VOL. 88, NO. 1, FEBRUARY 2015 1

A Search for Champion Boxers

Tien Chih
Montana State University-Billings
Billings MT 59101
tien.chih@msubillings.edu
Demitri Plessas
Independent Researcher
demitri.plessas@gmail.com

"Once the fans of history get an idea of the people he beat, then they will get a better perspective of him, I'm sure. He's got an All-Star list of victories. So they're gonna think '... damn he beat all these guys?' People are gonna know. He should be in the top 5, top 3 and stuff."

- Mike Tyson on Evander Holyfield (Chasing Tyson [3])

The 2015 documentary "Chasing Tyson" [3] describes the career of 4-time world heavyweight champion Evander Holyfield. One of the central themes of this film was how due to his low-key style and humble attitude, the boxing public did not accept Holyfield as a top fighter deserving of the accolade "world heavyweight champion". However, former Heavyweight Champion and widely accepted all-time great Mike Tyson is quoted above at the end of the film, making the case for Holyfield's greatness. Although not particularly flashy in or outside the ring, Holyfield's achievement in the sport of boxing may be measured in an objective way: by the list of quality boxers he has defeated.

This argument is reminiscent of Google's PageRank algorithm [2, 10]. PageRank is an algorithm that assigns to web pages a value, based on the web pages who link to it and *their* value, that is receiving an incoming link from a high-value website carries more value than a link from a low-level one. The analogy to boxing is clear; a boxer's status is determined by the opponents they beat, the higher quality the opponents, the more value those victories are worth. Of course, the value of those boxers are determined the same way, so a global analysis of all the fighters involved is necessary.

We describe the mathematics behind PageRank and give a scheme to adapt it to boxing. We illustrate how this algorithm would behave in several scenarios. Then 20 top heavyweight boxers from the 1990's are selected to see if Mike Tyson was justified in his praise of Evander Holyfield. We then make a comparison to the Colley Method of ranking sports teams that also depends on the quality of opponents, and describe some possible extensions.

Shifting Values, Stochastic Matrices, and Markov Chains

In this paper, our model will use Markov chains that are row-stochastic and act on the right of row vectors. In order to model our boxing network and apply the PageRank algorithm, it is necessary to keep two different views of what a Markov chain can represent in mind.

As an example, consider the Markov chain:

$$M = \begin{bmatrix} 1/3 & 1/3 & 1/3 \\ 1/4 & 1/2 & 1/4 \\ 1/2 & 0 & 1/2 \end{bmatrix}$$

which has a weighted digraph representation (Figure 1). One way of interpreting this matrix is as a value shifting scheme. If Alice, Bob and Chen are people each with some amount of water. Then, they pass water to each other in a way so that in each phase of this process, Alice gives $1/3$ of her water away to Bob and Chen, Bob gives away a $1/4$ of his water to Alice and Chen, and Chen gives half his water to Alice.

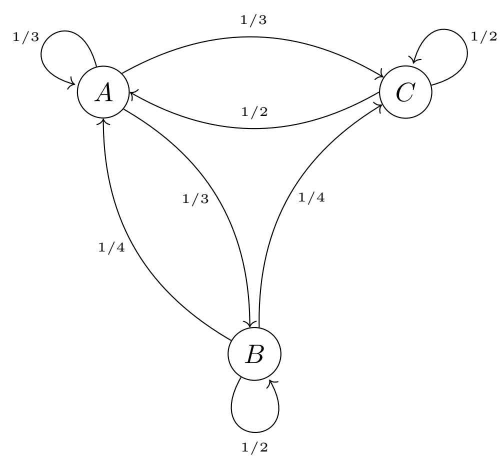

Figure 1 Digraph representation of a Markov chain

It would then be reasonable to ask, is there some quantity of water per person such that after each phase of this process, each person's quantity remains unchanged? In other words, are there solutions to

$$[a \quad b \quad c] \begin{bmatrix} 1/3 & 1/3 & 1/3 \\ 1/4 & 1/2 & 1/4 \\ 1/2 & 0 & 1/2 \end{bmatrix} = [a \quad b \quad c]$$

or is there a left eigenvector for eigenvalue $\lambda = 1$? We can see by inspection that

$$[3/8 \quad 1/4 \quad 3/8]$$

is one such solution.

