

O'Sullivan, Hendry, Davies, ...?

A complex networks approach to ranking professional Snooker players

Joseph O'Brien

4-Dec-2020 | 7th Annual Irish SIAM Student Chapter Conference



UNIVERSITY of LIMERICK

OLLSCOIL LUIMNIGH



joseph.obrien@ul.ie



@obrienj_



Fondúireacht Eolaíochta Éireann
Dá bhfuil romhainn

Science Foundation Ireland
For what's next

Outline of Today's Talk

- Ranking within complex systems
- **Snooker** and its historical ranking approach
- Networked representation
- **Mathematical framework** for rankings + **results**
- **Conclusions**

Outline of Today's Talk

- Ranking within complex systems
- **Snooker** and its historical ranking approach
- Networked representation
- **Mathematical framework** for rankings + **results**
- **Conclusions**

Outline of Today's Talk

- Ranking within complex systems
- **Snooker** and its historical ranking approach
- Networked representation
- **Mathematical framework** for rankings + **results**
- **Conclusions**

Outline of Today's Talk

- Ranking within complex systems
- **Snooker** and its historical ranking approach
- Networked representation
- **Mathematical framework** for rankings + **results**
- **Conclusions**

Outline of Today's Talk

- Ranking within complex systems
- **Snooker** and its historical ranking approach
- Networked representation
- **Mathematical framework** for rankings + **results**
- **Conclusions**

Motivation

- How to rank elements in a complex system?
- Ideally using a quantitative approach
- Requires detailed data of interactions
- Inherent networked representation

Motivation

- How to rank elements in a complex system?
- Ideally using a quantitative approach
- Requires detailed data of interactions
- Inherent networked representation

Motivation

- How to rank elements in a complex system?
- Ideally using a quantitative approach
- Requires detailed data of interactions
- Inherent networked representation

Motivation

- How to rank elements in a complex system?
- Ideally using a quantitative approach
- Requires detailed data of interactions
- Inherent networked representation

Motivation

- How to rank elements in a complex system?
- Ideally using a quantitative approach
- Requires detailed data of interactions
- Inherent networked representation

The PageRank Citation Ranking: Bringing Order to the Web

January 29, 1998

Abstract

The importance of a Web page is an inherently subjective matter, which depends on the readers interests, knowledge and attitudes. But there is still much that can be said objectively about the relative importance of Web pages. This paper describes PageRank, a method for rating Web pages objectively and mechanically, effectively measuring the human interest and attention devoted to them.

Motivation

- How to rank elements in a complex system?
- Ideally using a quantitative approach
- Requires detailed data of interactions
- Inherent networked representation

The PageRank Citation Ranking: Bringing Order to the Web

January 29, 1998

Abstract

The importance of a Web page is an inherently subjective matter, which depends on the readers interests, knowledge and attitudes. But there is still much that can be said objectively about the relative importance of Web pages. This paper describes PageRank, a method for rating Web pages objectively and mechanically, effectively measuring the human interest and attention devoted to them.

The Anatomy of a Large-Scale Hypertextual Web Search Engine

Sergey Brin and Lawrence Page

*Computer Science Department,
Stanford University, Stanford, CA 94305, USA*
sergey@cs.stanford.edu and page@cs.stanford.edu

Abstract

In this paper, we present Google, a prototype of a large-scale search engine which makes heavy use of the structure present in hypertext. Google is designed to crawl and index the Web efficiently and produce much more satisfying search results than existing systems. The prototype with a full text and hyperlink database of at least 24 million pages is available at <http://google.stanford.edu/>

Motivation

- How to rank elements in a complex system?
- Ideally using a quantitative approach
- Requires detailed data of interactions
- Inherent networked representation

The PageRank Citation Ranking: Bringing Order to the Web

January 29, 1998

Abstract

The importance of a Web page is an inherently subjective matter, which depends on the readers interests, knowledge and attitudes. But there is still much that can be said objectively about the relative importance of Web pages. This paper describes PageRank, a method for rating Web pages objectively and mechanically, effectively measuring the human interest and attention devoted to them.

The Anatomy of a Large-Scale Hypertextual Web Search Engine

Sergey Brin and Lawrence Page

*Computer Science Department,
Stanford University, Stanford, CA 94305, USA
sergey@cs.stanford.edu and page@cs.stanford.edu*

Abstract

In this paper, we present Google, a prototype of a large-scale search engine which makes heavy use of the structure present in hypertext. Google is designed to crawl and index the Web efficiently and produce much more satisfying search results than existing systems. The prototype with a full text and hyperlink database of at least 24 million pages is available at <http://google.stanford.edu/>



So... snooker?



So... snooker?



So... snooker?



So... snooker?



So... snooker?



So... snooker?



So... snooker?

- Each year there are a number of tournaments that the player's compete in
- Classically, ranked on a point basis obtained from results
- From 2013 instead based upon total prize money

So... snooker?

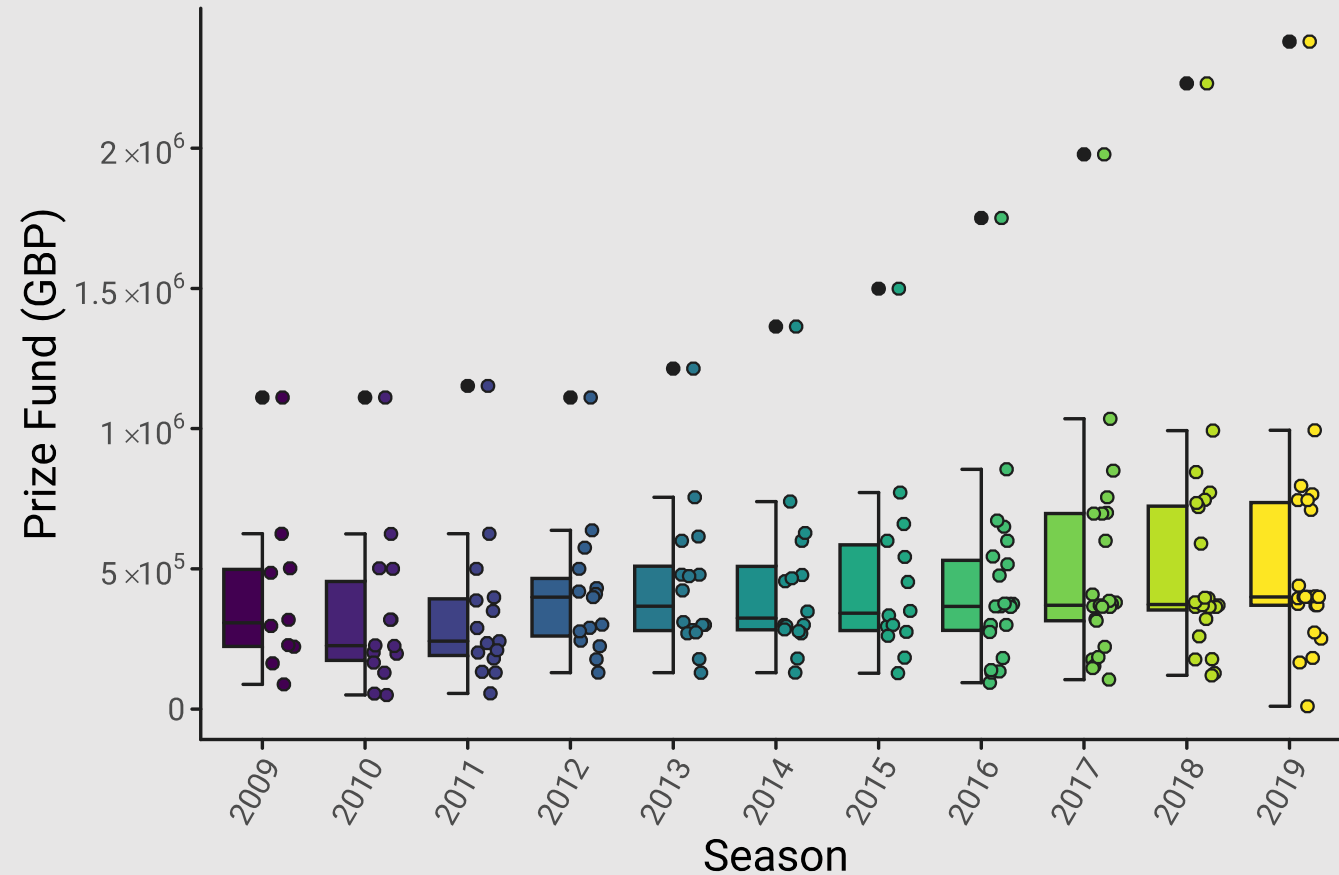
- Each year there are a number of tournaments that the player's compete in
- Classically, ranked on a point basis obtained from results
- From 2013 instead based upon total prize money

So... snooker?

