

Who was the greatest of all-time? A historical analysis by a complex network of professional boxing

ADAM G. TENNANT[†]

*Department of Engineering, 2030 Business and Engineering Center, University of Southern Indiana,
8600 University Boulevard, Evansville, IN 47712, USA*

[†]Corresponding author. Email: atennant@usi.edu

CHASE M. L. SMITH

*Kinesiology and Sport Department, Health Professions Center 3092, University of Southern Indiana,
8600 University Boulevard, Evansville, IN 47712, USA*

AND

JOTAM E. CHEN C

*Department of Engineering, 2030 Business and Engineering Center, University of Southern Indiana,
8600 University Boulevard, Evansville, IN 47712, USA*

Edited by: Ernesto Estrada

[Received on 16 October 2019; editorial decision on 23 January 2020; accepted on 3 February 2020]

This study seeks to examine and compare boxers throughout history creating a pound-for-pound list of the different fighters. A PageRank algorithm was utilized to rank the boxers from the network to determine a list of the top 10 fighters from 1897 to 2019. Two data sets were utilized, a truncated subset and a larger data set, to explore the impact of network size on the rank of boxers. Additionally, the researchers systematically varied the damping factor of the PageRank algorithm to determine the effects on the rankings. A discussion of the results includes a comparison of journalistic rankings and those from a points-based system from the respected boxing website BoxRec.

Keywords: boxing; PageRank; ranking.

1. Introduction

Pugilistic historians for more than a century have engaged each other in a theoretical debate on who was the greatest of all time within the four corners of the boxing ring. Often these debates fall short due to a lack of evidence, these arguments must encompass wide time frames, multiple weight classes of boxers and a massive data set. The debaters must rely on speculative arguments such as boxer A's speed was far superior to boxer B's and this would have won the fight for them. These debates can be entertaining in a pop culture sense but only address the small data set of fighters that one person's mind can bring to the debate. Additionally, they are tremendously biased to one individual's preferences for style of fighting or even ethnic bias.

The culture of sport in America is one that prides itself on undisputed, unquestioned, unanimous champions for both team sports and individual. Typically, fans can identify these champions from special events (e.g. Super Bowl, US Open, World Series, The Masters, etc.). While the boxing world contains

title fights, fans are only privy to identifying an undisputed champion within a weight class. This leads many to wonder who is the best of the best. The pound-for-pound (P4P) argument is probably one of the most contentious in the sport and everyone seems to have an opinion on it [1]. This study attempts to breakout this debate, from speculation to more quantitative arguments by sharing preliminary findings of a sports data analytics approach. The researchers in this study ranked individual boxers through the PageRank algorithm and created a historical model of the complexity of the sport. The specifics of the research are discussed in this study along with the results of the rankings.

Although the science of boxing for the combatants is brutally simple, to hit and not get hit, the sport additionally offers a rich source of data for complexity science. Boxing, with its long history, has a depth to its data that can be explored with complexity science to yield insights that analysis from standard statistics has left in the dark. The researchers will explore whether Google's PageRank algorithm sufficiently creates an all-time greatest P4P list regardless of weight class or historical time frame.

2. Background

The depth and breadth of complexity science and network science provides insight into such diverse set of research topics as: cities [2–5], infrastructure [6], health [7], biology [8] and even friendship [9].

Sports is a topic in the field of data analytics that is just beginning to be explored in earnest. Cassidy *et al.* [10] took the approach of using a quadratic-assignment model that could be customized by parameters selection values. A genetic algorithm and other search techniques were employed to rank NCAA Division I-A football with this technique. Deng *et al.* [11] employed a power law ranking method based on prize monies collected or points earned on 12 different sports. The results for the various sports yielded similar graphs with almost matching exponents for the power law.

The PageRank algorithm was developed by Larry Page in 1996 for ranking academic papers. It is a probability distribution that was quickly applied to webpages by use of a weighted network to optimize web searches, where quality of hyperlinks of a webpage produces an advantages PageRank score that pushes the webpage up on search results [12].

Lazova and Basnarkov [13] used FIFA World Cup soccer matches to populate a graph where the PageRank algorithm is applied as an effective ranking system by use of a method to assign weights to the graph based on matches won and goals scored.

Tennant *et al.* [14] produced a general model of the complexity of the sport of boxing exploring match play between welterweight boxers from 2004 to 2015. The PageRank algorithm was used to rank the boxers from a directed graph. The rankings produced were compared with the sport's notoriously corrupt sanctioning bodies, journalistic rankings and other ranking systems. This work supplied further confirmation of the value of the network-based analysis in athletic match play, but was limited in its timespan and had the narrow scope of only analysing welterweight fighters.

2.1 Pound for pound

The approach to deciding who the best-ever in boxing must include a list that consolidates all boxers from all weight classes. Currently, the debate on who is the best of the best within boxing no matter the era is labelled as the P4P champion [15–17]. These lists (i.e. rankings) comprise of fighters who competed directly with other fighters on the lists, and fighters who fought decades prior to other fighters. While there are several different lists publicly available, the majority of those lists come from non-academic sources. Moreover, many of the rubrics used for creating the lists can be argued as biased (e.g. how exciting the

fighter is) or subjective (e.g. quality of opposition). At any rate, the goal for each ranking is to identify the best P4P fighter; to ever fight.