This eignvector can also answer the question: if Alice were to have a cup of water, and this process were to be repeated ad infinitum, would there be some amount of water for each person that their quantity would converge to? This would be represented as

$$\lim_{n \to \infty} [1, 0, 0] M^n.$$

We can see that as we iterate this process:

$$[1 \quad 0 \quad 0] \begin{bmatrix} 1/3 & 1/3 & 1/3 \\ 1/4 & 1/2 & 1/4 \\ 1/2 & 0 & 1/2 \end{bmatrix} = [1/3 \quad 1/3 \quad 1/3],$$

$$[1 \quad 0 \quad 0] \begin{bmatrix} 1/3 & 1/3 & 1/3 \\ 1/4 & 1/2 & 1/4 \\ 1/2 & 0 & 1/2 \end{bmatrix}^2 = [13/36 \quad 5/18 \quad 13/36], \text{ and}$$

$$[1 \quad 0 \quad 0] \begin{bmatrix} 1/3 & 1/3 & 1/3 \\ 1/4 & 1/2 & 1/4 \\ 1/2 & 0 & 1/2 \end{bmatrix}^{100} \approx [3/8 \quad 1/4 \quad 3/8] \, .$$

We achieve the same values we had before.

Our second way of interpreting this matrix is as a representation of stochastic probability. An ant standing on node A would have a $1/3$ chance of staying on A and a $1/3$ chance of moving to nodes B or C. An ant on node B has a $1/2$ chance of staying put and a $1/4$ chance of moving to either A or C. If on node C, the ant has a $1/2$ chance of staying put and a $1/2$ chance of moving to A. It would be reasonable then to ask then if an ant were to start on a particular node, say A, as this process iterates, what probability would the ant have on landing on each node? Since the probability of the ant landing on any given node after n steps is

$$[1 \quad 0 \quad 0] \begin{bmatrix} 1/3 & 1/3 & 1/3 \\ 1/4 & 1/2 & 1/4 \\ 1/2 & 0 & 1/2 \end{bmatrix}^{n} ,$$

we can surmise that the probability the ant lands on any particular node as n tends to infinity is $3/8$ for A, $1/4$ for B and $3/8$ for C.

So the key to our ranking system is noting that these Markov matrices represent both stochastic probabilities and value shifting schemes. Moreover, reiteration of either process resulted in converging to a *stable state* or left eigenvector for eigenvalue $\lambda = 1$. More on these concepts can be found in [**12**].

Modeling a Boxing Network

We can see why we would want to use stochastic matrices to model the ranking of boxers. After all, our entire premise is that when a boxer defeats an opponent, the value of their opponent is transferred to them, and the more valuable the opponent, the more it should enhance the status of the victor.

Let us consider a simple scenario: let there be three boxers creatively named A, B and C. First A and B fight, and B emerges victorious. They then go on to fight C and lose. We can see a visualization of who loses to whom in the Figure 2.

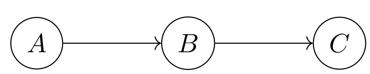

Figure 2 A simple boxing network

Now if there is any justice in the universe, any reasonable scheme to rank these boxers should rank C ahead of B ahead of A. Let's think of this as a value shifting scheme, where each boxer is carrying around a glass of water, and losing to other boxers forces them to share that water. Since A loses to B, A has to give B their water, and similarly, B loses to C. If we represent this as a matrix, we would obtain

$$\begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ ? & ? & ? \end{bmatrix} .$$

It's not entirely clear what is to be done for boxer C, after all, boxer C doesn't lose to anyone. Would we allow boxer C to keep all their water? If we do so, then this is equivalent to saying that boxer C defeated themselves, does this really make sense?