- Each year there are a number of tournaments that the player's compete in
- Classically, ranked on a point basis obtained from results
- From 2013 instead based upon total prize money

So... snooker?

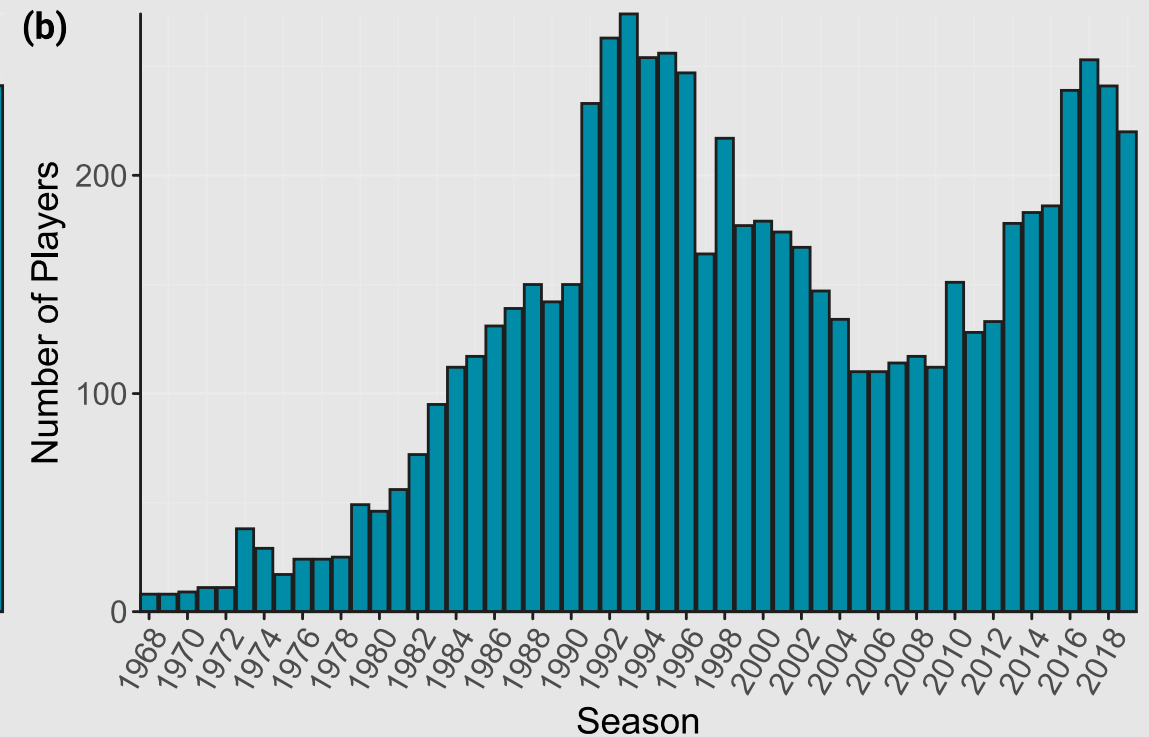
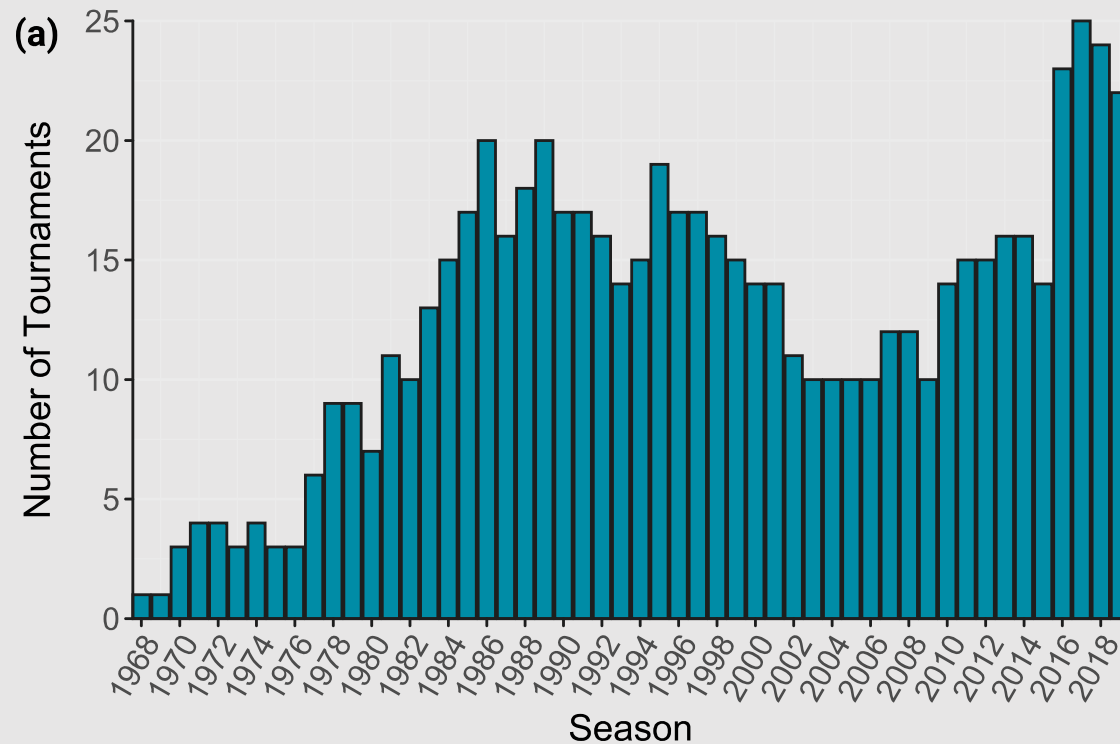
- Each year there are a number of tournaments that the player's compete in
- Classically, ranked on a point basis obtained from results
- From 2013 instead based upon total prize money



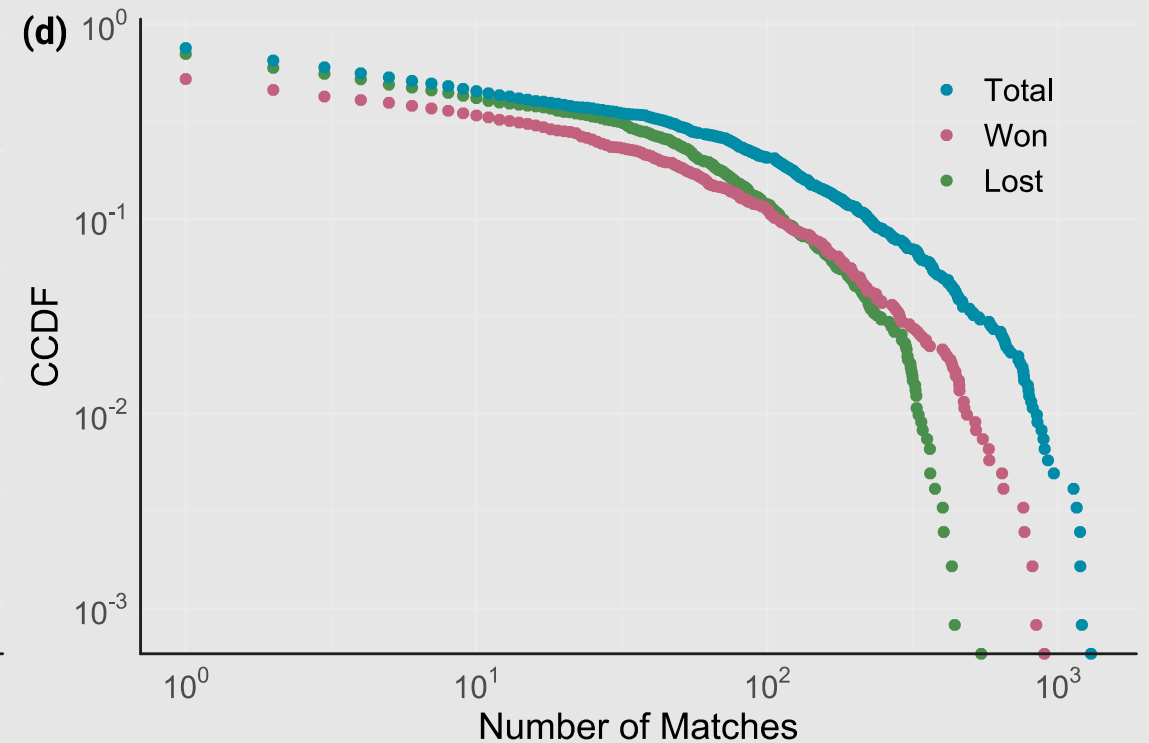
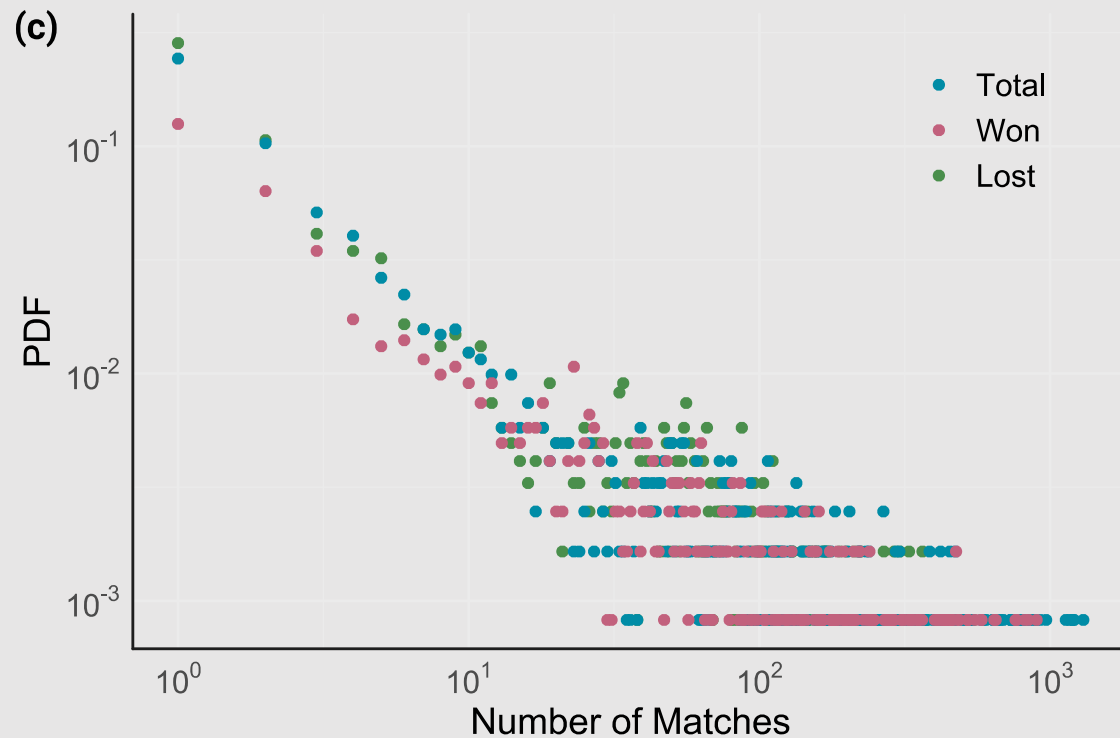
Data - collection

