As we do not have much space, we first answer general questions raised by two or more reviewers and then some specific issues.

1) Problem too narrow and solutions too specific to sport social networks.

Although we focused in one domain, the problem (i.e., ranking of network nodes) and solution (i.e., time decaying edge weights coupled with PageRank) has a fairly broad context. Ranking is a fundamental problem of many networks. Our solution is based on temporal events between nodes that are characterized by some type of advantage of one node over the other. The events can be any type of dispute or conflict of interest between pairs of nodes that, when resolved, gives advantage to one of the nodes. Alternatively, the conflicts could represent win-win situations where both nodes receive (non-equal) benefits. Other examples would be law firms (nodes) and judicial disputes (edges with time-decaying weights) or scientific collaboration between paper co-authors (edges with time-decaying weights). We have included these comments in the conclusions.

2) Comparison with T-Rank

There are significant conceptual differences between our approach and T-Rank and similar methods. They work on the algorithmic level and their solutions are hard-coded. We, otherwise, work on the data engineering level, modifying certain aspects of the data representation. Thus, we are potentially much more flexible and adaptable to different domains and applications, by modifying the parameterization via alpha, or changing the type of decaying function.

We included this argument in the related work and more thorough comparisons with other methods as future work.

3) Issues related to the alpha parameter.

We in fact do not provide an out-of-the-box way to choose the alpha parameter. However, we think that may there not be a single best alpha for all cases, it depends on how much of the past we want to carry to the predictions of the ranks. This depends ultimately in the current rules of the competition. In this sense, this may be thought as a flexibility of the method, since rules may change (in fact, depending on the sport, a lot!). Also, we have seen in some cases (e.g., MMA) that the best values of alpha are in a short range, with at most three best possible values to look for.

In any case, we are working on ways to improve this aspect, which we consider an important future work.

Response to reviewer#1's comment regarding problems with the MMA data and gold standard: We do not consider that the MMA data had problems; we only mentioned peculiar aspects of this sport that guided our experiments. Regarding the gold standard, it is an aggregation of 25 rankings with credibility among the participants of the sport, so we think it is trustable.

Response to reviewer#3's regarding the decaying function: We tested with other functions, and none produced better results. The reason is that, in this type of network, recent events are the most important to determine the best athletes in a certain point in time. For other applications this may change.

Time-Aware Ranking in Sport Social Networks

Péterson S. Procópio Jr, Marcos A. Gonçalves, Alberto H. F. Laender, Thiago Salles, Daniel Figueiredo

[1] Departamento de Ciência da Computação
Universidade Federal de Minas Gerais, Brazil
{peterson, mgoncalv, laender, tsalles}@dcc.ufmg.br
[2] Programa de Engenharia de Sistemas e Computação
Universidade Federal do Rio de Janeiro, Brazil
daniel@land.ufrj.br

Abstract. Sport social networks concern different types of relationships among athletes or teams in specific sports. Such networks have recently been used to address problems related to prediction of results of matches or championships and rankings of athletes or teams. In many cases, such analyses consider a complete and static view of the network that does not take into account the temporal nature of sports events. In this article, we present a time-aware ranking method for sport social networks that explicitly considers these temporal factors. In particular, we propose modeling such networks with edge weights that decay over time, in order to represent the relative importance of past interactions. We apply the proposed method to two sports, namely Mixed Martial Arts (MMA) and tennis. Our results show that our rankings are more accurate than a baseline ranking that ignores temporal factors, when both are compared to gold standards derived from well known rankings.

Categories and Subject Descriptors: J.4 [**Computer Applications**]: Social and Behavioral Sciencies

Keywords: Complex Networks, Sport Social Networks, Temporal Factors

1. INTRODUCTION

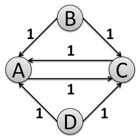
Network Science has emerged with the goal of understanding properties of dynamic and connected systems, providing several models and tools to characterize their behavior [Barabási 2009]. Facebook, the Web, protein interaction, and computer networks are all examples of networks that have been widely studied in the literature. Moreover, the last decade has evidenced a growing interest in the study of networks, partially due to the availability of large amounts of empirical data and the increase in computational power [Newman 2010].

Social Networks are among the most studied kind of network in part due to the surge of Online Social Networks and Social Media, such as Twitter, Facebook, YouTube and Google Plus. Among these, real sports social networks have also received attention recently. In such networks, nodes are athletes or teams and edges indicate some sort of interaction among them, such as direct matches. Soccer [Cotta et al. 2011; Onody and de Castro 2004], basketball [Vaz de Melo et al. 2008] and tennis [Radicchi and Perc 2011], for instance, are sport modalities that have been studied considering player-level interactions.

An important problem related to network study is vertex ranking, where the goal is to determine an ordering of network nodes to reflect some relative aspect of their importance. For example, ranking in social networks usually attempts to order individuals by their importance [Freire and Figueiredo 2011; Newman 2004]. Ranking has also been applied to sport social networks to identify the better teams

Time	**Match**	
	Winner	**Loser**
2007	A	D
2008	A	C
2009	A	B
2010	C	B
2011	C	D
2012	C	A

(a) Dataset.



(b) Network.

Fig. 1. Example of a network of matches.

or athletes in a particular sport [Radicchi and Perc 2011]. More specifically, Radicchi and Perc have considered the problem of identifying the best tennis players of all times in a recent study [Radicchi and Perc 2011]. They consider a network where nodes correspond to athletes and directed edges correspond to match results. There is an edge from athlete i to athlete j if j has defeated i in at least one official tennis match. Figure 1(b) illustrates an example of such a network. The authors apply a variation of the well-known PageRank algorithm [Page et al. 1999] to such a network created using thousands of official matches in order to obtain a ranking of the best players in History. The intuition behind this idea is analogous to that of Web page ranking. Each player has an associated "prestige" that defines her importance and flows through the network according to the direction of the edges. In such a configuration, players that defeat other players with high prestige are highly benefited.