It's useful here to imagine this matrix representing the probabilities of a Markov chain. We know A loses to B and barring any further data, we can conclude that A has a 100% chance of losing to B. Similarly B loses to C 100% of the times they fought. What about C though? Since C has suffered no losses, technically everyone has an equal chance to defeat C, including himself! After all, A, B **and** C all have defeated C an equal number of times. With this in mind, we can fill out this last row:

$$\begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1/3 & 1/3 & 1/3 \end{bmatrix}.$$

How do we rank our boxers then? Well, just as with our water sharing friends Abby, Ben and Chen, we have a value sharing scheme represented by a stochastic matrix. This matrix has an eigenvector with eigenvalue 1, representing a stable state for the amount of water each person approaches after reiterating this value sharing process. Since we now have a water sharing scheme for our 3 boxers, we can solve for this state in a similar way:

$$\begin{bmatrix} a & b & c \end{bmatrix} \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1/3 & 1/3 & 1/3 \end{bmatrix} = \begin{bmatrix} a & b & c \end{bmatrix}, \text{ and thus:}$$

$$\frac{1}{3}c = a,$$

$$a + \frac{1}{3}c = b, \text{ and}$$

$$b + \frac{1}{3}c = c.$$

which has solution $a[1, 2, 3]$. Note that for any positive value we choose to give a, we always have that b will have greater value, and c greater value still. Thus we obtain a ranking of these boxers that is sensible.

For a more in-depth example, suppose there are 5 boxers, A, B, C, D, E. Boxer A loses to B who then loses to C. Then E fights B to a draw, and D also fights A and wins. Afterwards E and B have a rematch and E wins; this is represented in Figure 3.

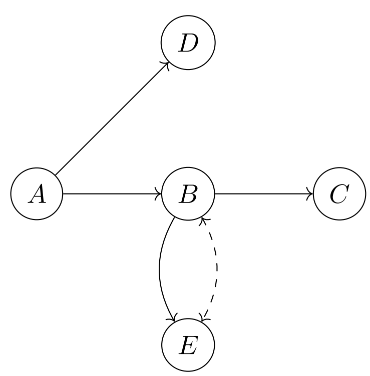

Figure 3 A more complex boxing network

What would be the respective ranking of these boxers? Would D be ranked the same as C as they are both undefeated? How does E compare to everyone?

Here we address some of our increased complexity. To begin with a draw should not be ignored by our scheme, fighting a champion fighter to a draw implies good things about a fighter. In our scheme, we will count each draw as half a loss for each fighter. Doing so gives us the following raw data matrix:

$$R = \begin{bmatrix} 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 3/2 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 1/2 & 0 & 0 & 0 \end{bmatrix}.$$

This is certainly **not** a stochastic matrix. It is simply a record of who lost to whom, and how many "times". We also have 0's in lieu of the question marks ? earlier in this section, since this is not meant to represent a stochastic matrix, and C, D are undefeated. Here we note that since B tied E once and lost to E once, the B, E entry is $1/2 + 1 = 3/2$, and similarly the E, B entry is $1/2$.

Our first step here is to normalize the results of multiple fights. Re-matches are a common occurrence in boxing, especially if the public or one of the parties disputes the results of a previous match. Ultimately, our goal is to compare the boxers to each other, and two boxers who fight against each other multiple times shouldn't have their mutual comparison have an inordinate impact on their ranking. In our case, boxers B and E fought twice with B achieving 1/2 a victory, and E had 3/2's victories. To normalize this, we take the total number of fights and divide the number of victories by that total. So B would be understood to have $(1/2)/2 = 1/4$ victories over E and E would have a $(3/2)/2 = 3/4$ victories over B. Thus our normalized matrix would be:

$$N = \begin{bmatrix} 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 3/4 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 1/4 & 0 & 0 & 0 \end{bmatrix}.$$

The next step is to stochasticize this matrix. For a boxer with ties and losses, we simply divide their row by their row sum, guaranteeing a row sum of 1. For boxers with only victories, they would have 0 rows, and as no one has defeated them, everyone has an equal chance of defeating them, and so each entry of their row would be 1/5.