One limitation worth mentioning is the competitive structure professional boxing adopts that influences the different levels of competition a modern-era boxer may face over their career. There are significant factors identified by past scholars [18] that explains reasons why the best boxers seldom fight each other within a particular weight class. The governance structure for professional boxing is not exclusive to one sanctioning body like other individual professional sports (e.g. golf, tennis, bowling). Professional boxing has [at least] four major boxing sanctioning bodies that have significant influence in the industry: World Boxing Council (WBC), World Boxing Association (WBA), International Boxing Federation (IBF) and World Boxing Organization (WBO). Each of these sanctioning bodies claim to have a world champion in each of the respective weight classes [18].

The main implication of having different sanctioning bodies for professional boxing is a lack of ability to identify the best boxer within a weight class. Additionally, there seems to be a lack of oversight in the governance and enforcement of title fights for each sanctioning body, along with corruption within the decisions of fight organizers and promoters [18]. This essentially affects contractual agreements which leads to champion fighters to avoid fighting the best opponents when defending their title. Thus, the limitation exists for the researchers to establish a surface-level boundary to identify which fighters had comparable fighting resumes.

2.2 *Past rubrics for all-time best*

The P4P lists contain substantial popularity for active boxers. The Transnational Boxing Rankings Board [19] has consistently [since 2012] put out the perennial top-10 P4P rankings. They utilize a point-system to attempt to quantify the answers to the following questions to determine their list: how high is the quality of the fighter's recent opposition? How strong is the fighter's career-long resume? How advanced is the fighter's ring generalship? How willing is the fighter to accept all-comers? The authors of this list acknowledge the subjectivity of the list by mentioning the frequent occurrence of a lively debate.

When considering the best fighters in history, McRae [15] acknowledged the element of subjectivity in the list for the top 50 P4P boxers of all time. Factors like wins, losses, world titles and the quality of opposition were variables incorporated. The appeal of the list seemed to be the looseness of the list itself, thus allowing for enough ambiguity to spark a discussion. Mulvaney [16] seemed to clarify reasoning for a more objective rubric by including significant variables such as: in-ring performance, achievements, dominance and mainstream appeal. However, even by attempting to quantify these significant variables, the author appears to acknowledge that choosing these particular variables is, in fact, subjective. This list, then, is not just the 50 greatest fighters of all time. It is the 50 fighters who were the greatest in their time [16].

When considering a fighter's current status, Benson [1] explains that four criteria are considered when ranking fighters P4P. The factors are result, performance, resume and activity. The result of a fight trumps the other factors and essentially is the most influential. The other three factors are equivalent in value. The performance of a fighter is the manner in which a fighter's result is determined. The resume is simply the opponents the fighter beat, and titles achieved. The activity is the frequency a fighter fights against quality competition.

To date, BoxRec [20] appears to quantify the subjective nature of determining the best P4P boxer of all time with their rubrics and methods. Their current ratings structure considers 35 different points of measurement. The ratings structure not only considers performance, results and timing of the fights, but also applies ceilings and limits to the amount a boxer's rating can differentiate between fights (e.g.

The rating of a boxer is reduced by up to 50% in proportion to the difference of two times the rating points of his best opponent in this time period minus his own rating). While BoxRec.com acknowledges that there are inaccuracies and anomalies due to missing historical data, the foundation for their ratings system stands on the unbiased approach to their formulas. However, it appears there still lies ambiguity in the ceilings imposed within the formula, and total points awarded (e.g. 1 point, when a boxer defeats an opponent, who already won a bout against a winning opponent within 18 months) for a number of the criteria points. This gives reason to analyse further using different methods to determine a P4P boxing ranking (e.g. Page Rank).

2.3 *Data set*

BoxRec is an online database that is the preferred source for boxing data post the Marquess of Queensberry Rules. Both male and female, past and present professional boxer's records are catalogued on this website for the public to explore. There are in excess of over 2.1 million bouts entered into the database that have been researched and validated by global volunteers. Besides the results, the boxing data bout history contained on the site can include venues, referees, judges, official weights, knockout times, promoters and descriptions of many bouts. While small subsets are available categorized by date or boxer, the complete data set is not available for download in its entirety. The researchers approached the staff of the website to obtain the whole data set, and they graciously agreed to send the data set in .csv format. Researchers took a historical perspective emphasizing the significance of the sport's long history by requesting a data set that included bouts from 1897 to 2019.

The investigation that follows utilized two data sets provided by BoxRec.com. These two data sets were already filtered by BoxRec's scoring system for bouts that involved fighters with at least 10 points or 50 points. This scoring system that BoxRec employs is essentially an algebraic formula that involves several conditions that are checked and then re-calculated daily, incorporating any new bouts into the database. This filtering was needed as a data set involving all 2.1 million bouts would be unmanageable to work with. Nonetheless, the scale of this data involved up to 307,407 bouts and 61,232 boxers and should be considered more than adequate.

3. **Methods**

3.1 *Network creation-directed graph*

A directed graph which is analogous to hyperlinks (edges) within webpages (nodes) was used by employing the idea into the boxing world by linking fighters through bouts. The researchers started by preprocessing the data in spreadsheet form. This included preparing the data to assist in the creation of the directed graph with some data management clean-up techniques. This involved positioning the data in a source and target arrangement for the nodes to reflect winners and losers making this a directed graph. The edges are the bouts between the fighters that travel from the loser to the winner. The source for the data was the loser of the boxing match and the target was the winner of the bout. A small subset of the entire directed graph can be seen below for understanding in Fig. 1.