- Historical results from 1908-2020 stored on **cuetracker.net**
- Entire website scraped we focus on **professional era from 1968**
 - 657 tournaments
 - 1221 players
 - > 47,000 matches
 - Official rankings from 1975 onwards

Data – initial analysis

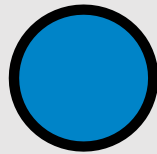
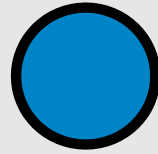


Data – initial analysis

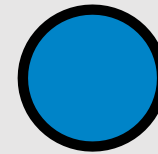


Snooker's complex network

Ronnie O'Sullivan

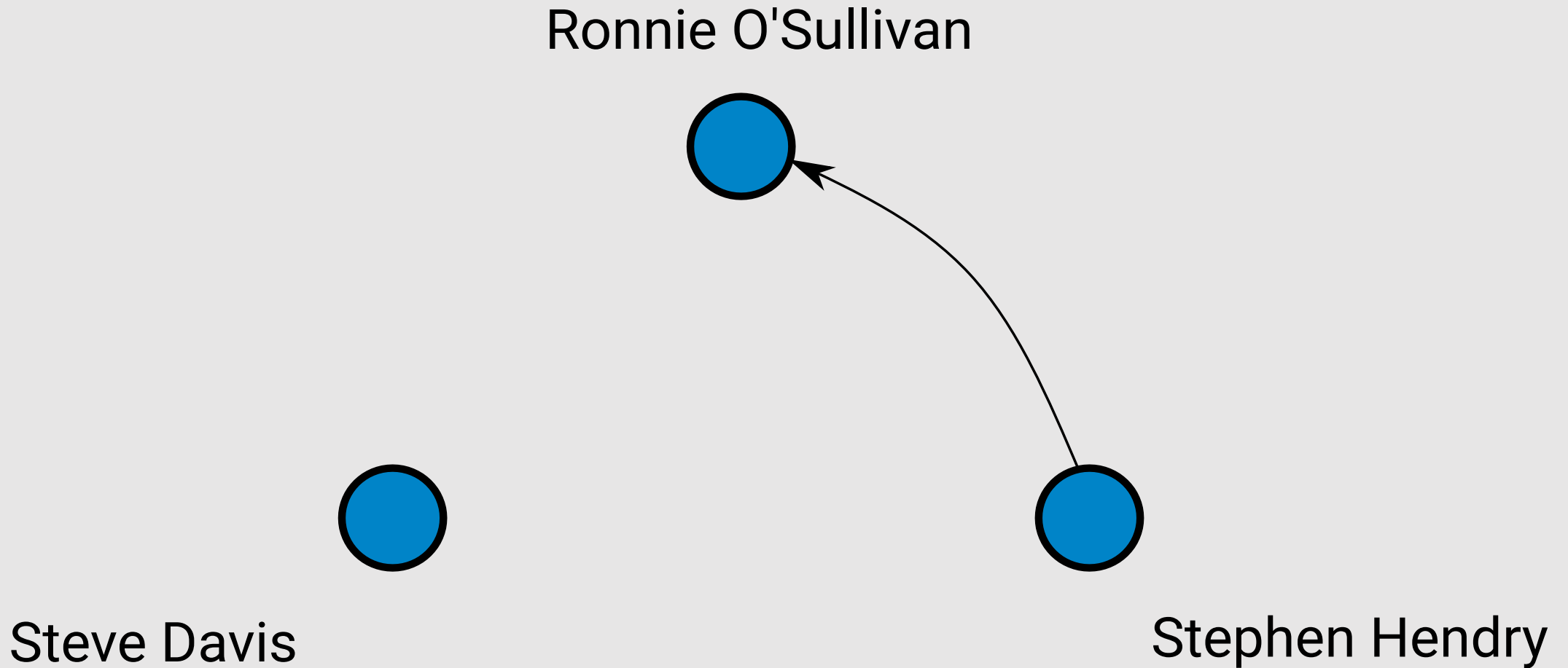


Steve Davis

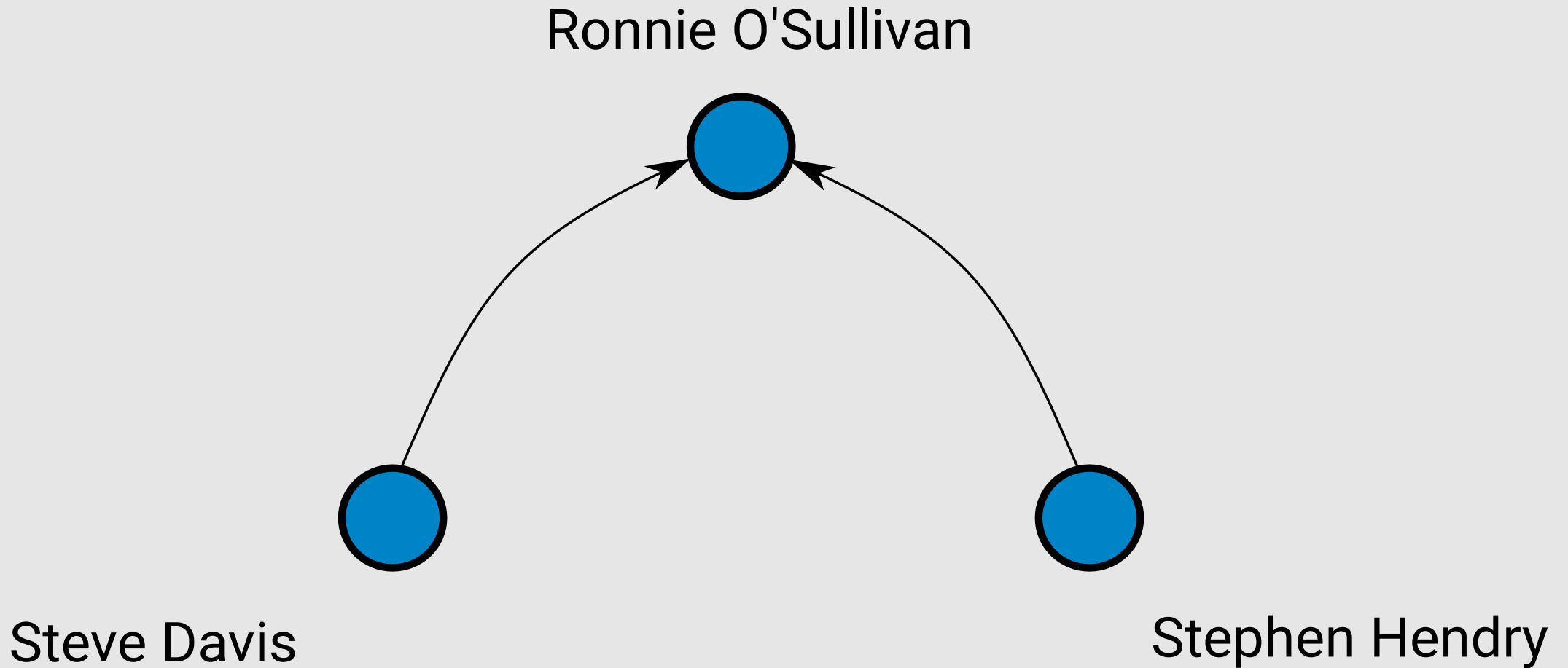


Stephen Hendry

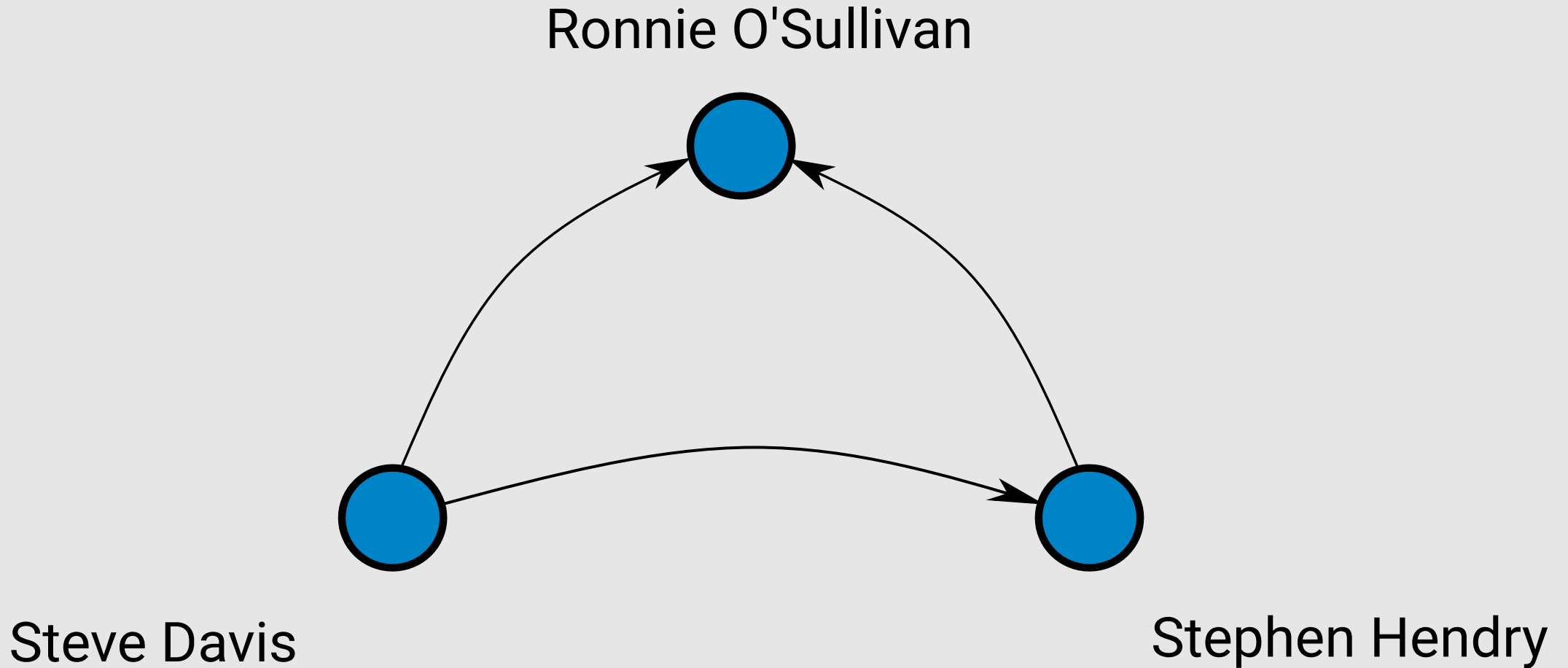
Snooker's complex network



Snooker's complex network

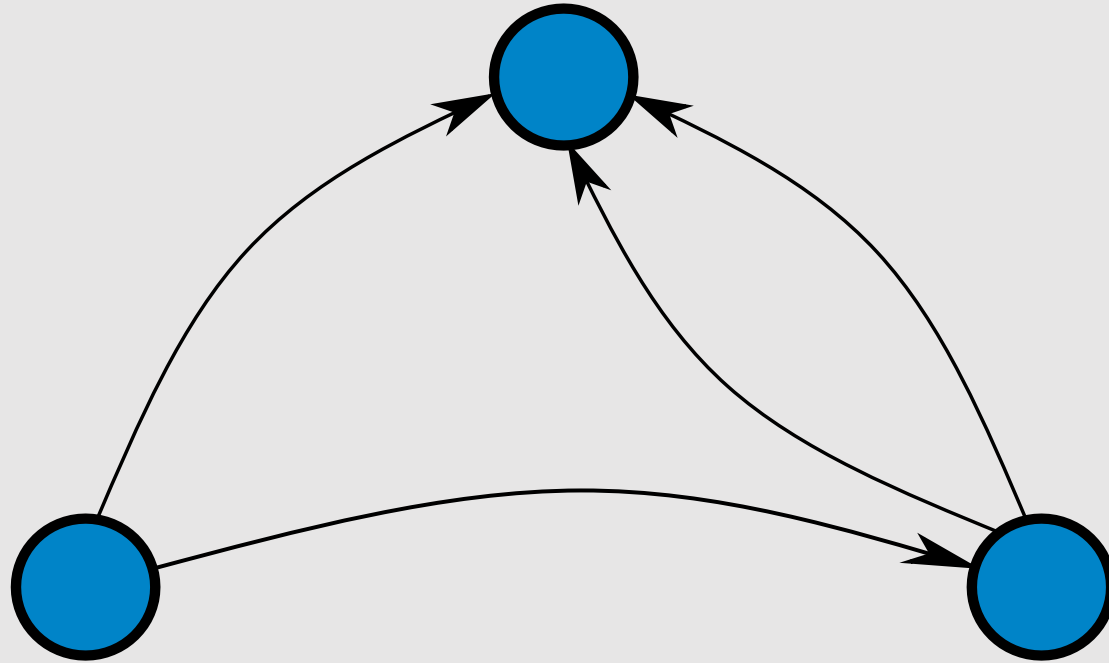


Snooker's complex network



Snooker's complex network

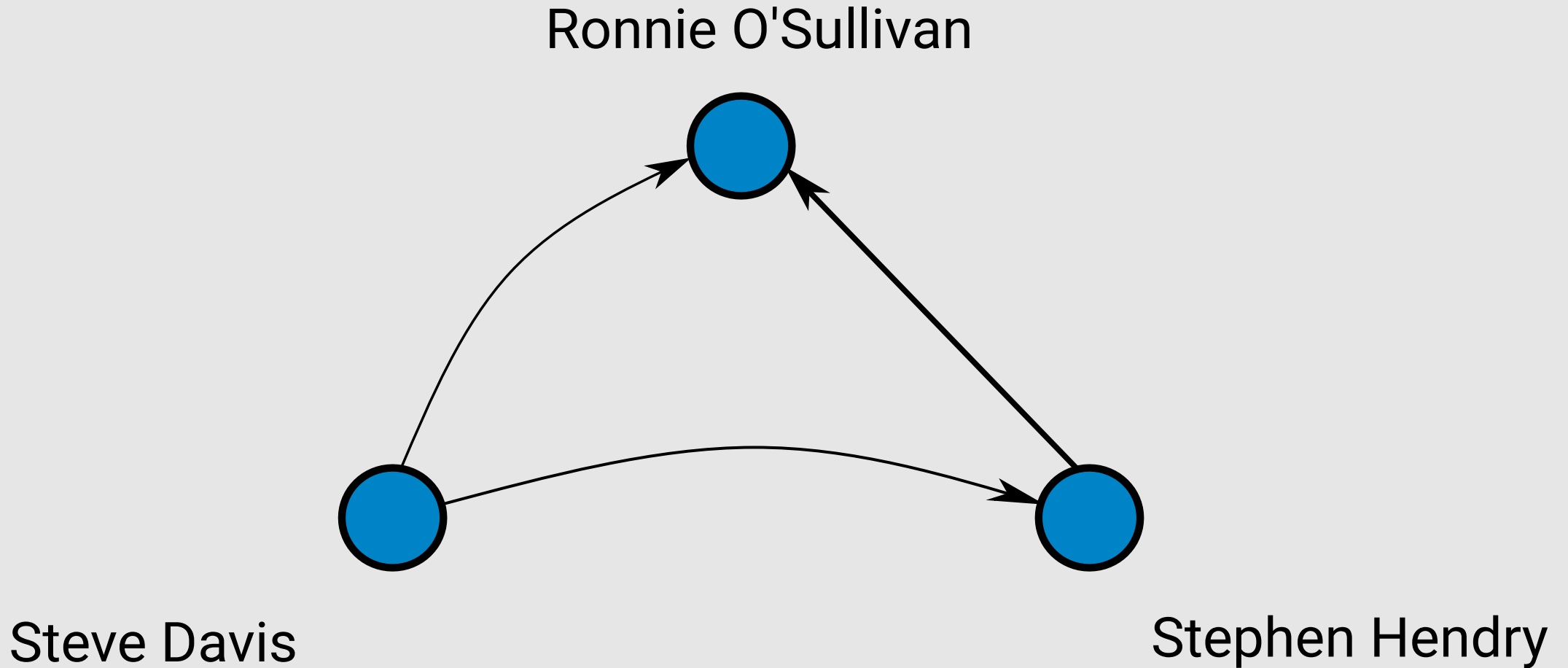
Ronnie O'Sullivan



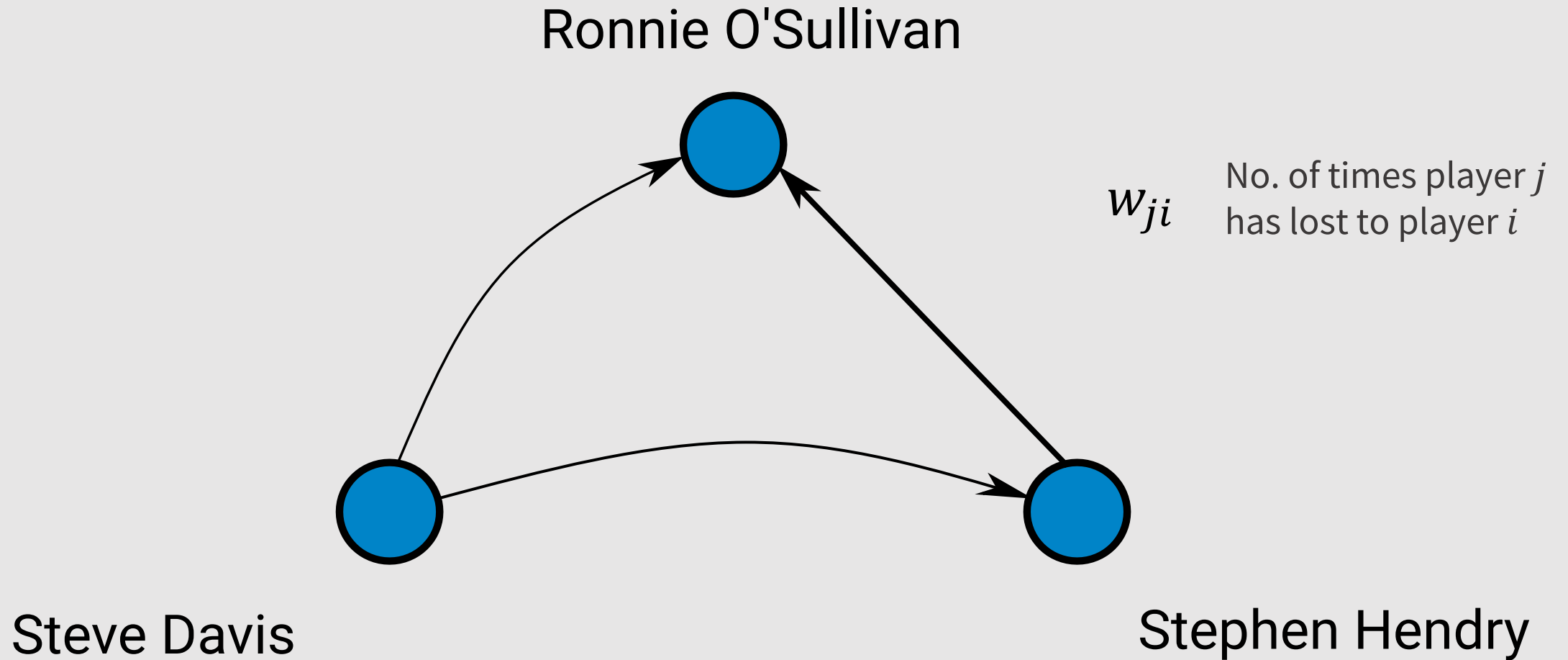
Steve Davis

Stephen Hendry

Snooker's complex network



Snooker's complex network



Mathematical framework

- Prestige score for player i , P_i , obtained from N coupled equations

$$P_i = (1 - q) \sum_j P_j \frac{w_{ji}}{k_j^{out}} + \frac{q}{N} + \frac{1 - q}{N} \sum_j P_j \delta(k_j^{out})$$

Mathematical framework

- Prestige score for player i , P_i , obtained from N coupled equations

$$P_i = (1 - q) \sum_j P_j \frac{w_{ji}}{k_j^{out}} + \frac{q}{N} + \frac{1 - q}{N} \sum_j P_j \delta(k_j^{out})$$

Player i 's
level of
prestige

Mathematical framework