Although good results have been obtained when compared to some official rankings, this strategy does not consider the temporal factor inherent to the events that make up the network. For instance, a match that took place recently is more important than one that happened far in the past, with respect to establishing a current ranking. Intuitively, sport social networks are dynamic with respect to athletes and matches and this time factor clearly influences official rankings. For example, an athlete that has won more matches than any other athlete in the past is not necessarily at the top of the current ranking. The importance of her victories diminishes over time when considering the ranking problem. This is also illustrated in Figures 1(a) and 1(b). Looking only at the network no clear winner emerges, but based on the matches results it is clear that in the last three years athlete C is the dominant one.

In this article we address this limitation by explicitly considering temporal aspects when constructing a sport social network. Our approach is to define a time-varying weight for the edges of the network to allow for more accurate rankings. Intuitively, an edge weight decreases with time and increases when a match is won by an athlete. This approach can be used to produce a ranking for any point in time, taking into consideration any particular period of time (e.g., the ranking of the 80s). We apply our approach to data obtained from two popular sports, Mixed Martial Arts (MMA) and tennis. Experimental results indicate that the rankings produced with our approach are superior to those provided by a classic approach (where edge weights do not decay) when compared to gold standards derived from credible known rankings. In some cases, the rankings produced by the proposed approach are very similar to the gold standard ones.

The rest of this article is organized as follows. Section 2 addresses related work. Section 3 describes the methodological aspects of our work. Section 4 details our experiments. Finally, Section 5 presents our conclusions and comments on future work.

2. RELATED WORK

Within the area of Network Science, the problem of ranking nodes in a network is of fundamental importance due to its various applications in different contexts. The problem consists in establishing an ordering or identifying a small set of nodes according to some predominant characteristic. For example, ranking individuals according to their social status within a social network; or ranking scientists according to their influence in a collaboration network [Freire and Figueiredo 2011; Newman 2004].

Due to the availability of large relational data and increasingly powerful computational resources, new challenging aspects naturally emerge, potentially harming the accuracy of current ranking strategies. One such challenge, considered here, has to do with the dynamic behavior of the interactions found in the network, usually observed when data span long periods. Centrality metrics such as degree, closeness, betweenness and PageRank are some of well known metrics for establishing an ordering of nodes in any given network. However, when facing temporal dynamics, such common network properties may change as time goes and various works have considered the time dimension when representing the network [Barabási et al. 2002; Huang et al. 2008; Kudelka et al. 2011; Leskovec et al. 2008; Sharan and Neville 2007]. In fact, considering the entire data without regarding the temporal dimension may indeed lead to misleading conclusions about some characteristics of the network. Recently, in [Mourão et al. 2009] the authors proposed a methodology to quantitatively characterize time-varying relational data. They show that neglecting the temporal dimension do indeed negatively impact the accuracy of prediction strategies.

This article focuses on a time-aware ranking method for sports social networks, which may also be applied to other networks that are subject to temporal factors, such as, for instance, collaboration networks. The idea of having time varying weights on network edges to capture the fact that intensity of relationships can increase and decrease over time has appeared in the literature [Kudelka et al. 2011; Sharan and Neville 2007]. However, these approaches have been applied to characterizing different types of social ties [Kudelka et al. 2011] or applied to the problem of topic classification [Sharan and Neville 2007].

In [Berberich et al. 2004], the authors propose T-Rank, a method that extends PageRank to improve node ranking by considering the activity and freshness of nodes to compute their importance, given a temporal window of interest. The quality of the method was assessed by comparing rankings produced by T-Rank with those produced by PageRank.

We focus explicitly on the *network ranking problem* and propose an edge weight function specific for this purpose. Our proposed method does not restrict the importance of events to a given temporal interval of interest, but instead explicitly changes the importance of events according to time. We also focus on sport social networks and, unlike previously mentioned papers, validate our method by quantitatively comparing its produced rankings to credible gold standards.

Moreover, there are significant conceptual differences between our approach and T-Rank and similar methods. They work on the algorithmic level and their solutions are hard-coded in the algorithm. We, otherwise, work on the data engineering level, modifying certain aspects of the data representation, which means that we are potentially much more flexible and adaptable to different domains and applications, by modifying for instance, the edge weight function.

3. METHODOLOGY

3.1. The Sports

We have chosen Mixed Martial Arts (MMA) and tennis as the sports of interest for testing our proposed time-aware ranking strategy for a number of reasons. First, they are both high popular sports. Although tennis is a more traditional sport, with years of history, MMA in its turn has gained a lot of popularity in the last few years. Second, and more importantly, there are several sites available on the Web with information about events, confronts, etc., as well as with rankings of athletes that can serve as a possible gold standard, thus making our work feasible.

Next, we describe some design choices we made regarding each sport.

3.1.1. MMA

Mixed Martial Arts (MMA) is a full contact combat sport that allows several types of fighting techniques. Confronts usually take place in events in which fighters try to defeat each other. Although interesting for our study for the reasons mentioned above, there are also some challenges in exploring this particular sport. First, there are several organizations related to the sport that promote events all around the world, with slightly different rules. The current main organization is UFC, although there are several other organizations such as Pride Fighting Championships (now extinct), DREAM, and WEC. The second challenge relates to the existence of several fighter categories, based on the body weight. Examples include Welterweight and Heavyweight. An important aspect is that each organization adopts its own rules to define the categories. For instance, UFC and Strikeforce consider that the Welterweight category covers body weights between 71–77 kg while in Pride FC the same category covers fighters who weight up to 83 kg. Besides that, another important aspect is that a fighter may compete in different categories in different points in time. For instance, in a given moment a fighter may belong to category Welterweight while in a different moment in his carrier he fighted under the Middleweight category.