For illustrative purposes, a subgraph of the entire directed graph was constructed from heavyweight champion Jack Dempsey's last six fights. This subgraph is made up of six nodes (boxers) and six edges (bouts). The nodes are displayed as the boxer's images with edges depicted as arrows. If an edge is present between nodes this reflects that a bout was fought between these boxers. For example, Dempsey's node has an edge entering this node originating from Firpo, Gibbons, Sharkey and Carpentier. These edges represent victories for Dempsey as the target node with the depicted arrow originating from the losing



FIG. 1. Directed graph example of Jack Dempsey's last six fights.

fighter, or more accurately, defined as the source node. This arrangement gives the graph its directed nature. Conversely, Dempsey shares two other edges with Gene Tunney shown as dotted arrows. These edges are both directed to enter Tunney's node representing two bouts fought with Dempsey and both being losses. Any boxing matches that were fought to a draw, disqualification or no contest were eliminated from the data set as no one won the match and a directed graph could not use this data.

Furthermore, identical edges were combined so that only a single edge could connect a source and a target. An illustrative example of this is seen in Fig. 2 first showing the original arrangement of edges between just two nodes. This example relies on the all-time great saga of fights between Sugar Ray Robinson and Jake LaMotta who fought a staggering six times between 1942 and 1951 with Robinson winning five out of the six. Duplicate edges were combined and then given a weight to reflect that multiple fights status. For example, the original six bouts were combined into two by taking Robinson's five edges that were wins and converting it to one single edge with a weight of 5. The single LaMotta win was left alone and given a weight of 1. These weights will be utilized when calculating the PageRank values for the individual nodes.

The original size of the data sets and the reduced size of the now weighted-directed networks can be seen in Table 1. After the removal of bouts without a winner and the reduction due to weighting the rematches, the networks were both reduced by about 20% of their edges and 1% of their nodes.

3.2 PageRank calculation

Once the directed graph was created and the edges are weighted the teams can be ranked. To rank the boxers, the PageRank algorithm was used, which was originally developed in the webpage search engine

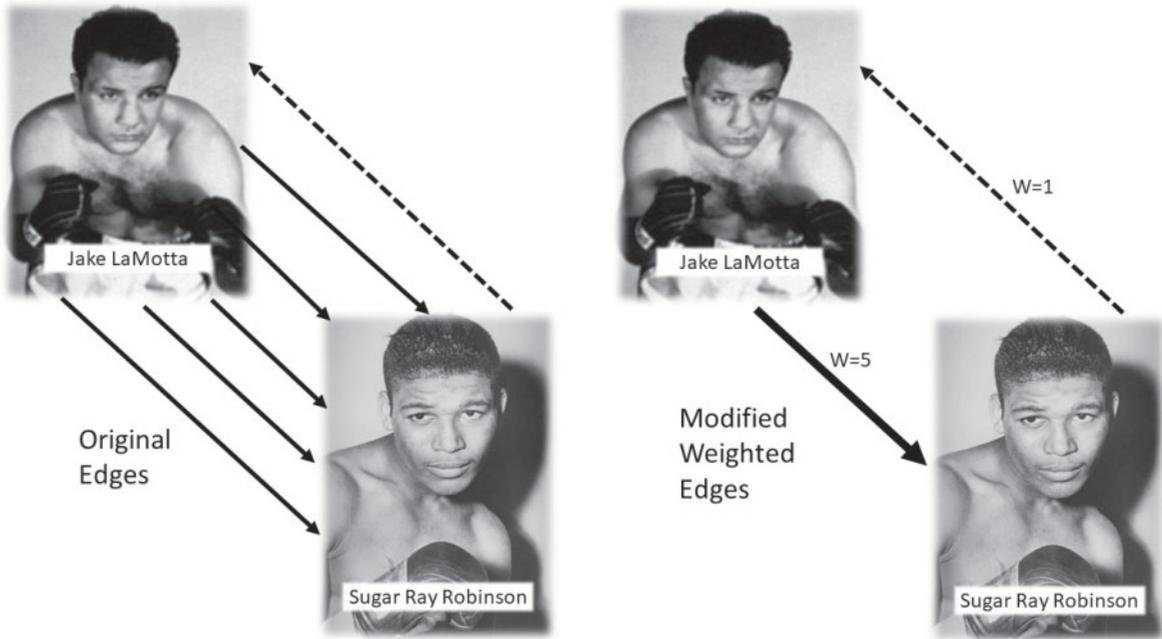


FIG. 2. Directed graph and modified weighted-directed graph example of fights between Sugar Ray Robinson and Jake LaMotta.

TABLE 1 *Network creation statistics*

	Original		Weighted-directed	
	Nodes	Edges	Nodes	Edges
10 pts.	62,033	388,801	61,232	307,407
50 pts.	12,521	58,744	12,425	47,309

revolution of the late 1990's. The concept is that the webpages are assigned higher PageRank values based not only on the number, but also the quality of the links. Page *et al.* (1999) gave the PageRank algorithm in early publications as shown in equation 1:

In terms of boxers, $PR(A)$ is the PageRank of boxer A, $PR(T_i)$ is the PageRank of boxer T_i linked to boxer A with a loss, $C(T_i)$ is the number of edges leaving boxer T_i ; that is, number of losses. Finally, d is a damping factor, empirically preferred to be 0.85 [14].