$$S = \begin{bmatrix} 0 & 1/2 & 0 & 1/2 & 0 \\ 0 & 0 & 4/7 & 0 & 3/7 \\ 1/5 & 1/5 & 1/5 & 1/5 & 1/5 \\ 1/5 & 1/5 & 1/5 & 1/5 & 1/5 \\ 0 & 1 & 0 & 0 & 0 \end{bmatrix}.$$

This matrix makes intuitive sense. Since A loses to C and D they split their value. As C and E clearly defeated B but E also suffered a draw, C is able to take more value from B than E is. Finally, E only suffers a draw to B, but if E were to lose to anyone, it would be B. When we find the stable state, we obtain

$$t\,[1 \quad 4.375 \quad 3.5 \quad 1.5 \quad 2.875]$$

where t is a real number. To make sense of this ranking, it is clear that A would be ranked the lowest. Although C and D are both undefeated, C has beaten a boxer with better performance than the one D defeated so C has a higher ranking than D.

Although both C and E bested B, they've only fought 1 fighter, and B has a win and a partial victory against 2 fighters, allowing more value to flow to B. If E or C wishes to increase their ranking, they would need to fight boxers other than B.

A Possible Complication Suppose we once again had boxers A, B, C, D, E, where A loses to B who loses to C who loses to A. Moreover, D and E fight and draw. How would we rank these boxers?

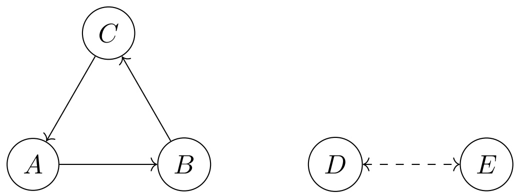

Figure 4 A disconnected boxing network

It's clear here that no boxer seems particularly dominant. In fact it wouldn't be a stretch to say that all the boxers should be ranked equally. When we compute the raw data, normalized and stochastic matrices, we obtain:

$$R, N = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \frac{1}{2} \\ 0 & 0 & 0 & \frac{1}{2} & 0 \end{bmatrix}$$

$$S = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 \end{bmatrix}$$

However, when we try to find our stable states, we notice the vector $[3 \quad 3 \quad 3 \quad 1 \quad 1]$ is a stable state for this system. But $[1 \quad 1 \quad 1 \quad 100 \quad 100]$ is also a stable state. What's going on?

When we consider Figure 4, we can see that there are basically two completely separated networks here. Each network could be dealt with on its own, and within those networks, it's easy to see everyone has equal value. However, as there are no fights between fighters A, B, C and D, E, there is no way to compare the value of those 2 groups. The fighters in the first group could be much better or much worse than the fighters in the second, but without a match or comparison, there's no way to tell. This is reflected in our matrix S. We see that S is a block matrix, where $s\,[1 \quad 1 \quad 1 \quad 0 \quad 0]\,, t\,[0 \quad 0 \quad 0 \quad 1 \quad 1]$ and thus $[s \quad s \quad s \quad t \quad t]$ will be an eigenvector with eigenvalue 1 for any real numbers s and t:

$$S = \left[\begin{array}{ccc|cc} 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 \end{array}\right].$$

However, in this situation, it is still necessary to be able to produce a ranking.

Our solution here is to introduce J, the all 1's matrix. It is clear that if J is a $n \times n$ matrix, then $\frac{1}{n}J$ is a stochastic matrix. Then given any $p, 0 \leq p \leq 1$, we observe that

$$M = (1-p)S + p(\frac{1}{n}J)$$

is stochastic as well. Moreover, if p is positive, then every entry of M will also be positive. This value p is called the *damping factor*, [2].

Why do we create this M matrix? If p is sufficiently small. then $\frac{p}{n}J$ is unlikely to change the relative rankings that would have being established by S. More importantly, this allows us to invoke the following theorem, which only applies to matrices with positive, rather than non-negative, values.