For these reasons we made the following design choices:

—We have not considered existing differences among organizations and events. We considered all confronts from all events equivalent, since the fighters were competing under the same conditions.

—The analysis of results are perfomed per category, i.e., we consider all fights within a given category to produce a rank.

3.1.2. Tennis

Tennis is a largely popular traditional sport, dating back to the 19^{th} century. Matches are usually held between two players or double teams and, unlike MMA, they occur much more frequently in a per year basis—consequently, there is a higher number of confronts per player than in MMA. The Association of Tennis Professionals (ATP) organizes several tennis tournaments around the world. The most important ones are the four Grand Slams, namely, the Australian Open, the French Open, Wimbledon, and the US Open.

With respect to tennis, we made the following design choices:

—We have not considered existing differences among tournaments. We considered all confronts from all tournaments equivalent, since the players were competing under the same conditions.

—The analysis of results are performed per time span, i.e., we consider all confronts that occurred within a given period of time.

3.2. The Network Representation

The results of the confronts (fights or matches) are materialized through a network represented as a weighted directed graph. Each node corresponds to an athlete and there is an edge connecting nodes i and j if the athlete represented by node i was defeated in some confront by the athlete represented by node j. The weight of an edge represents the weight associated with one or more confronts between two athletes. This will be further discussed in Section 3.4. As an example, we can observe in Figure 2 one network constructed using our dataset with the results of confronts among some of the main MMA fighters according to the ranking presented in Session 5.2.2.

It is important to notice that we organized the data in order to generate a network of confronts according to various criteria, such as confronts in the same category(MMA), in the same tournament(tennis), that happened in the same year, etc. for future developments.

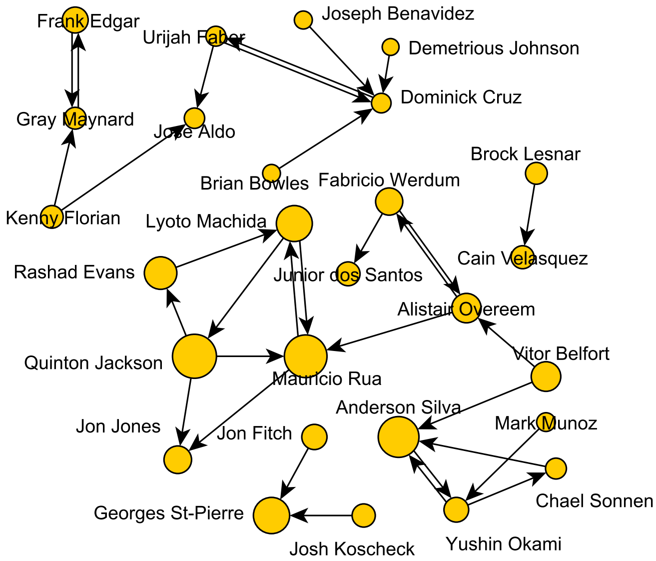


Fig. 2. Network of confronts among some of the main MMA fighters.

3.3. The Ranking Method

We want to obtain a ranking of athletes based on the results of the confronts up to a given point in time. By analyzing the network of confronts, we can, intuitively, make the conclusion that nodes with high indegree represented those with more wins, and, consequently, with higher chance of being highly ranked. This intuition, however, takes into account only the number of wins, ignoring the "quality" of these wins. For instance, using only this criteria an athlete that has defeated several other not so "prestigious" athletes would be better ranked than another athlete that has obtained fewer victories but over stronger opponents. A way to address this problem is to apply the well known PageRank algorithm [Page et al. 1999] in the network of confronts. The intuition here is analogous to that of Web page rank. With each athlete is associated a certain "prestige" that defines its importance and that flows through the network according to the direction of the edges. Thus, athletes who defeat other athletes with higher prestige are more benefited.

3.4. Time-Varying Edge Weights

The time-dependent interactions among athletes clearly has a fundamental impact on their rankings. For example, a series of recent confronts is likely to be more important in determining the current ranking of set of athletes than a series of confronts among the same set of athletes that occurred far in the past. Intuitively, the importance of confronts towards ranking of athletes decays over time: the older a confront is, the less important is its result. Thus, a good ranking for today's athletes should give more importance to recent confronts.

The importance of relationships is usually captured by assigning weights to network edges. Thus, based on this intuition, we propose time-varying edge weights to reflect the fact that the importance of confront results decays with time. In particular, we will consider an exponentially decaying weight, controlled by a parameter that determines how fast importance decays over time [1]. Consider two athletes i and j and let $t^k_{i,j}$ denote the time instant of their k-th confront, for $k = 1, 2, \ldots$. Moreover,

[1] We tested with other functions, and none produced better results.

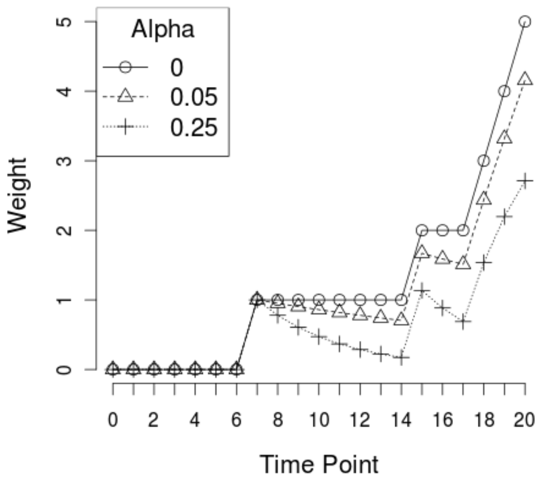


Fig. 3. Evolution of edge weight $w_{i,j}(t)$ over time (for $\alpha \in \{\mathbf{0}, \mathbf{0.05}, \mathbf{0.25}\}$).