The algorithm is an iterative process that can be explained by the concept of a random walk through the directed graph. The walker will randomly enter the graph on any source node (boxer) then randomly travel along one of the outward-bound edges (match that was lost) to the target node (winning boxer). This process has been modified in our directed graph to favour certain edges based on the weights discussed previously. This gives the ability to place more emphasis on those edges that represent multiple wins. For instance, if a weight of 3 was present on a particular edge leaving a node, that would give that edge a three times more likely chance of being selected opposed to an edge leaving that node weighted with

TABLE 2 *PageRank top1 10 P4P boxers from 1897 to 2019 for 50 point data set*

Rank	First name	Last name	Career start	Career end	PageRank
1	Sugar Ray	Robinson	4 October 1940	11 October 1965	0.00242
2	Evander	Holyfield	15 November 1984	7 May 2011	0.00211
3	Floyd	Mayweather Jr.	11 October 1996	26 August 2017	0.00206
4	Emile	Griffith	2 June 1958	30 July 1977	0.00204
5	Manny	Pacquiao	22 January 1995	NA	0.00202
6	Harry	Greb	29 May 1913	19 August 1926	0.00190
7	Muhammad	Ali	29 October 1960	11 December 1981	0.00189
8	Archie	Moore	3 September 1935	15 March 1963	0.00187
9	Ezzard	Charles	12 March 1940	1 September 1959	0.00174
10	Lennox	Lewis	27 June 1989	21 June 2003	0.00163

a 1. This random walk will continuously occur and essentially build a probability distribution on what node (boxer) the walker would be located if the algorithm was stopped. The algorithm was run on both the smaller quantity of bouts in the 50-point and larger 10-point data sets.

To force convergence of the PageRank algorithm, a damping factor was utilized, this is often referred to as teleportation when applied to the random walk paradigm. For instance, with a damping factor of $d = 0.85$ on the current node position the walker will 15% of the time not take an outward-bound edge and randomly teleport to another node (boxer) in the directed graph. This process of teleportation avoids the walker being trapped on a node with no edges out (i.e. undefeated boxer), commonly referred to as a dangling node in network science, in this technique non-convergence of the algorithm is avoided. In this study, the standard damping factor of 0.85 was modified for larger 10-point data set to both 0.90 and 0.95 to explore the variance in the results.

Alternatively, the Eigenvector centrality was also calculated for the network to compare later to the results to help explain the advantages of using PageRank. Centrality has been historically used in various networks as an evaluation of individual node importance. The method employed in this research work departs from Tennant *et al.* [14] in the much larger historical data set, the inclusion of all weight classes, and the use of weighted edges.

4. Results and discussion

Utilizing the PageRank algorithm can provide an unbiased way to rank boxers based solely on their PageRank score. For both the smaller 50-point and larger 10-point data set networks the top 100 boxers according to their PageRank are presented in the subgraph visualization displayed in Figs 3 and 4. These directed graphs are filtered by only including the top 100 PageRank results and edges associated between these nodes. The larger diameter and darker green colour nodes represent the higher PageRank values. The top 10 boxers by PageRank score have been labeled in both visualizations to display the results in a relatable form. Additionally, Tables 2 and 3 gives a more detailed view of these results with some supplementary career data for 50- and 10-point data set networks, respectively.

This depiction in Figs 3 and 4 assisted in understanding the significance of the analysis. For instance, Archie Moore in both graphs has many edges leaving or entering his node. Which reflects the historical fact that he was not only an extremely active fighter with 220 bouts, but one that was fighting a high degree

TABLE 3 *PageRank top 10 P4P boxers from 1897 to 2019 for 10 point data set*

Rank	First name	Last name	Career start	Career end	PageRank
1	Floyd	Mayweather Jr.	11 October 1996	26 August 2017	0.00143
2	Manny	Pacquiao	22 January 1995	NA	0.00129
3	Sugar Ray	Robinson	4 October 1940	10 November 1965	0.00104
4	Archie	Moore	3 September 1935	15 March 1963	0.00098
5	Emile	Griffith	2 June 1958	30 July 1977	0.00089
6	Muhammad	Ali	29 October 1960	11 December 1981	0.00086
7	Harry	Greb	29 May 1913	19 August 1926	0.00084
8	Wladimir	Klitschko	16 November 1996	29 April 2017	0.00083
9	Evander	Holyfield	15 November 1984	7 May 2011	0.00076
10	Bernard	Hopkins	11 October 1988	17 December 2016	0.00075

of opposition since these edges appear in this subgraph with only high PageRank boxers. Additionally, these graphs tend to show the clustering of fighters by their weight class and era they fought. For instance, Harry Greb in both subgraph figures, is set aside from the other top 10 fighters. This is due to the era he fought in (i.e. the early 1900s) being significantly different than the eras of the other nine fighters.

By inspecting the list of names provided by the PageRank analysis in Table 2, it becomes immediately evident to even the casual fan of the sport, the power of the technique to rank boxers. The highest ranked boxer, Sugar Ray Robinson, is a well-known pugilistic wizard of the golden era of boxing and is often considered by journalist and fans as the greatest middleweight boxer of all time. Robinson wore the crown of the world middleweight title on five occasions. His career was prolific with a record of 174 victories, 19 defeats and 6 draws from debuting in 1940 and entering the ring for the final time in 1965. Robinson was a well-known celebrity in his time and had gained the respect of the American public for his speed, power, ring generalship, toughness and graceful style [21].