Theorem 1 (Perron-Frobenius Theorem [8]). *If M is a row stochastic matrix where each entry is positive then:*

1. *1 is an eigenvalue of multiplicity one.*
2. *1 is the largest eigenvalue: all the other eigenvalues have an absolute value smaller than 1.*
3. *The eigenvectors corresponding to the eigenvalue 1 have either only positive entries or only negative entries. In particular, for the eigenvalue 1, there exists a unique eigenvector with the sum of its entries equal to 1.*

Items 1 & 3 are of particular importance, it guarantees the existence of a unique ranking. A value of $p = 0.15$ is typical for many implementations of this program. In our case, if $p = 0.15$:

$$M = (1-0.15)S + (0.15)\frac{1}{5}J = \begin{bmatrix} 0.03 & 0.88 & 0.03 & 0.03 & 0.03 \\ 0.03 & 0.03 & 0.88 & 0.03 & 0.03 \\ 0.88 & 0.03 & 0.03 & 0.03 & 0.03 \\ 0.03 & 0.03 & 0.03 & 0.03 & 0.88 \\ 0.03 & 0.03 & 0.03 & 0.88 & 0.03 \end{bmatrix},$$

which has stable-state $t\begin{bmatrix}1 & 1 & 1 & 1 & 1\end{bmatrix}$ where t is a real number, as we suspected.

The Final Algorithm Putting everything together, our final algorithm has the following steps.

1. Start with a raw data matrix R where $R_{i,j}$ is the number of times boxer i losses to boxer j, with draws counting as a half-loss.
2. Create a normalized data matrix N, where $N_{i,j} = R_{i,j}/(R_{i,j} + R_{j,i})$ if $R_{i,j} + R_{j,i} \neq 0$, and $N_{i,j} = 0$ otherwise.
3. Create a stochastic data matrix S, where $S_{i,j} = N_{i,j}/(\sum_{k=1}^{n} N_{i,k})$ if $\sum_{k=1}^{n} N_{i,k} \neq 0$ and $N_{i,j} = 1/n$ otherwise.
4. Choose a damping factor p, $0 < p < 1$ and create $M = (1-p)S + p\frac{1}{n}J$.
5. Find a positive left eigenvector of M that has eigenvalue 1.

This roughly describes the implementation of PageRank [2, 10]. When used to rank web pages, web pages values are determined by incoming links and the respective value of those pages. Some modification was required to address the issues of multiple fights and draws which are not relevant to ranking websites. More on using PageRank in sports can be found [9, 10, 11].

So, was Mike Tyson right or not?

Evander Holyfield's best-remembered fights took place in the 1990's, and so looking at a list of the top 20 boxers of the 1990's http://heavyweightaction.com/Decade%201990s.html [15], we are given a list:

1. Lennox Lewis
2. Riddick Bowe
3. Evander Holyfield
4. Mike Tyson
5. George Foreman
6. Ike Ibeabuchi
7. Michael Moorer
8. Buster Douglas
9. Oliver McCall
10. Larry Holmes
11. Ray Mercer
12. David Tua
13. Tommy Morrison
14. Frank Bruno
15. Razor Ruddock
16. Michael Grant
17. Andrew Golata
18. Tony Tucker
19. Tim Witherspoon
20. Henry Akinwande

The author gives their reasons for this particular ranking. However, we will determine our own ranking with our algorithm. Looking at fights that took place between these fighters in the 1990's as recorded by http://boxrec.com/ [14], we obtain the following network for fights (Figure 5):