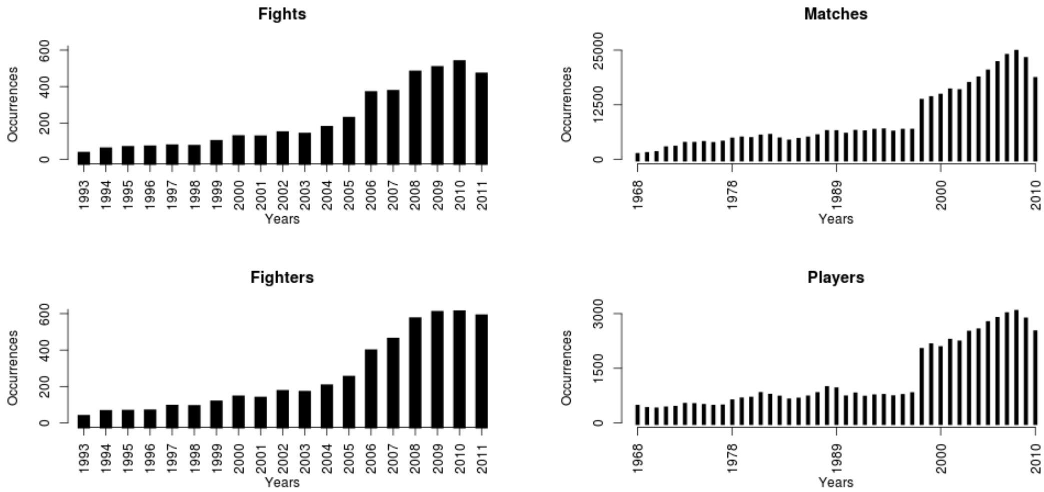
let $w_{i,j}(t)$ denote the weight of the directed edge from i to j at time $t \geq 0$. We define the edge weight $w_{i,j}(t)$ recursively as follows:

$$w_{i,j}(t) = \begin{cases} 0 & t = 0 \\ w_{i,j}(t-\epsilon) + I & t = t_{i,j}^k \text{ and } j \text{ defeated } i \\ & \text{in confront at } t_{i,j}^k \\ w_{i,j}(t_{i,j}^k)e^{-\alpha \cdot (t - t_{i,j}^k)} & t_{i,j}^k \leq t < t_{i,j}^{k+1} \end{cases} \tag{1}$$

Note that the edge weight is zero at time zero. At the time of a confront, namely at the instants $t_{i,j}^k$ for $k = 1, \ldots$, the weight of the edge increases by I if j defeats i. Note that I is a constant that indicates how much weight is added to the edge when an athlete defeats another. This value could depend on the importance of the confront (e.g., a final), but in this paper we assume $I = 1$ for all confronts. Note that ϵ is a small constant (e.g., 10^{-6}) and is used to capture the edge weight just before a confront between i and j occurs. Finally, in between confronts the weight decreases exponentially with parameter α according to the amount of time elapsed since the last confront (note that $t - t_{i,j}^k$ is the time elapsed since the k-th confront). Note that if j never again defeats i, then the weight $w_{i,j}$ will eventually approach zero.

A key parameter of the formulation above is α which denotes how fast the weight of an edge decreases over time. If α is too small, then the edge weight will have a long memory. In particular, if $\alpha = 0$ then edge weights do not decrease with time. On the other hand if α is too large, then the weights have very short memory, quickly going to zero. Intuitively, α should be set according to the timescale of the sport to reflect how fast real rankings change. Moreover, it should also be related to the number of confronts per unit of time (e.g., year or month) of the sport. In what follows, we investigate the impact of various values for α.

Figure 3 illustrates the function $w_{i,j}(t)$ for two athletes in our dataset, considering three values for α, namely, 0, 0.05 and 0.25. Note that there was a confront at time points 7, 15, 18, 19 and 20 when j defeated i, thus indicating the increments ($I = 1$). Between confronts, the weight decreases exponentially. Note that for $\alpha = 0.05$ the edge accumulates more weight over time, indicating its longer memory.



(a) MMA dataset.

(b) Tennis dataset.

Fig. 4. Number of confronts and athletes through time.

4. DATASETS

In this section we describe the datasets we used in this article.

4.1. Mixed Martial Arts

As far as we know, there is no publicly available dataset with results of MMA fights. To obtain structured data of fight results one has to look for websites that make this kind of data available and create wrappers [Laender et al. 2002] to extract it. The main and most complete site related to MMA in terms of fight results is Sherdog[2]. However, Sherdog does not make available information about the category of the fights, only the fighters' main categories. As fighters may compete in distinct categories at different points in time and we wanted to compare rankings obtained from fights of each category, we could not use this data.

Instead, we used data available in the FightMetrics site[3] which also makes available the results of the most important fights in the main MMA events around the world. Although it covers less fights than Sherdog, it has detailed information of each match, including the category of each one. Moreover, FightMetrics has also a cleaner and more structured interface which facilitates the creation of the wrapper for data extraction. Table I summarizes the data collected.

Number of fighters	1838
Number of fights	3648
Number of organizations	59

Table I. Summary of the extracted data.

Figure 4(a) shows the number of MMA fighters and fights through time. As we can see, only recently MMA has started to have a considerable number of matches per year. This shows that this

[2]http://www.sherdog.com/
[3]http://fightmetric.com/

is a new and growing popular sport.

