li, J., Muthukrishnan, S., Iftode, L.: Finding hierarchy in directed online social networks. In: *WWW*. pp. 557–566 (2011)
8. Hanauer, K.: *Linear Orderings of Sparse Graphs*. Ph.D. thesis, University of Passau (2018)
9. Hanauer, K.: *Linear orderings of sparse graphs supplementary* (2018), <https://sparselo.algo-rhythmics.org>
10. Hubáček, O., Šourek, G., Železný, F.: Learning to predict soccer results from relational data with gradient boosted trees. *Machine Learning* (2018). <https://doi.org/10.1007/s10994-018-5704-6>
11. Hvattum, L.M., Arntzen, H.: Using elo ratings for match result prediction in association football. *International Journal of forecasting* **26**(3), 460–470 (2010)
12. Karp, R.M.: Reducibility among combinatorial problems. In: *Symposium on the Complexity of Computer Computations*. pp. 85–103 (1972)
13. Leitner, C., Zeileis, A., Hornik, K.: Forecasting sports tournaments by ratings of (prob)abilities: A comparison for the euro 2008. *International Journal of Forecasting* **26**(3), 471–481 (2010)
14. Skinner, G.K., Freeman, G.H.: Soccer matches as experiments: how often does the ‘best’ team win? *Journal of Applied Statistics* **36**(10), 1087–1095 (2009). <https://doi.org/10.1080/02664760802715922>

15. Tatti, N.: Faster way to agony - discovering hierarchies in directed graphs. In: ECML PKDD. pp. 163–178 (2014)
16. Tatti, N.: Hierarchies in directed networks. In: ICDM. pp. 991–996 (2015)
17. Tatti, N.: Tiers for peers: a practical algorithm for discovering hierarchy in weighted networks. *Data Min. Knowl. Discov.* **31**(3), 702–738 (2017)