Further down the rankings, the only active fighter on this list, Manny Pacquiao, appears at the number five position. Given the moniker of the Pac Man^ž due to his record of gobbling up opponents his current record stands at 62 victories, 7 defeats and 2 draws and has been active since 1995. Pacquiao had success early on as a champion as a flyweight (112 lb), super bantamweight (122 lb), super featherweight (130 lb), lightweight (135 lb) and currently as a welterweight (147 lb). This success across multiple weight classes connected Pacquiao with many other high-ranking nodes within the directed graph creating a scenario where higher PageRank scores can be expected [22].

As a comparison to the PageRank results, Table 4 displays the Eigenvector centrality results. The names of the top 10 boxers from this analysis are fairly recognizable to any devotee of the history of the sport. Henry Armstrong held the top spot followed by Sugar Ray Robinson, two of the most decorated boxers in the history of the sport, both multi-divisional champs. Though upon a closer look of these results it was discovered that these top 10 spots were completely made up of boxers who haven't fought in at least 50 years. More specifically, a timeline clustering of the Eigenvector centrality results was noticeable. The list was predominantly comprised of boxers who fought in the 1930s and 40s. This is interesting as this is considered by historians to be the era known as the golden age of boxing. Nevertheless, it is highly unlikely that any list of the best of all time should be that biased towards a specific era. Additionally, the list of the top 10 was exclusively composed of lightweights, welterweights and middleweights, the

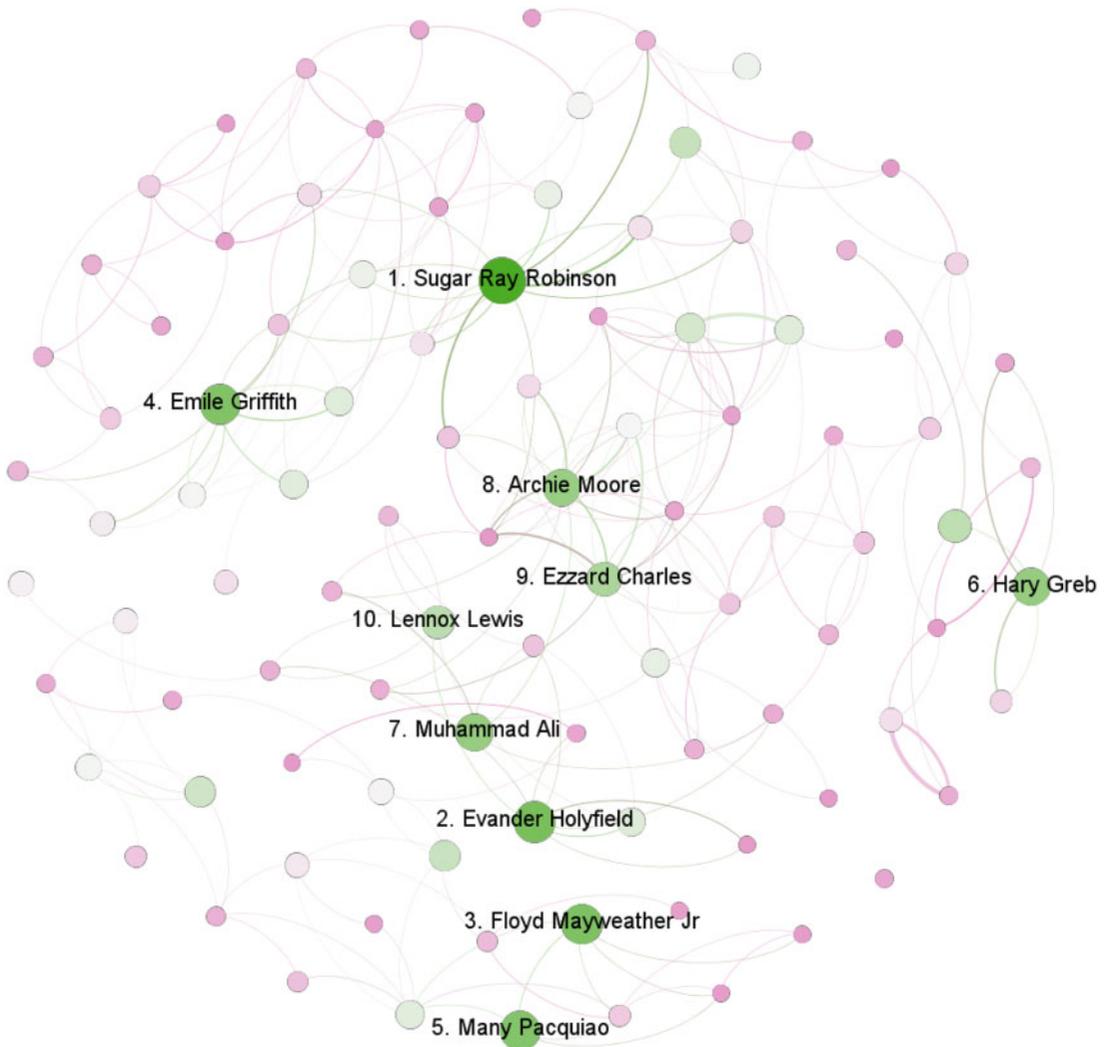


FIG. 3. Weighted directed subgraph filtered by top 100 boxers by PageRank value from 50-point data set.

high value Eigenvector centrality results completely excluded flyweights, bantamweights, featherweights, light heavyweights, cruiserweights and heavyweights. It is improbable that the exclusions of all these weight classes accurately represent the history of the sport in any all-time P4P ranking list.