4.2. Tennis

We collected data about tennis matches from the web site of the Association of Tennis Professionals (ATP)[4]. We collected all matches played by professional tennis players from January 1968 to October 2010, summing up to a total of 378566 matches. Table II summarizes the data collected.

Figure 4(b) shows the number of tennis players and matches through time. As we can see, tennis is already a consolidated sport, presenting a much higher number of confronts per year than MMA.

Number of players	13703
Number of matches	378566

Table II. Summary of the extracted data.

5. EXPERIMENTS

In this section we describe the experiments we conducted using the two datasets. For each experiment we present the experimental design, describe the gold standard and the evaluation metrics used and discuss the results.

Our main goal with the experiments was to verify whether our time-aware enhanced strategy was more suitable to produce rankings for specific time points than a baseline that ignores the time factor. In other words, we would like to evaluate how effective our strategy with the temporal decaying function is for this specific task.

5.1. Mixed Martial Arts Experiments

5.1.1. Experimental Design

For the MMA experiments, we built several fighting networks divided per category and varied the weighting method for the edges. Next we applied the PageRank algorithm [White and Smyth 2003] in all networks to rank the athletes. We then compared the generated rankings with a gold standard to verify which weighting strategy produced the best results. More specifically we tested two weighting strategies:

(1) Cumulative: No time-aware function is applied to the weight of the edges, which is our baseline and corresponds to the method proposed in [Radicchi and Perc 2011]: a PageRank variation that does not consider the time factor. Notice that this is the same as considering $\alpha = 0$ in our edge weighting function described in Section 3.4.

(2) Time-aware: The time-aware function is applied to the weight of the edges (for this, we varied α from 0.05 a 0.35 with a 0.05 step).

Overall, we built and compared the rankings produced by 56 different fighting networks: 7 categories $\times$ 8 variations of the edge weighting function (one cumulative and seven functions with temporal decay).

[4] http:www.atpworldtour.com

5.1.2. Gold Standard

Since our ranking method and the baseline's [Radicchi and Perc 2011], which completely ignores the temporal factors, produce ranked lists of the best athletes, we need an external list of the best fighters, ideally produced by a third-party, to serve as a gold standard with which we could compare our results. This existing rank should also be credible enough so that the results could be trusted.

Accordingly, we chose to use the USA Today / SB Nation MMA 2012 Consensus Ranking[5], a ranking of the 25 best fighters of 2011 for each category built from the aggregation of other 20 most important rankings of the MMA community, which may use different criteria. This was supposed to be a consensual ranking, therefore being a good candidate for a gold standard. We have considered the rankings published in the date of 10/27/2011, and thus have also considered only fights that happened up to this date in order to produce the rankings.

5.1.3. Evaluation Metrics

To compare the rankings generated by our method and the baseline's with the respective gold standard, we used two metrics: Recall and the Spearman's Rank Correlation Coefficient[6]. Recall captures how many of the top-25 best fighters in each category we are able to retrieve with each ranking method.

Spearman's Rank Correlation Coefficient captures how well two rankings correlate in terms of the positions each fighter appears in the respective rankings. It returns values between 1 and -1 for perfectly coincident and inverse rankings. Positive values indicate some correlation, negative values indicate negative correlation and a value of 0 means there is no correlation at all between the ranks.

One disadvantage of using the Spearman's Rank Correlation Coefficient to compare two ranks, in our case specifically, is that the gold standard considers only the top 25 positions while the ranking methods return a ranked list with all the fighters in the dataset, totalizing at least 50 fighters per category. As the set of all possible rank values is different, this may cause some anomalies in the computation of this metric. To minimize this problem, we implemented a small variation of this coefficient, in which fighters whose real position in the gold standard ranking are below the 25th, but that are not retrieved by the analyzed ranking methods in their top 25 results, are ordered according to the position they appear in the respective rankings and are then given rank positions 26, 27 and so on. This small change smooths the distortions while at the same time penalizes the method that retrieves more of those "below 25" fighters.

5.1.4. Results

Table III shows the results obtained for the 56 generated fighting networks. Each line corresponds to the results for one specific network. Columns *Weighting Class*, *Num. Compared*, α, *Recall* and *Spearman's Coef* represent, respectively, the fighting category, the number of fighters present in the gold standard in each category for which we have information in our dataset, the respective α parameter for the time decaying strategies, and Recall and Spearman's Coefficient results obtained by comparing the generated rank and the gold standard.

By analyzing the results, we can make some general observations:

—In all categories, our time-aware alternatives outperformed the baseline ($\alpha = 0$) according to the Recall and Spearman's Coefficient metrics. Recall figures for the respective best α value in each category varied between 74%-96% which can be considered an excellent result. In fact, in most

[5] http://www.bloodyelbow.com/rankings
[6] http://mathworld.wolfram.com/SpearmanRankCorrelationCoefficient.html