- Prestige score for player i , P_i , obtained from N coupled equations

$$P_i = (1 - q) \sum_j P_j \frac{w_{ji}}{k_j^{out}} + \frac{q}{N} + \frac{1 - q}{N} \sum_j P_j \delta(k_j^{out})$$

Player i 's level of prestige

Damping factor for unif. redistribution of prestige

Mathematical framework

- Prestige score for player i , P_i , obtained from N coupled equations

$$P_i = (1 - q) \sum_j P_j \frac{w_{ji}}{k_j^{out}} + \frac{q}{N} + \frac{1 - q}{N} \sum_j P_j \delta(k_j^{out})$$

Player i 's level of prestige

No. of times player j has lost to player i

Damping factor for unif. redistribution of prestige

Mathematical framework

- Prestige score for player i , P_i , obtained from N coupled equations

$$P_i = (1 - q) \sum_j P_j \frac{w_{ji}}{k_j^{out}} + \frac{q}{N} + \frac{1 - q}{N} \sum_j P_j \delta(k_j^{out})$$

Player i 's level of prestige

No. of times player j has lost to player i

Damping factor for unif. redistribution of prestige

Total no. of time player j has lost

Mathematical framework

- Prestige score for player i , P_i , obtained from N coupled equations

$$P_i = (1 - q) \sum_j P_j \frac{w_{ji}}{k_j^{out}} + \frac{q}{N} + \frac{1 - q}{N} \sum_j P_j \delta(k_j^{out})$$

Diagram illustrating the mathematical framework for the prestige score P_i for player i .

The equation is:

$$P_i = (1 - q) \sum_j P_j \frac{w_{ji}}{k_j^{out}} + \frac{q}{N} + \frac{1 - q}{N} \sum_j P_j \delta(k_j^{out})$$

Annotations:

- P_i : Player i 's level of prestige