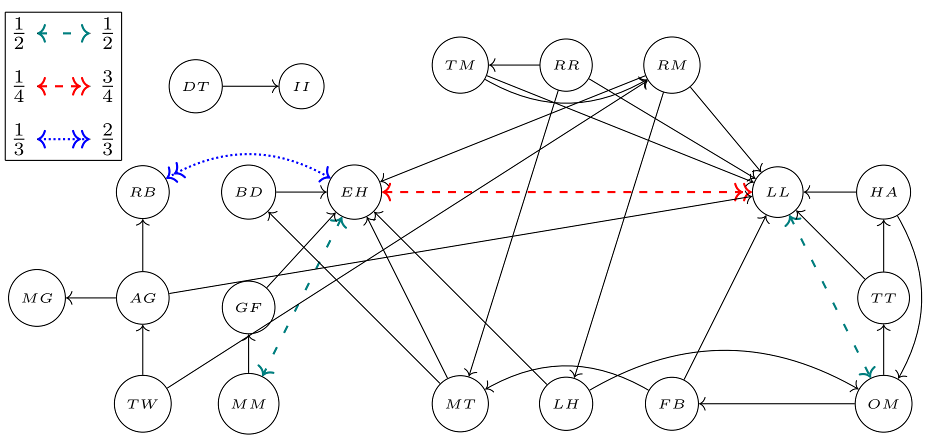

Figure 5 The boxing network for the top 20 boxers of the 90's

Then, allowing $p = 0.15$, the stable state for $M = (1 - p)S + p\frac{1}{20}J$ is approximately:

$$t[1 \;\; 0.394 \;\; 1.104 \;\; 0.188 \;\; 0.215 \;\; 0.089 \;\; 0.293 \;\; 0.128 \;\; 0.726 \;\; 0.086,$$
$$0.134 \;\; 0.048 \;\; 0.153 \;\; 0.295 \;\; 0.048 \;\; 0.068 \;\; 0.069 \;\; 0.295 \;\; 0.048 \;\; 0.174]$$

So by our method, the ranking is:

1. Evander Holyfield
2. Lennox Lewis
3. Oliver McCall
4. Riddick Bowe
5. Frank Bruno
5. Tony Tucker
7. Michael Moorer
8. George Foreman
9. Mike Tyson
10. Henry Akinwande
11. Tommy Morrison
12. Ray Mercer
13. Buster Douglas
14. Ike Ibeabuchi
15. Larry Holmes
16. Andrew Golata
17. Michael Grant
18. Razor Ruddock
18. David Tua
18. Tim Witherspoon

As Mike Tyson intimated, Holyfield has an impressive list of victories, that includes Mike Tyson himself. Moreover, no fighter who bested Holyfield was able to do so unilaterally, the only fighters who beat Holyfield were Riddick Bowe, Michael Moorer, and Lennox Lewis. Holyfield lost to Bowe in 1992, defeated Bowe in 1993, and lost to him a second time in 1995. He lost to Moorer in 1997 and avenged his loss in 1997. His initial fight against Lennox was a draw in 1999, before being defeated by Lennox later that year [**14**]. So even though these fighters defeated Holyfield, the draws and comeback victories make those only partial victories and allows for some of their value to flow back to Holyfield. Thus, by our analysis, Mike Tyson was right, and Holyfield's preeminence should not be ignored.

Comparison to the Colley Method

PageRank is not the only matrix based ranking algorithm to be applied to sports. One of these methods is the Colley Method developed by Astrophysicist Dr. Wessley Colley [**6**, **5**, **11**]. The general principles of this method are as follows:

1. The algorithm is *bias-free* meaning it only uses win/loss/tie information to determine the rankings. It does not consider ad hoc adjustments likes leagues, points scored, or rounds won.
2. It follows a principle of *conservation* in that each team/athlete starts with a value of 1/2. Each time a team/athlete defeats another, part of the losers value is transferred to the winner. The stronger the opponent defeated, the more value is transferred to the winner.

As we can see, this method shares a lot of the same principles of PageRank. The ranking vector is found by solving for $\mathbf{r}$, for $C\mathbf{r} = \mathbf{b}$ where

$$C_{i,j} = \begin{cases} 2 + t_i & i = j \\ -n_{i,j} & i \neq j \end{cases}, \qquad b_i = 1 + \frac{w_i - l_i}{2},$$

t_i is the total games by team/athlete i, w_i, l_i are their wins and losses, and $n_{i,j}$ represents the number of matches between team/athletes i and j.

Perhaps the most significant difference between the Colley method and PageRank is that while Colley transfers some value from the loser to the victor in each defeat, PageRank uses the win/loss/tie information to establish a hierarchy between the team/athletes. PageRank does not count the number of matches between two teams/athletes, only the normalized outcomes.