The difference between the PageRank and Eigenvector centrality is the inclusion of the teleportation phenomenon in the PageRank algorithm. This teleportation phenomena explains the superior wide-ranging results of the PageRank algorithm. Where the Eigenvector centrality identified the broader era, when boxing was potentially at its peak. The PageRank results capture the individual boxer's career achievements superiorly. For a percentage of the iterations of the analysis, the act of teleporting randomly to another node (boxer) likely outside of any cluster the node belongs to prevents the results from being

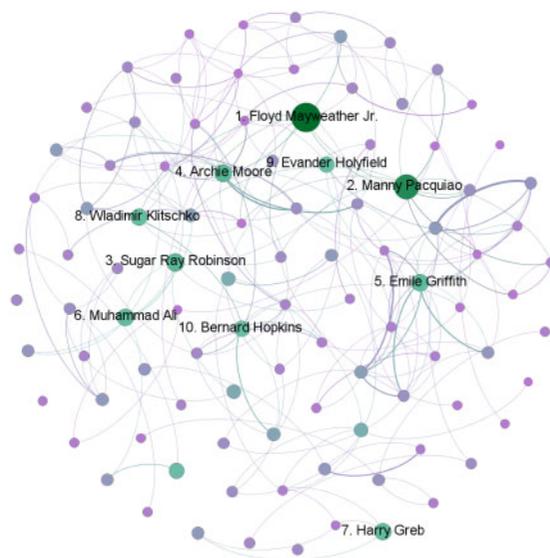


FIG. 4. Weighted directed subgraph filtered by top 100 boxers by PageRank value from 10-point data set.

TABLE 4 *Eigenvector centrality top 10 P4P boxers from 1897 to 2019 for 50 point data set*

Rank	First name	Last name	Career start	Career end	E. centrality
1	Henry	Armstrong	27 July 1931	14 February 1945	1.00000
2	Sugar Ray	Robinson	4 October 1940	10 November 1965	0.96297
3	Tony	Canzoneri	2 July 1925	1 November 1939	0.76707
4	Wesley	Ramey	14 August 1929	20 June 1941	0.75452
5	Cocoa	Kid	27 May 1929	24 August 1948	0.71421
6	Holman	Williams	12 August 1932	30 June 1948	0.71405
7	Fritzie	Zivic	5 October 1931	17 January 1949	0.69543
8	Sammy	Angott	9 March 1935	8 August 1950	0.69104
9	Ike	Williams	15 March 1940	12 August 1955	0.62509
10	Lou	Ambers	9 June 1932	28 February 1941	0.60380

centralized to any specific era or grouping of weight classes in the sport. This spreads the higher ranks out to various weight classes and to a considerably less biased time frame.

In Table 5, the results of modifying the damping factor for the larger 10-point network is presented. Floyd Mayweather Junior does stay in the top spot for each of these rankings, but one noticeable trend took place. Successful heavyweight boxers moved up in the rankings as the damping factor is increased, the most noticeable being Klitschko, who in the 0.85 analysis was ranked 8 but at the 0.95 analysis had climbed to the number 2 position. Within these three separate rankings for the top 10 fighters, the heavyweight fighters either climbed in rankings or remained constant.

TABLE 5 *The ranking results with varying PageRank damping factors*

Rank	PageRank algorithm damping factors		
	0.85	0.9	0.95
1	Mayweather Jr.	Mayweather Jr.	Mayweather Jr.
2	Pacquiao	Pacquiao	Klitschko
3	Robinson	Robinson	Ali
4	Moore	Moore	Moore
5	Griffith	Griffith	Robinson
6	Ali	Ali	Pacquiao
7	Greb	Klitschko	Holyfield
8	Klitschko	Holyfield	Griffith
9	Holyfield	Greb	Charles
10	Hopkins	Charles	Lewis

The assumption behind this phenomenon is that boxers tend to move up in weight due to age or looking for a big money fight. When a naturally smaller boxer moves up in weight to fight a larger boxer he is at a disadvantage and will be more likely to lose. This basic fact will eventually produce more edges leading to larger boxers eventually terminating at a heavyweight boxer. The lower the damping factor measures, the more likely the random walker is to be teleported and not be directed along this trend. Since this is a P4P list and the researchers are attempting to avoid bias based on the size of prize fighter, the 0.85 damping factor is preferred as it lessens this tendency.

In Table 6, the PageRank results for the 10-point data set with a damping factor of 0.85 are compared with the corresponding rank from BoxRec's all-time list. The top two spots directly match with Mayweather and Pacquiao, validating the high skill level of both these fighters who have been active the last 20 years. A two-dimensional plot containing the BoxRec rankings and the PageRank results is shown in Fig. 5 a rather weak correlation is shown between the two through a simple linear curve fit. Nevertheless, this visualization is for data exploration which helps with the corresponding table to identify the presence of outliers to investigate.

Taking an inquisitive look at those rankings with the highest discrepancy boxers like Holman Williams and Cocoa Kid seem to be ranked either excessively high or low by the PageRank algorithm or BoxRec, respectively. A thorough inspection of these fighter's data shows that Williams and Kid fought a staggering 12 times against each other, an anomaly even before the 1960's when rematches were more common. The high weight on these edges makes them much more likely to be selected during the random walk thus potentially skewing the results. A limit could be worked into the network to reduce excessively high weights due to rematches to a prescribed max weight. It is worth noting that Williams and Kid are highly reputable all-time great fighters that have been inducted into the International Boxing Hall of Fame [23].