- $(1 - q) \sum_j P_j \frac{w_{ji}}{k_j^{out}}$: Transfer of prestige from player j to i
- $\frac{q}{N}$: No. of times player j has lost to player i
- $\frac{1 - q}{N}$: Damping factor for unif. redistribution of prestige
- $\sum_j P_j \delta(k_j^{out})$: Total no. of time player j has lost

Mathematical framework

- Prestige score for player i , P_i , obtained from N coupled equations

$$P_i = (1 - q) \sum_j P_j \frac{w_{ji}}{k_j^{out}} + \frac{q}{N} + \frac{1 - q}{N} \sum_j P_j \delta(k_j^{out})$$

Diagram illustrating the mathematical framework for the prestige score P_i for player i .

The equation is:

$$P_i = (1 - q) \sum_j P_j \frac{w_{ji}}{k_j^{out}} + \frac{q}{N} + \frac{1 - q}{N} \sum_j P_j \delta(k_j^{out})$$

Annotations:






- P_i : Player i 's level of prestige
- $(1 - q) \sum_j P_j \frac{w_{ji}}{k_j^{out}}$: Transfer of prestige from player j to i
- $\frac{q}{N}$: No. of times player j has lost to player i
- $\frac{1 - q}{N}$: Damping factor for unif. redistribution of prestige
- $\delta(k_j^{out})$: Account for cases where player j has never lost
- k_j^{out} : Total no. of time player j has lost

Rankings – all-time

The Top 10 Snooker Players

1968 - 2020

RANK ¹	PLAYER	SCORE	NATION
-------------------	--------	-------	--------

6	Judd Trump	0.0136	
7	Neil Robertson	0.0134	
8	Steve Davis	0.0129	
9	Shaun Murphy	0.0126	
10	Jimmy White	0.0116	

¹ Calculated via a complex networks approach.