For example, suppose two unranked boxers, A and B, enter a match, and A defeats B. Consider the Colley method to find a ranking vector for these boxers. Prior to this match, we would have valued each boxer at $1/2$. Then after the match we would have:

$$\begin{aligned}\begin{bmatrix}3 & -1\\ -1 & 3\end{bmatrix}\begin{bmatrix}r_A\\ r_B\end{bmatrix} &= \begin{bmatrix}3/2\\ 1/2\end{bmatrix}\\ \begin{bmatrix}r_A\\ r_B\end{bmatrix} &= \begin{bmatrix}5/8\\ 3/8\end{bmatrix}\end{aligned}$$

A being the victor, has taken some of B's value for themselves. If they had a rematch, and A won again, we would obtain a new ranking vector:

$$\begin{aligned}\begin{bmatrix}4 & -2\\ -2 & 4\end{bmatrix}\begin{bmatrix}r_A\\ r_B\end{bmatrix} &= \begin{bmatrix}2\\ 0\end{bmatrix}\\ \begin{bmatrix}r_A\\ r_B\end{bmatrix} &= \begin{bmatrix}2/3\\ 1/3\end{bmatrix}\end{aligned}$$

Here, we would have even more of B's value transferred to A.

On the other hand, PageRank would not distinguish between the first and second wins. The first match establishes A as the dominant fighter over B. The rematch then merely confirms this dominance, and in the eyes of PageRank, does nothing to change the relationship between A and B. In either case, we have:

$$\begin{aligned}\begin{bmatrix}a & b\end{bmatrix}\begin{bmatrix}1/2 & 1/2\\ 1 & 0\end{bmatrix} &= \begin{bmatrix}a & b\end{bmatrix}\\ \begin{bmatrix}a & b\end{bmatrix} &= t\begin{bmatrix}1 & 1/2\end{bmatrix},\end{aligned}$$

for t a real number. If we left $t = 2/3$, we achieve the same vector as the Colley method after 2 back to back wins for A.

This leads to a second difference between these methods. As noted in [4], the PageRank algorithm is much more sensitive to upset victories. This sensitivity is also related to the fact that PageRank doesn't distinguish between 1 or 5 or 100 victories between 2 opponents, whereas Colley transfers value from victors to losers in a more conservative manner. Moreover, since the damping factor chosen for the PageRank algorithm is somewhat arbitrary, we should note that the ranking vector produced by this process is for purposes of creating an ordinal ranking, and that the exact values of the vector are themselves of little significance.

The ranking of the 90's Heavyweights by the Colley Method would be:

1. Lennox Lewis
2. Riddick Bowe
3. Evander Holyfield
4. Oliver McCall
5. Buster Douglas
6. Tommy Morrison
7. Michael Grant
8. Ray Mercer
9. Ike Ibeabuchi
10. Mike Tyson
11. Frank Bruno
12. Henry Akinwande
13. Tony Tucker
13. Larry Holmes
15. Michael Moorer
16. George Foreman
17. Andrew Golata
18. David Tua
19. Razor Ruddock
20. Tim Witherspoon

Further Questions

Here we pose some further questions.

Although our ranking system showed that Evander Holyfield was ranked higher than Mike Tyson, and in fact Holyfield beat Tyson twice, very few people would say that Holyfield was a better fighter than Tyson. Many would argue that Holyfield fought Tyson at the tail end of Tyson's career, while Tyson was showing signs of age [3]. Certainly, during the 80's where Tyson was considered at his best, every fight resulted in victory, which can't be said about Holyfield in the 90's.

Problem. Can one find a way to weight the value or importance of fights, given where each fighter is in their career?

Problem. Is there a way to fairly compare boxers who may not have fought in the same time frame, or in overlapping time frames?

Our analysis here is close to binary, wins and losses with ties splitting the difference. However, commentators and boxing enthusiasts not only place value on who wins, but how they won. One of Mike Tyson's most infamous victories was against Michael Spinks in 1988, knocking out the previously undefeated champion in 91 seconds [7].

Problem. Is there a way to weight the value of victories, knockouts versus decisions, factoring in point-spreads and round on knockouts (if they occur) to have a less binary analysis? How would this change to our algorithm compare to the Massey method of ranking? How would winning because your opponent bit off a part of your ear factor into this?