Maxie Rosenbloom and Henry Armstrong are another two outliers in the linear curve fit. First exploring Rosenbloom's career, it is remarkable to see he fought a staggering 274 times from 1923 to 1939 which overlaps with the timeframe of the sport's golden age. Nicknamed Slapsie Maxie due to his softer open fist punching style he amassed a record of 207 wins, 39 losses and 26 draws. Rosenbloom is considered one of the most elusive fighters to punch in the history of the sport [21]. His high degree of activity

TABLE 6 *Comparison of PageRank and associated BoxRec ranking*

Boxer	Rankings	
	PageRank	BoxRec
Floyd Mayweather Jr.	1	1
Manny Pacquiao	2	2
Sugar Ray Robinson	3	5
Archie Moore	4	8
Emile Griffith	5	43
Muhammad Ali	6	4
Harry Greb	7	19
Wladimir Klitschko	8	36
Evander Holyfield	9	11
Bernard Hopkins	10	6
Ezzard Charles	11	63
Roy Jones Jr.	12	20
Saul Alvarez	13	50
Andre Ward	14	33
Maxie Rosenbloom	15	342
Holman Williams	16	408
Dick Tiger	17	66
Henry Armstrong	18	96
Cocoa Kid	19	923
Lennox Lewis	20	79

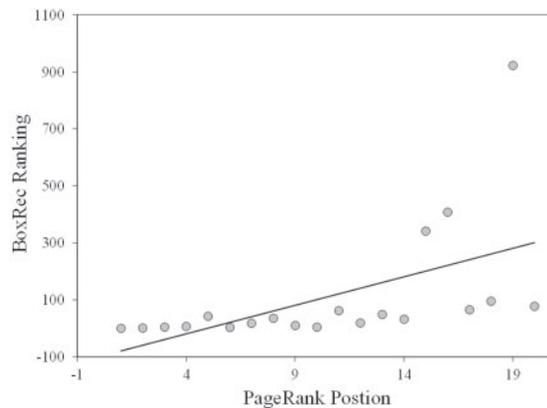


FIG. 5. BoxRec rankings vs. PageRank positions.

and success in this timeframe with higher eigenvector centrality values correspondingly explain his high PageRank score.

TABLE 7 *Comparison of PageRank and various boxing journalist rankings*

Rank	PageRank	Kieran Mulvaney (2007)	Bert Sugar (1988)	Kevin McRae (2012)
1	Floyd Mayweather Jr.	Sugar Ray Robinson	Sugar Ray Robinson	Sugar Ray Robinson
2	Manny Pacquiao	Muhammad Ali	Henry Armstrong	Henry Armstrong
3	Sugar Ray Robinson	Henry Armstrong	Harry Greb	Willie Pep
4	Archie Moore	Joe Louis	Jack Dempsey	Muhammad Ali
5	Emile Griffith	Willie Pep	Benny Leonard	Joe Louis
6	Muhammad Ali	Roberto Duran	Joe Louis	Roberto Duran
7	Harry Greb	Benny Leonard	Mickey Walker	Jack Johnson
8	Wladimir Klitschko	Jack Johnson	Sam Langford	Jack Dempsey
9	Evander Holyfield	Jack Dempsey	Tony Canzoneri	Benny Leonard
10	Bernard Hopkins	Sam Langford	Muhammad Ali	Sugar Ray Leonard
11	Ezzard Charles	Joe Gans	Joe Gans	Harry Greb
12	Roy Jones Jr.	Sugar Ray Leonard	Willie Pep	Joe Gans
13	Saul Alvarez	Harry Greb	Jack Johnson	Sam Langford
14	Andre Ward	Rocky Marciano	Barney Ross	Gene Tunney
15	Maxie Rosenbloom	Jimmy Wilde	Jimmy Wilde	Rocky Marciano
16	Holman Williams	Gene Tunney	Gene Tunney	Archie Moore
17	Dick Tiger	Mickey Walker	Roberto Duran	Jimmy Wilde
18	Henry Armstrong	Archie Moore	Johnny Dundee	Mickey Walker
19	Cocoa Kid	Stanley Ketchel	Rocky Marciano	Julio Cesar Chavez
20	Lennox Lewis	George Foreman	Joe Walcott	George Foreman

Reviewing the career of Henry Armstrong a fighter who was ranked 96th by BoxRec and 18th by the authors gives further proof of the validity of the PageRank Algorithm. As will be seen below Armstrong is considered one of the all-time greats by historians of the sport, he is the sole boxer to simultaneously hold world championships in three weight classes (feather-, light- and welter-weight). Nicknamed Homicide Hank he was the polar opposite to Rosenbloom crafty style by fighting with a come forward reckless hard hitting style [21]. He also was active in the timeframe 1931–1945 that overlapped with the sport's golden age that encapsulated the higher eigenvector centrality values for that era. BoxRec 96th rank is rather low for such an accomplished boxer, the PageRank results align more accurately with this higher rank expectation based on the historical significance of Armstrong's career.