Rankings – all-time

The Top 10 Snooker Players

1968 - 2020

RANK ¹	PLAYER	SCORE	NATION
-------------------	--------	-------	--------

5	Mark Selby	0.0149	
6	Judd Trump	0.0136	
7	Neil Robertson	0.0134	
8	Steve Davis	0.0129	
9	Shaun Murphy	0.0126	
10	Jimmy White	0.0116	



¹ Calculated via a complex networks approach.

Rankings – all-time

The Top 10 Snooker Players

1968 - 2020

RANK ¹	PLAYER	SCORE	NATION
-------------------	--------	-------	--------

5	Mark Selby	0.0149	
6	Judd Trump	0.0136	
7	Neil Robertson	0.0134	
8	Steve Davis	0.0129	
9	Shaun Murphy	0.0126	
10	Jimmy White	0.0116	




¹ Calculated via a complex networks approach.

Rankings – all-time

The Top 10 Snooker Players

1968 - 2020

RANK ¹	PLAYER	SCORE	NATION
-------------------	--------	-------	--------

4	Stephen Hendry	0.0164	
5	Mark Selby	0.0149	
6	Judd Trump	0.0136	
7	Neil Robertson	0.0134	
8	Steve Davis	0.0129	
9	Shaun Murphy	0.0126	
10	Jimmy White	0.0116	






¹ Calculated via a complex networks approach.

Rankings – all-time

The Top 10 Snooker Players

1968 - 2020

RANK ¹	PLAYER	SCORE	NATION
-------------------	--------	-------	--------

4	Stephen Hendry	0.0164	
5	Mark Selby	0.0149	
6	Judd Trump	0.0136	
7	Neil Robertson	0.0134	
8	Steve Davis	0.0129	
9	Shaun Murphy	0.0126	
10	Jimmy White	0.0116	









¹ Calculated via a complex networks approach.

Rankings – all-time

The Top 10 Snooker Players

1968 - 2020

RANK ¹	PLAYER	SCORE	NATION
-------------------	--------	-------	--------

3	Mark Williams	0.0169	
4	Stephen Hendry	0.0164	
5	Mark Selby	0.0149	
6	Judd Trump	0.0136	
7	Neil Robertson	0.0134	
8	Steve Davis	0.0129	
9	Shaun Murphy	0.0126	
10	Jimmy White	0.0116	









¹ Calculated via a complex networks approach.

Rankings – all-time

The Top 10 Snooker Players

1968 - 2020

RANK ¹	PLAYER	SCORE	NATION
-------------------	--------	-------	--------










3	Mark Williams	0.0169	
4	Stephen Hendry	0.0164	
5	Mark Selby	0.0149	
6	Judd Trump	0.0136	
7	Neil Robertson	0.0134	
8	Steve Davis	0.0129	
9	Shaun Murphy	0.0126	
10	Jimmy White	0.0116	

¹ Calculated via a complex networks approach.

Rankings – all-time

The Top 10 Snooker Players

1968 - 2020










RANK ¹	PLAYER	SCORE	NATION
2	Ronnie O'Sullivan	0.0201	
3	Mark Williams	0.0169	
4	Stephen Hendry	0.0164	
5	Mark Selby	0.0149	
6	Judd Trump	0.0136	
7	Neil Robertson	0.0134	
8	Steve Davis	0.0129	
9	Shaun Murphy	0.0126	
10	Jimmy White	0.0116	

¹ Calculated via a complex networks approach.

Rankings – all-time

The Top 10 Snooker Players

1968 - 2020











RANK ¹	PLAYER	SCORE	NATION
2	Ronnie O'Sullivan	0.0201	
3	Mark Williams	0.0169	
4	Stephen Hendry	0.0164	
5	Mark Selby	0.0149	
6	Judd Trump	0.0136	
7	Neil Robertson	0.0134	
8	Steve Davis	0.0129	
9	Shaun Murphy	0.0126	
10	Jimmy White	0.0116	

¹ Calculated via a complex networks approach.

Rankings – all-time

The Top 10 Snooker Players

1968 - 2020











RANK ¹	PLAYER	SCORE	NATION
1	John Higgins	0.0204	
2	Ronnie O'Sullivan	0.0201	
3	Mark Williams	0.0169	
4	Stephen Hendry	0.0164	
5	Mark Selby	0.0149	
6	Judd Trump	0.0136	
7	Neil Robertson	0.0134	
8	Steve Davis	0.0129	
9	Shaun Murphy	0.0126	
10	Jimmy White	0.0116	

¹ Calculated via a complex networks approach.