Acknowledgment

The authors would like to thank the Editor Dr. Michael Jones and the Referees for their suggestions which improved the paper.

Tien Chih would like to thank his son Maxton Dun Chih, born April 19th, 2016. Without the many sleepless nights following his birth, the author would not have watched the appropriate documentaries on both boxing [3] and mathematics [13] which inspired this paper.

This paper is dedicated to the memory of Muhammad Ali (January 17, 1942 - June 3, 2016). He was a champion of not only boxing but of justice and human rights as well.

REFERENCES

1. Adrian Bondy, U.S.R. Murphy, *Graph Theory*, Springer, 2008.
2. Kurt Bryan, Tanya Leise, *The $25,000,000,000 eigenvector. The linear algebra behind Google*, SIAM Review, 48 (3), 569-81. 2006.
3. *Chasing Tyson*, dir. Steven Cantor, ESPN 30-for-30, 2015.
4. Timothy P. Chartier, Erich Kreutzer, Amy N. Langville , Kathyrn E. Pedings, *Sensitivity and Stability of Ranking Vectors*, SIAM Journal on Scientific Computing, Vol. 33, No. 3, pp. 1077–1102, 2011.
5. Timothy P. Chartier, Erich Kreutzer, Amy N. Langville , Kathyrn E. Pedings, *Bracketology: How can math help?*, https://doi.org/10.5948/UPO9781614442004.006, 2010.
6. Wesley N. Colley, *Colley's Bias Free College Football Ranking Method: The Colley Matrix Explained*, https://www.colleyrankings.com/matrate.pdf, 2002.
7. William Gildea, *Tyson-Spinks Is More Than a Fight, It's History*, LA Times, 1988.
8. Chris Godsil, Gordon Royle, *Algebraic Graph Theory*, Springer, 2001.
9. Anjela Y. Govan, Carl D. Meyer, Russell Albright, *Generalizing Google's PageRank to Rank National Football League Teams*, SAS Global Forum 2008: Data Mining and Predictive Modeling, 2008.
10. Langville, Amy N. and Meyer, Carl D. *Google's PageRank and Beyond: The Science of Search Engine Rankings*, Princeton University Press, 2006.
11. Langville, Amy N. and Meyer, Carl D. *Who's # 1?: The Science of Rating and Ranking*, Princeton University Press, 2012.
12. David C. Lay, Steven R. Lay, Judi J. McDonald, *Linear Algebra and it's Applciations*, Pearson, 2016.
13. *The Secret Rules of Modern Living: Algorithms*, Marcus Du Sautoy, BBC Four, 2015.
14. *The Boxing Database*, http://boxrec.com/, 2016.
15. *Decade-by-Decade Rankings: 1990-1999*, http://heavyweightaction.com/Decade%201990s.html, 2016.

Summary In the world of boxing, fighters are judged by the quality of their opponents. This way of determining value is reminiscent of Google's PageRank algorithm, which assigns values to websites based on the value of sites linked to it. We describe the mathematics behind PageRank and give a scheme to adapt it to boxing. We illustrate how this algorithm would behave in several scenarios. Then 20 top heavyweight boxers from the 1990's are selected and ranked using this method. A comparison is made to the Colley method of ranking. We describe some possible extensions of this method.

TIEN CHIH (MR Author ID: 1223485) earned a B.A. from the University of Hawaii at Hilo, and an M.A. & Ph.D. from the University of Montana. He enjoys his current appointment as Assistant Professor at Montana State University-Billings where he works on integrating high impact practices into his courses, mentoring undergraduates, running a Math Circle and conducting research on Categorical Graph Theory. He enjoys cooking for his wife and two sons.

DEMITRI PLESSAS (MR Author ID: 1223509) earned a B.S. from Montana Tech, and an M.A. & Ph.D. from the University of Montana. He loves blending Mathematics and Statistics to tackle unique modelling challenges as a Lead Data Scientist at Zulily. As the proud father of a son and a daughter, his greatest joy in life is seeing the world through their eyes.