Weight Class	Num. Compared	α	Recall	Spearman's Coef.
Bantamweight	20	0.05	0.90	0.26
		0.10	0.90	0.28
		0.15	0.90	0.23
		0.20	0.90	0.27
		0.25	0.90	0.29
		0.30	0.90	0.35
		0.35	**0.90**	**0.36**
		0	0.90	0.19
Featherweight	19	0.05	0.68	-0.99
		0.10	**0.74**	**-0.82**
		0.15	0.74	-0.85
		0.20	0.74	-0.89
		0.25	0.68	-1.05
		0.30	0.63	-1.24
		0.35	0.63	-1.26
		0	0.68	-0.99
Heavyweight	23	0.05	0.78	0.01
		0.10	0.87	0.21
		0.15	0.87	0.27
		0.20	0.96	0.35
		0.25	**0.96**	**0.45**
		0.30	0.91	0.34
		0.35	0.91	0.32
		0	0.74	-0.25
Light Heavyweight	23	0.05	0.78	0.30
		0.10	0.83	0.46
		0.15	0.83	0.58
		0.20	**0.83**	**0.63**
		0.25	0.78	0.59
		0.30	0.78	0.53
		0.35	0.74	0.45
		0	0.78	0.26
Lightweight	24	0.05	0.75	0.24
		0.10	0.75	0.28
		0.15	0.79	0.36
		0.20	**0.79**	**0.36**
		0.25	0.75	0.23
		0.30	0.75	0.14
		0.35	0.75	0.08
		0	0.67	-0.22
Middleweight	22	0.05	0.64	-0.50
		0.10	0.68	-0.21
		0.15	0.73	0.03
		0.20	0.77	0.25
		0.25	**0.77**	**0.35**
		0.30	0.73	0.29
		0.35	0.68	0.22
		0	0.59	-0.67
Welterweight	23	0.05	0.70	-0.06
		0.10	0.74	0.07
		0.15	0.74	0.07
		0.20	0.74	0.00
		0.25	0.74	0.00
		0.30	**0.78**	**0.03**
		0.35	0.74	-0.11
		0	0.61	-0.29

Table III. Results of the 56 fighting networks. Notice that when $\alpha = 0$ the cumulative edge weighting strategy is in use. When $\alpha > 0$, the decaying strategy is used.

cases our Recall results are equal or better than the baselines', with gains of up to 30%. Notice also the Spearman's Coefficient for the best α values are positively correlated with the gold standard ranking in 5 out of 7 categories and that in some cases there are very significant correlations (e.g., for categories "Heavyweight, and "Light Heavyweight")[7].

[7]Notice that in a few cases, the Spearman Coefficient's results are lower than the inferior limit of -1. This occurs because the universe of possible ranking values are different between the generated ranks and the gold standard, although this problem has

(a) Ranking results for the Light Heavyweight category.

Fighter	**Gold Standard**	**Ours**	$\mathbf{D^2}$	**Baseline**	$\mathbf{D^2}$
Jon Jones	1	1	0	6	25
Mauricio Rua	2	8	36	10	64
Rashad Evans	3	2	1	1	4
Quinton Jackson	4	4	0	5	1
Lyoto Machida	5	3	4	2	9
Dan Henderson	6	22	**256**	20	**196**
Phil Davis	7	7	0	19	**144**
Forrest Griffin	8	6	4	4	16
Gegard Mousasi	9	14	25	21	**144**
Rafael Cavalcante	10	9	1	22	**144**

(b) Ranking results for the Heavyweight category.

Fighter	**Gold Standard**	**Ours**	$\mathbf{D^2}$	**Baseline**	$\mathbf{D^2}$
Cain Velasquez	1	5	16	12	**121**
Junior dos Santos	2	3	1	13	**121**
Alistair Overeem	3	1	4	7	16
Brock Lesnar	4	18	**196**	16	**144**
Fabricio Werdum	5	10	25	5	0
Frank Mir	6	4	4	4	4
Shane Carwin	7	22	**225**	25	**324**
Josh Barnett	8	2	36	8	0
Daniel Cormier	9	8	1	26	**289**
Antonio R. Nogueira	10	9	1	2	64

Table IV. Detailed results.

—In all cases but the Featherweight category, in which the overall results were not good, the best results were obtained with α varying between 0.20 and 0.30 for the strategies that use the edge weighting function[8]. This indicates that in order to produce good results one could look only for values within a small range, reducing the problem of searching for the best parameters for the function.

If we take an even closer look at the rankings, we may see some interesting results. Tables IV(a) and IV(b) show the "top 10" rankings produced by our proposed method and the baseline's for the "Light Heavyweight" and "Heavyweight" categories, respectively. In these figures, the first column corresponds to the fighter, the second to his rank position in the gold standard, the third corresponds to the rank position produced by our method, and the fourth to the square of the difference—as used by the Speaman Rank correlation metric—between the predicted rank position and the "correct" one. The fifth and sixth colums show similar information but for the baseline ranking. Differences in rank positions which are higher than 10 ($D^2 > 100$) are marked in bold.

We can see in the those tables that our ranking and the gold standard have significant similarities. Moreover, when there is a difference in the rank position of a given fighter when compared to his position in the gold standard, this difference is usually small (see values in column D^2), with cases of exact prediction. Furthermore, these differences are in most cases much lower than those produced by the baseline's ranking. Results in terms of recall are also very good. For the Heavyweight category from 23 fighters present in the gold standard rank, we were are able to retrieve 22 with our proposed method, an excellent result (96% of recall). For the Light Heavyweight category there are 19 out of 23 fighters in our ranking, also a significant result (83% of recall). In sum, we can consider that the ranks produced by our time-aware ranking method are, in fact, very satisfactory in several cases.

been alleviated by our small adaptation of the metric as explained in Section 5.2.3.

[8]In fact, in the category Bantamweight, best results were obtained with $\alpha = 0.35$ but these are basically the same as those obtained with $\alpha = 0.30$.

5.2. Tennis Experiments

5.2.1. Experimental Design

In the second round of experiments we applied our method to the tennis dataset. We adopted the same experimental design, building match networks for different time spans instead of categories, and varying the weighting method for the edges. We next applied the PageRank algorithm to rank athletes and compared the generated rankings with a gold standard.

We generated networks considering 8 points of interest for which we generated and compared rankings, ranging from 1975 to 2010 with a 5 year step. Each point of interest corresponds to a time span having 1968 as its first year and the point of interest as the final year. Therefore, we generated networks with matches that occurred during the periods 1968-1975, 1968-1980 and so on. For each time span we tested two weighting strategies, as we did in the previous experiment:

(1) Cumulative: No time-aware function is applied to the weight of the edges. Notice that this is the same as considering $\alpha = 0$ in our edge weighting function described in Section 3.4.