Table 7 takes a comparative look at the 10-point data set with a damping factor of 0.85 PageRank listings with three separate journalists top 20 boxers. These three lists have been composed by highly respected boxing writers and historians, with Burt Sugar's list the most highly esteemed [15–17]. These lists have many similarities and it should be remembered that they were not created in a vacuum and were likely impacted by each other either consciously or subconsciously. All three lists are highly skewed with 15 out of the 20 boxers fighting before the 1960's. Some boxers on the list have very little film of their events and their rankings are based on second-hand accounts. It is highly likely that these lists are falling under an overly nostalgic look on a period when the sport of boxing was far more culturally significant. These boxers were household names and commanded the world's attention with demigod like status. A qualitative non-analytic list would be hard pressed to ignore this and not over emphasize a boxer's career or era.

The PageRank listings are slightly skewed to boxers who fought after the 1980's with 9 out of the 20 boxers coming from this time frame. This result is likely caused by the fact that as boxer's age they are more likely to lose as they run into younger opposition. This would have a cascading effect and tend to push higher PageRank scores in a directed graph to more current boxers. Further research could use modified weights adjusted to place more emphasis on wins from the past by use of factors from a timeline scale. In general, the PageRank results had significantly better distribution of boxers from the many eras included in the analysis than previous journalistic rankings.

5. Equations

$$PR(A) = (1 - d) + d \left[\frac{PR(T_1)}{C(T_1)} + \dots + \frac{PR(T_n)}{C(T_n)} \right] \quad (1)$$

Acknowledgements

The authors acknowledges the support of the BoxRec.com and especially Martin Reichert within that organization.

REFERENCES

1. BENSON, M. (2019) Ranked top ten pound-for-pound boxers in the world, including vasily lomachenko and canelo alvarez. <https://talksport.com/sport/boxing/466713/top-ten-pound-for-pound-boxers-vasily-lomachenko-canelo/>, August 2019.
2. BATTY, M. (2007) *Cities and Complexity: Understanding Cities with Cellular Automata, Agent-based Models, and Fractals*. Cambridge, MA: The MIT Press.
3. BATTY, M. (2013) *The New Science of Cities*. Cambridge, MA: MIT Press.
4. DERRIBLE, S. (2017) Urban infrastructure is not a tree: integrating and decentralizing urban infrastructure systems. *Environ. Plan. B*, **44**, 553–569.
5. PEIRAVIAN, F. & DERRIBLE, S. (2017) Multi-dimensional geometric complexity in urban transportation systems. *J. Transp. Land Use*, 10.
6. DERRIBLE, S. & KENNEDY, C. (2010) The complexity and robustness of metro networks. *Physica A*, **389**, 3678–3691.
7. STURMBERG, J. P. & MARTIN, C. (2013) *Handbook of Systems and Complexity in Health*. New York, NY: Springer Science & Business Media.
8. RUAL, J.-F., VENKATESAN, K., HAO, T., HIROZANE-KISHIKAWA, T., DRICOT, A., LI, N., BERRIZ, G. F., GIBBONS, F. D., DREZE, M. & AYIVI-GUEDEHOSSOU, N. (2005) Towards a proteome-scale map of the human protein interaction network. *Nature*, **437**, 1173.
9. BORGE-HOLTHOEFER, J., BANOS, R. A., GONZÁLEZ-BAILÓN, S. & MORENO, Y. (2013) Cascading behaviour in complex socio-technical networks. *J. Complex Netw.*, **1**, 3–24.
10. CASSADY, C. R., MAILLART, L. M. & SALMAN, S. (2005) Ranking sports teams: a customizable quadratic assignment approach. *INFORMS J. Appl. Anal.*, **35**, 497–510.
11. DENG, W., LI, W., CAI, X., BULOUL, A. & WANG, Q. A. (2012) Universal scaling in sports ranking. *N. J. Phys.*, **14**, 093038.
12. PAGE, L., BRIN, S., MOTWANI, R. & WINOGRAD, T. (1999) The PageRank citation ranking: bringing order to the web. *Technical Report*.
13. LAZOVA, V. & BASNARKOV, L. (2015) PageRank approach to ranking national football teams. *arXiv preprint arXiv:1503.01331*.

14. TENNANT, A. G., AHMAD, N. & DERRIBLE, S. (2017) Complexity analysis in the sport of boxing. *J. Complex Netw.*, **5**, 953–963.
15. MCRAE, K. (2012) Bleacher report: the top 50 pound-for-pound boxers of all time. <https://bleacherreport.com/articles/1436191> December 2012.
16. MULVANEY, K. (2007) Espn:50 greatest. <http://www.espn.com/sports/boxing/greatest/featureVideo?page=greatest4150>, May 2007.
17. SUGAR, B. R. (1988) *100 Greatest Boxers of All Time*, revised edition. New York, NY: Crescent.
18. ANDREF, W. & SZYMANSKI, S. (2006) *Handbook on the Economics of Sport*. Cheltenham, UK: Edward Elgar Publishing.
19. Transnational Boxing Rankings Board. (2019) P4P. <http://www.tbrb.org/p4p/>, September 2019.
20. BOXREC. (2019) Boxrec ratings description. https://boxrec.com/media/index.php/BoxRec_Ratings_Description, September 2019.
21. ROBERTS, J. B., SKUTT, A. G. & INTERNATIONAL BOXING HALL OF FAME (2006) *The Boxing Register: International Boxing Hall of Fame Official Record Book*. Boxing Register. Ithaca, NY: McBooks Press.
22. BOXREC. (2019) Manny pacquiao's boxing record. <https://boxrec.com/en/proboxer/6129>, July 2019.
23. INTERNATIONAL BOXING HALL OF FAME. (2019) Inductees. <http://www.ibhof.com/pages/about/inductees/inducteeindex.html>, June 2019.