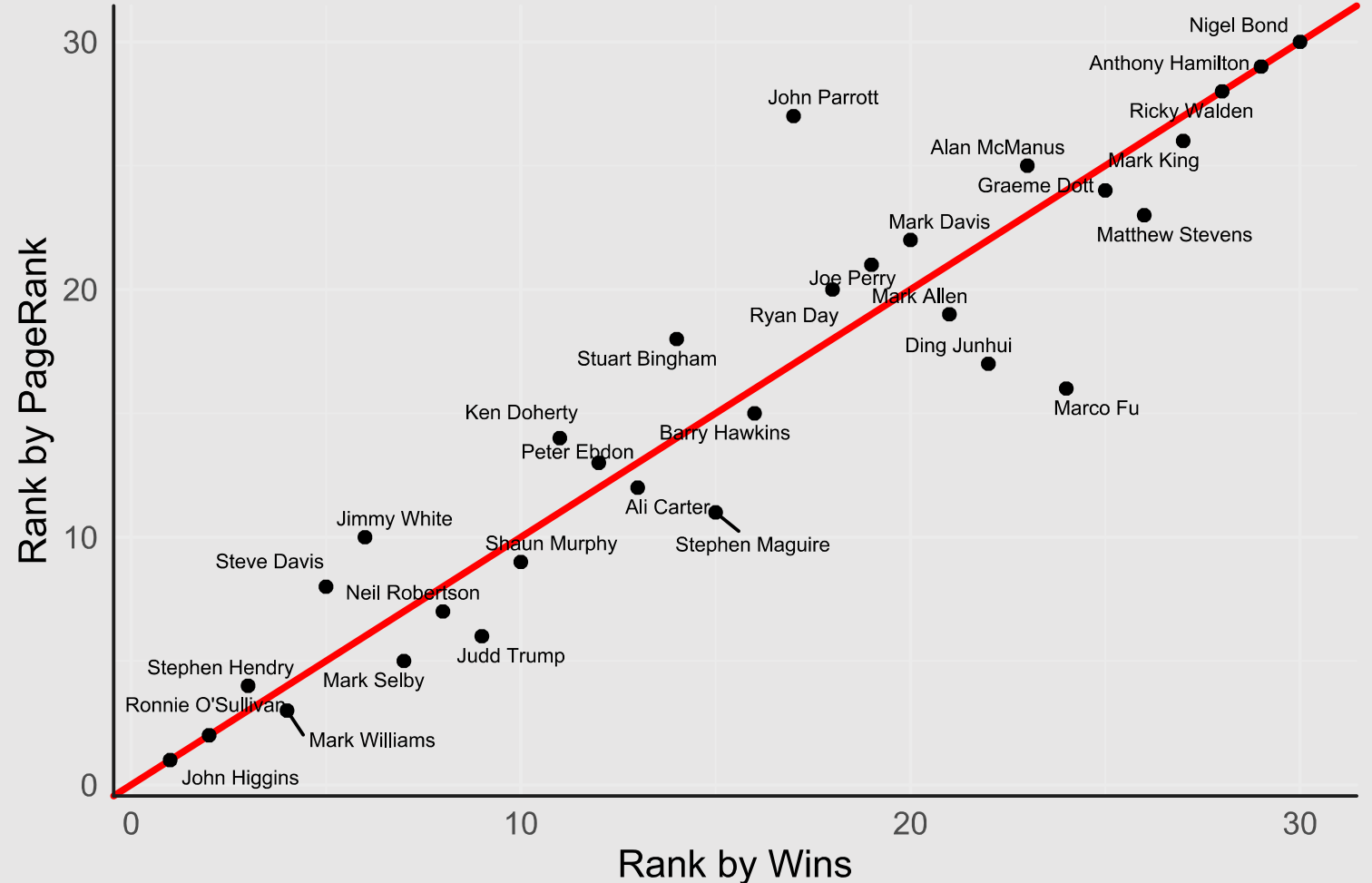
Rankings – all-time

The Top 10 Snooker Players

1968 - 2020

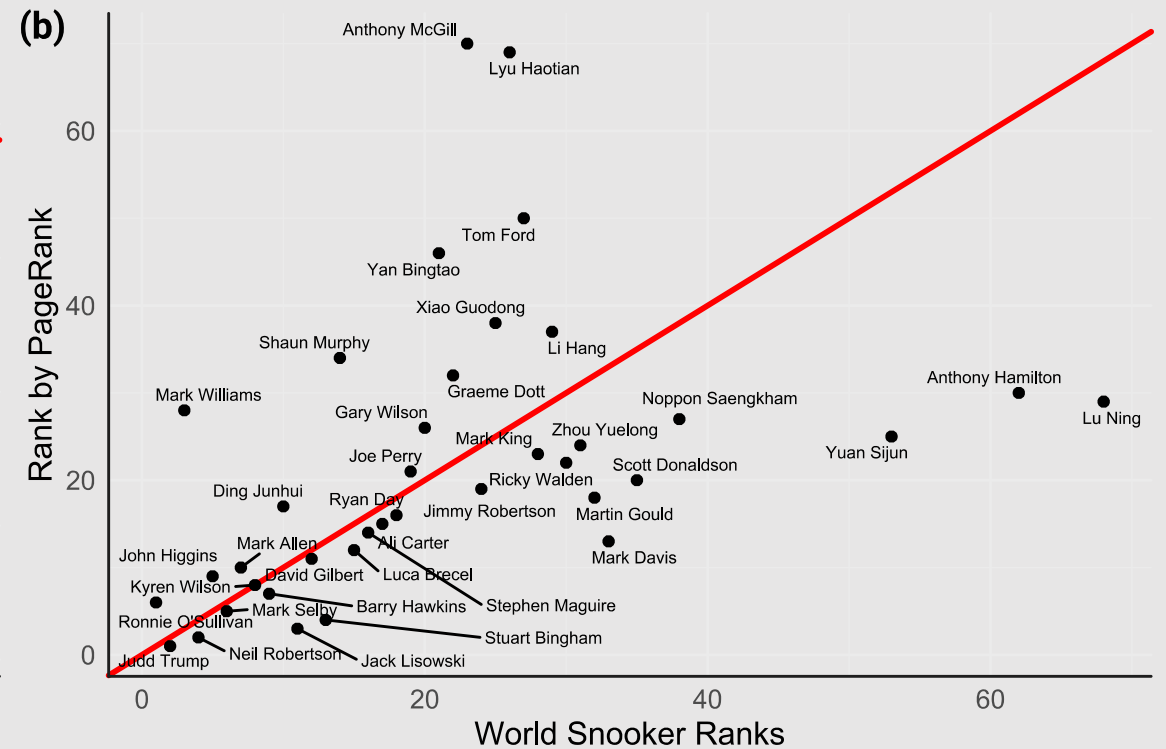
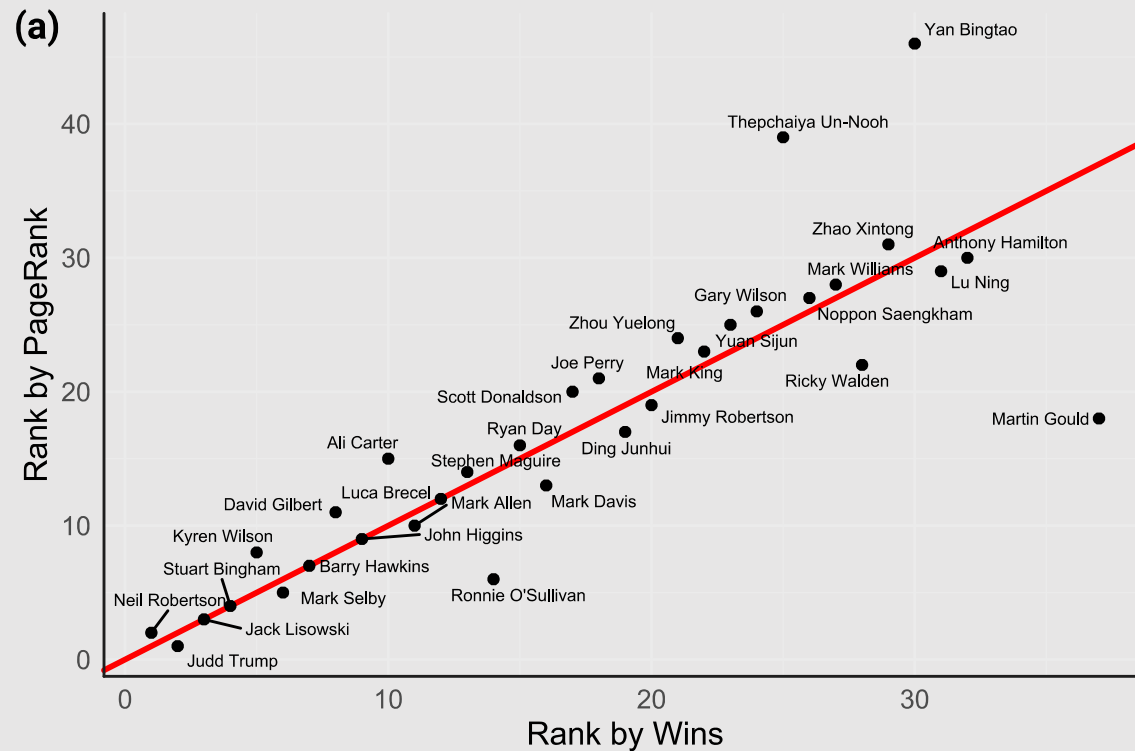
RANK ¹	PLAYER	SCORE	NATION
1	John Higgins	0.0204	
2	Ronnie O'Sullivan	0.0201	
3	Mark Williams	0.0169	
4	Stephen Hendry	0.0164	
5	Mark Selby	0.0149	
6	Judd Trump	0.0136	
7	Neil Robertson	0.0134	
8	Steve Davis	0.0129	
9	Shaun Murphy	0.0126	
10	Jimmy White	0.0116	

¹ Calculated via a complex networks approach.



Rankings – temporal periods

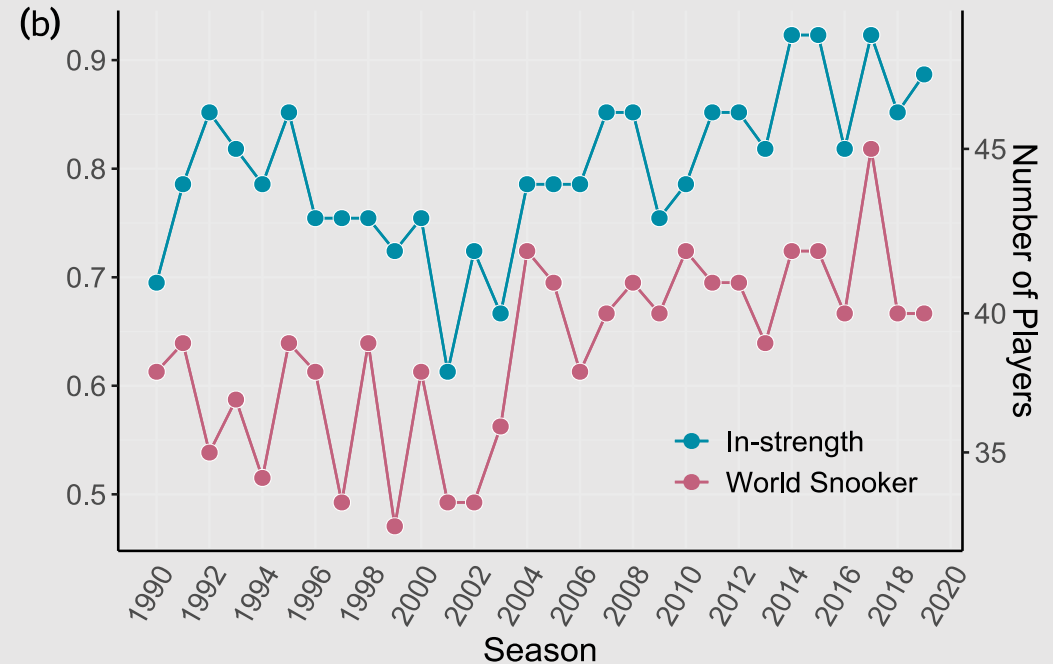
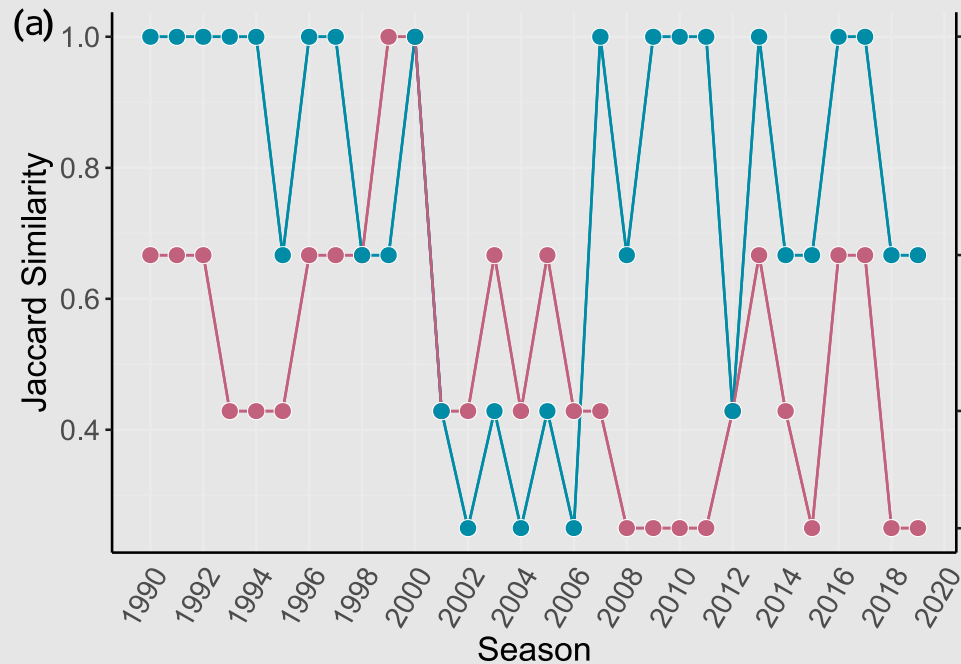
- Any time period may be considered by filtering the matches