(2) Time-aware: The time-aware function is applied to the weight of the edges (for this, we varied α from 0.25 to 1.75 with a 0.25 step). Notice that α is tested with higher values because of the larger time span of the networks.

Therefore, we generated a total of 64 networks: 8 time spans $\times$ 8 variations of the edge weighting function (one cumulative and seven functions with temporal decay).

In this second round of experiments we also wanted to verify whether generating networks considering just the matches that occurred near to a given point of interest could lead us to better rankings than considering all matches up to that point. For example, given the year 2000 as a point of interest, we would like to know whether building a network with matches that occurred during the years 1999 to 2000 would give us better rankings than considering matches that occurred during the whole period of 1968 to 2000. Therefore, we also generated networks that considered shorter periods (called time spans), including only matches that occurred near to a point of interest. More specifically, for each point of interest, we considered six time spans, which consisted of fixing the point of interest as the final year and varying the initial year in an one year step. For example, given 2000 as the year of interest, we built six networks, considering the matches that occurred during the time spans 1995-2000, 1996-2000,...,2000-2000. For each time span we varied the edge weighting strategy similarly as described above, but varying α from 0.05 to 0.35 with a 0.05 step.

In this way, for this experiment we generated 384 networks: 48 time spans (8 points of interests $\times$ 6 time spans) $\times$ 8 variations of the edge weighting function (one cumulative and seven functions with temporal decay). The eight points of interested in our experiment are the years of 1975, 1980, 1985, 1990, 1995, 2000, 2005 and 2010, which, as we will see below, correspond to the years that we want to contrast the rankings generated by our method for specific time spans with the corresponding ATP official ranking.

Thus, all together, we generated a total of 448 networks for the tennis dataset.

5.2.2. Gold Standard

The natural gold standard to compare the rankings of tennis players is the official ATP ranking. Thus, for each network we consider as the gold standard the ATP ranking of the respective year (point) of interest.

5.2.3. Evaluation Metrics

Likewise the MMA experiments, we used the same metrics, Recall and the Spearman's Rank Correlation Coefficient, to compare the rankings generated by our method and the baseline's with the respective gold standard.

5.2.4. Results

Table V shows the best results obtained for the 448 tennis matches networks, sorted by descending order of recall and Spearman's Coefficient values, considering the following strategies:

(1) Full period network with cumulative edge weighting function: The network is built considering matches from the whole period, i.e., from the year 1968 to the point of interest. No time-aware function is applied to the weight of the edges. Notice that this is the same as considering $\alpha = 0$ in our edge weighting function described in Section 3.4.
(2) Full period network with time-aware edge weighting function: The network is built considering matches from the whole period, i.e., from the year 1968 to the point of interest. The time-aware function is applied to the weight of the edges for this by varying α from 0.25 to 1.75 with a 0.25 step.
(3) Time span network with cumulative edge weighting function: The network is built considering matches that occurred at most five years before the point of interest. No time-aware function is applied to the weight of the edges. Notice that this is the same as considering $\alpha = 0$ in our edge weighting function described in Section 3.4.
(4) Time span network with time-aware edge weighting function: The network is built considering matches that occurred at most five years before the point of interest. The time-aware function is applied to the weight of the edges for this by varying α from 0.05 to 0.35 with a 0.05 step.

By analyzing the results, we can make the following comments:

—Considering the cumulative edge weighting function in the full period network leads to poor results. This is expected, since ATP considers only the respective year's results to produce its final ranking, whereas that strategy considers all matches equally important, no matter when they occurred.
—In all points of interest analyzed, the strategies that use the time-aware edge weighting function ($\alpha > 0$) outperform the ones with cumulative edge weighting function, although we may consider a tie in some cases.
—In general, considering small time spans (two to three years) to produce a ranking leads to good results, i.e., the produced rankings are very close to the official ATP rankings.

6. CONCLUSIONS

We have extended a recently proposed method [Radicchi and Perc 2011] that exploits complex network metrics to produce ranks in sport social networks. Our extensions cover issues related to the temporal aspects inherent to sports events, mainly when the goal is to generate rankings for specific points in time (e.g., current rankings).

Our strategy applies a time-aware function to the weights of the edges of the network to capture the notion that the results of older confronts are not so important as newer ones to predict a more recent ranking (or rankings in a specific point in time). This notion may also be true to other scenarios, like, for instance, the collaborations in scientific networks. Thus, our strategy may also be relevant to other kinds of networks.

Time Span	α	Recall	Spearman's Coef.
1968 - 1975	0.50	0.88	0.61
1972 - 1975	0.35	0.88	0.60
1973 - 1975	0	0.88	0.54
1968 - 1975	0	0.76	-0.12
1980 - 1980	0.05	0.88	0.87
1980 - 1980	0	0.88	0.87
1968 - 1980	1.00	0.80	0.88
1968 - 1980	0	0.36	-1.43
1968 - 1985	0.75	0.92	0.82
1984 - 1985	0.25	0.92	0.81
1984 - 1985	0	0.92	0.72
1968 - 1985	0	0.12	-3.82
1968 - 1990	1.25	0.92	0.72
1990 - 1990	0	0.88	0.85
1990 - 1990	0.05	0.88	0.84
1968 - 1990	0	0.08	-4.32
1995 - 1995	0.05	0.96	0.83
1995 - 1995	0	0.96	0.83
1968 - 1995	1.00	0.88	0.90
1968 - 1995	0	0.08	-4.42
1968 - 2000	1.75	0.80	0.55
1999 - 2000	0.35	0.80	0.48
2000 - 2000	0	0.76	0.53
1968 - 2000	0	0.12	-3.84
1968 - 2005	0.75	0.88	0.82
2004 - 2005	0.25	0.84	0.80
2005 - 2005	0	0.84	0.73
1968 - 2005	0	0.04	-5.12
2009 - 2010	0.15	0.84	0.90
2010 - 2010	0	0.84	0.70
1968 - 2010	0.25	0.60	0.43
1968 - 2010	0	0.12	-3.66

Table V. Best results of the 484 tennis confront networks. The results are grouped by points of interest. Notice that when $\alpha = 0$ the cumulative edge weighting strategy is in use. When $\alpha > 0$, the decaying strategy is used.

We applied our proposed time-aware method to networks built based on results of confronts in to different sports, Mixed Martial Arts (MMA) and tennis. We compared the results produced by our method and the baseline's against a gold standard generated credible rankings and verified that ours outperformed the baseline in all situations, sometimes by very large margins. Moreover, for MMA the results obtained for the parameterization of the methods show that the best values are stable in a small range of values (between 0.20 and 0.30) and that the ranks we produced are very satisfactory, being very similar in some cases to the corresponding gold standard. For tennis, the produced rankings are also very close to the ATP ones, leading us to conclude that, as raised by [Radicchi and Perc 2011], we need only to take into account a recent sequence of match results for for establishing a ranking of players, without considering the actual tournaments and a complex punctuation system.

In this work, we chose to work with a specific type of social networks focused on sports. However, we argue that our solutions can be applied to other contexts. The solution is based on the notion of temporal events between nodes that are characterized by some type of advantage of one node over the other. The events can be any type of dispute or conflict of interest between pairs of nodes that, when resolved, gives advantage to one of the nodes. Judicial disputes, for instance, could be one such an example in which nodes would represent the lawyers involved. Others could be easily thought. Moreover, the approach could also be easily adapted to win-win situations such as collaborations in which the weights are proportional to the level of collaboration but that may decay over time. Scientific social networks are one example of this type of network and the rank methods could be used, for instance, for recommending new collaborations.

As future work, besides applying our solutions to other contexts, as we just discussed, we intend to experiment with alternative time-aware edge weighting functions, investigate better ways to determine the α parameter and perform a more thorough comparison of our method with others under several criteria.

REFERENCES

Barabási, A.-L. Scale-free networks: A decade and beyond. *Science* 325 (5939): 412–413, 2009.

Barabási, A.-L., Jeong, H., Nda, Z., Ravasz, E., Schubert, A., and Vicsek, T. Evolution of the social network of scientific collaborations. *Physica A* 311 (3-4): 590–614, 2002.

Berberich, K., Vazirgiannis, M., and Weikum, G. T-rank: Time-aware authority ranking. In *Algorithms and Models for the Web-Graph*, S. Leonardi (Ed.). Lecture Notes in Computer Science, vol. 3243. Springer Berlin Heidelberg, pp. 131–142, 2004.

Cotta, C., Mora, A. M., Merelo-Molina, C., and Julin Merelo, J. FIFA World Cup 2010: A Network Analysis of the Champion Team Play. *ArXiv e-prints*, Aug., 2011.

Freire, V. P. and Figueiredo, D. R. Ranking in collaboration networks using a group based metric. *J. Braz. Comp. Soc.* 17 (4): 255–266, 2011.

Huang, J., Zhuang, Z., Li, J., and Giles, C. L. Collaboration over Time: Characterizing and Modeling Network Evolution. In *Proc. of WSDM*. pp. 107–116, 2008.

Kudelka, M., Hork, Z., Snasel, V., Kromer, P., Platos, J., and Abraham, A. Social and swarm aspects of co-authorship network. *Logic Journal of IGPL*, 2011.

Laender, A. H. F., Ribeiro-Neto, B. A., da Silva, A. S., and Teixeira, J. S. A brief survey of web data extraction tools. *SIGMOD Record* 31 (2): 84–93, 2002.

Leskovec, J., Backstrom, L., Kumar, R., and Tomkins, A. Microscopic evolution of social networks. In *Proc. of ACM SIGKDD*. pp. 462–470, 2008.

Mourão, F., Rocha, L., Miranda, L., Almeida, V., and Meira, Jr., W. Quantifying the impact of information aggregation on complex networks: A temporal perspective. In *Proc. of the Workshop on Algorithms and Models for the Web-Graph*. pp. 50–61, 2009.

Newman, M. E. J. Who Is the Best Connected Scientist? A Study of Scientific Coauthorship Networks. In *Complex Networks*. Springer, pp. 337–370, 2004.

Newman, M. E. J. *Networks: An Introduction*. Oxford University Press, 2010.

Onody, R. N. and de Castro, P. A. Complex network study of Brazilian soccer players. *Phys. Rev. E* 70 (3): 037103, Sept., 2004.

Page, L., Brin, S., Motwani, R., and Winograd, T. The PageRank Citation Ranking: Bringing Order to the Web. Technical Report 1999-66, Stanford InfoLab. Nov., 1999.

Radicchi, F. and Perc, M. Who is the best player ever? A complex network analysis of the history of professional tennis. *PloS one* 6 (2): e17249, 2011.

Sharan, U. and Neville, J. Exploiting time-varying relationships in statistical relational models. In *Proc. of SNA-KDD Workshop on Web Mining and Social Network Analysis*. pp. 9–15, 2007.

Vaz de Melo, P. O., Almeida, V. A., and Loureiro, A. A. Can Complex Network Metrics Predict the Behavior of NBA Teams? In *Proc. of ACM SIGKDD*. pp. 695–703, 2008.

White, S. and Smyth, P. Algorithms for estimating relative importance in networks. In *Proc. of ACM SIGKDD*. KDD '03. Washington, D.C., pp. 266–275, 2003.