Rankings – Similarity of approaches

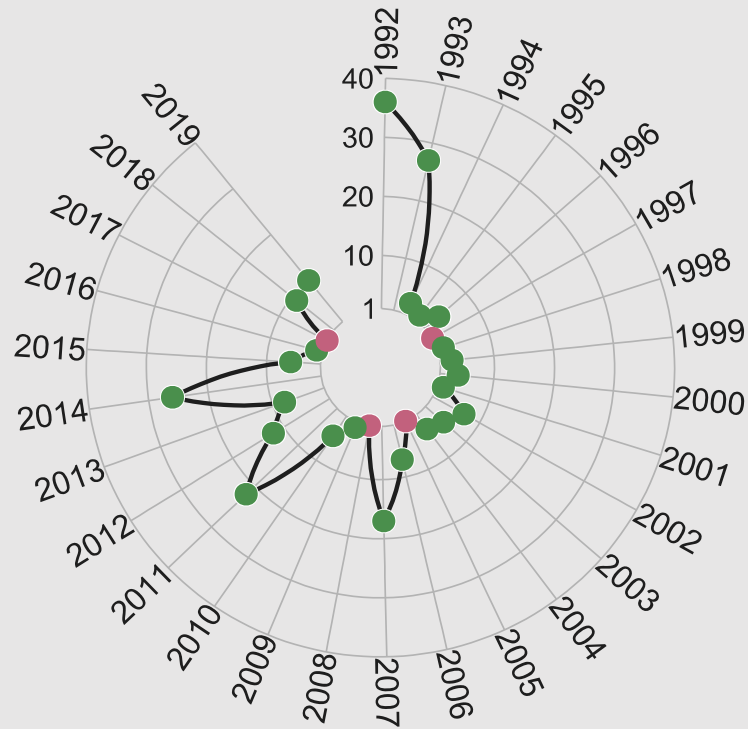
- Two rankings from approaches A and B the Jaccard similarity is

$$J(A, B) = \frac{|A \cap B|}{|A \cup B|}$$

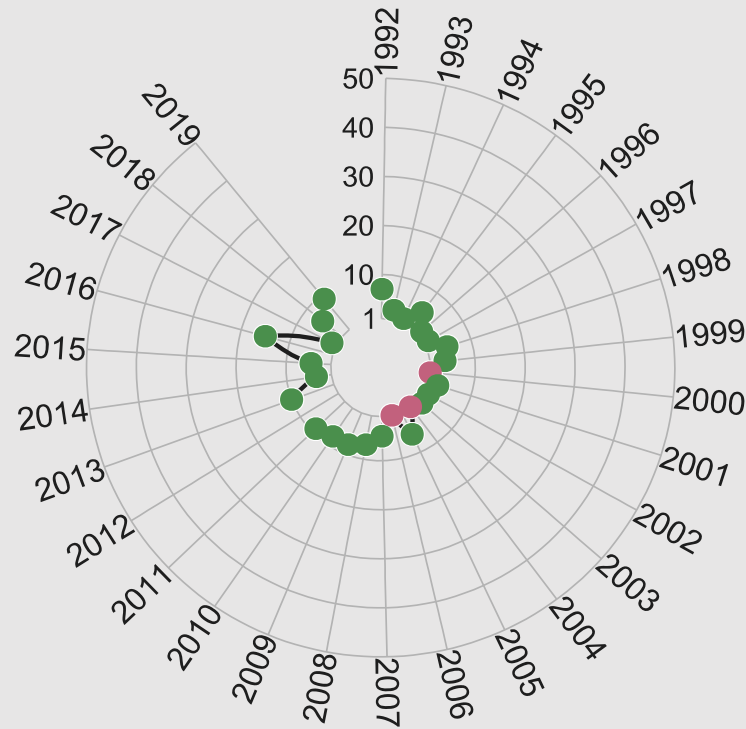


Visualizations – Rank Clocks

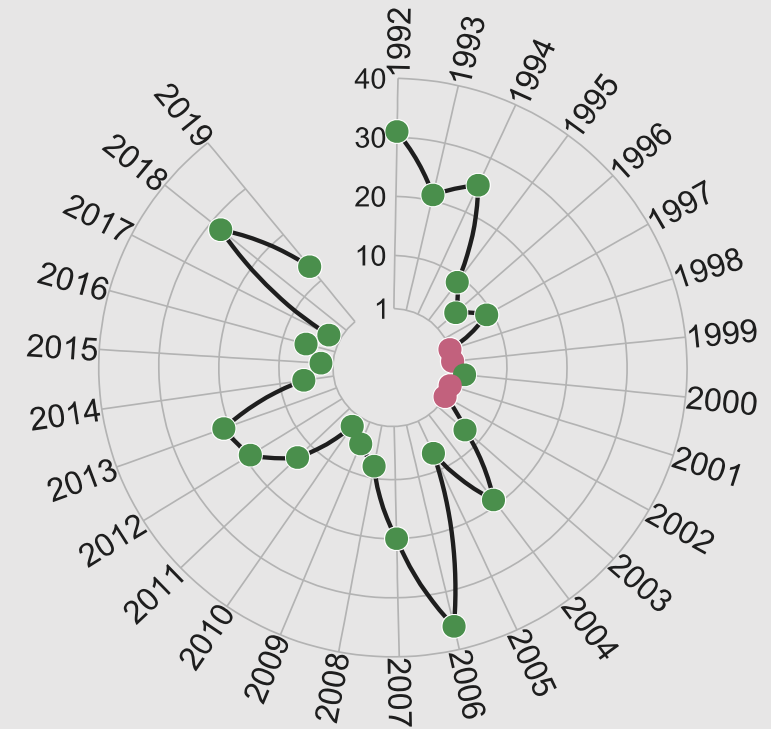
John Higgins



Ronnie O'Sullivan



Mark Williams



Conclusions

- Ranking is difficult – **networks** can help
- **Detailed data** of competitions is required
- Our approach allows **detailed inferences** to be made including **all-time ranks**, arbitrary **temporal periods**, ...
- Captures **stronger features** of other approaches while also considering the **quality of opponent** defeated

Collaborators & Thanks



James Gleeson
(University of Limerick)

A complex networks approach to ranking professional Snooker players


Joseph D. O'Brien* and James P. Gleeson
*MACSI, Department of Mathematics and Statistics,
University of Limerick, Limerick V94 T9PX, Ireland*


A detailed analysis of matches played in the sport of Snooker during the period 1968-2020 is used to calculate a directed and weighted dominance network based upon the corresponding results. We consider a ranking procedure based upon the well-studied PageRank algorithm that incorporates details of not only the number of wins a player has had over their career but also the quality of opponent faced in these wins. Through this study we find that John Higgins is the highest performing Snooker player of all time with Ronnie O'Sullivan appearing in second place. We demonstrate how this approach can be applied across a variety of temporal periods in each of which we may identify the strongest player in the corresponding era. This procedure is then compared with more classical ranking schemes. Furthermore, a visualization tool known as the rank-clock is introduced to the sport which allows for immediate analysis of the career trajectory of individual competitors. These results further demonstrate the use of network science in the quantification of success within the field of sport.

Keywords: Network analysis | PageRank | sports | science of success

Thank you for listening!

arXiv:2010.08395 (2020)

 obrienjoey/snooker_rankings

 joseph.obrien@ul.ie