

This is an Open Access document downloaded from ORCA, Cardiff University's institutional repository:<https://orca.cardiff.ac.uk/id/eprint/129863/>

This is the author's version of a work that was submitted to / accepted for publication.

Citation for final published version:

Lewis, Rhyd 2020. Five degrees of separation from De Niro - charting the social networks of movie stars. *The Conversation*

Publishers page: <https://theconversation.com/five-degrees-of-separa...>

Please note:

Changes made as a result of publishing processes such as copy-editing, formatting and page numbers may not be reflected in this version. For the definitive version of this publication, please refer to the published source. You are advised to consult the publisher's version if you wish to cite this paper.

This version is being made available in accordance with publisher policies. See <http://orca.cf.ac.uk/policies.html> for usage policies. Copyright and moral rights for publications made available in ORCA are retained by the copyright holders.



# Five degrees of separation from De Niro – charting the social networks of movie stars

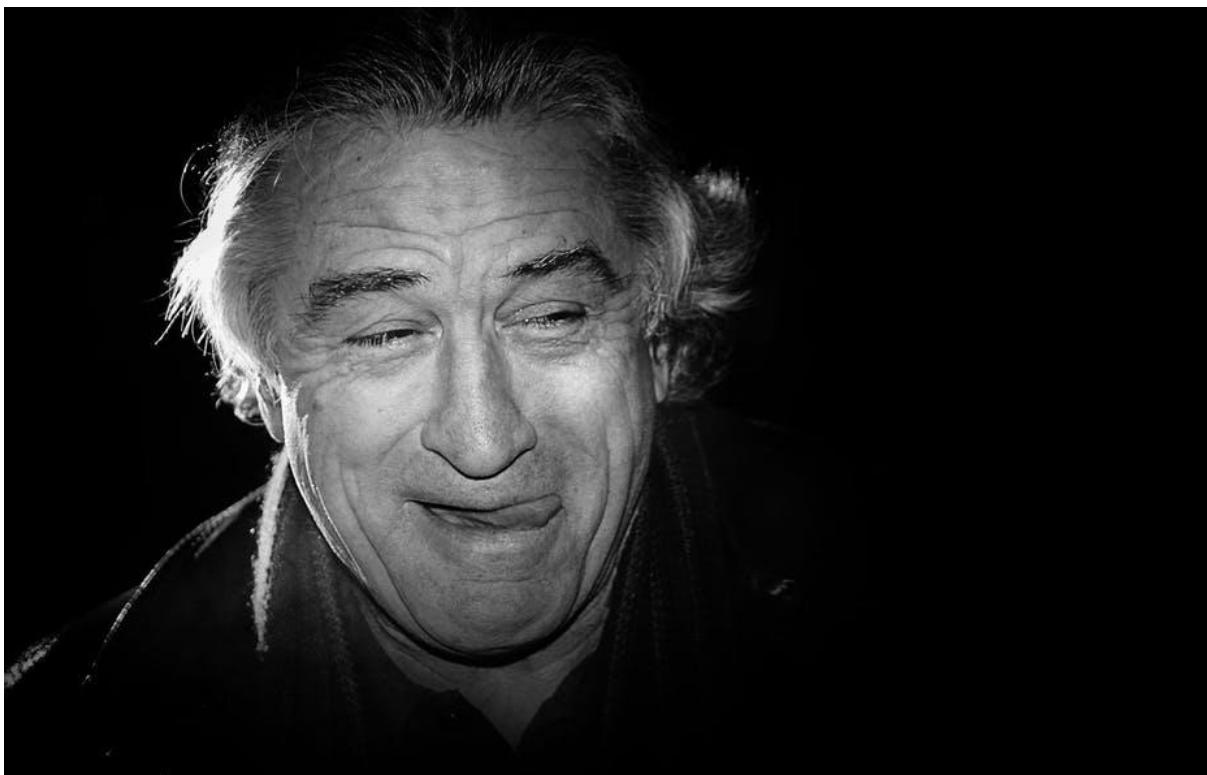
February 20, 2020 7.40am GMT

Author



Rhyd Lewis

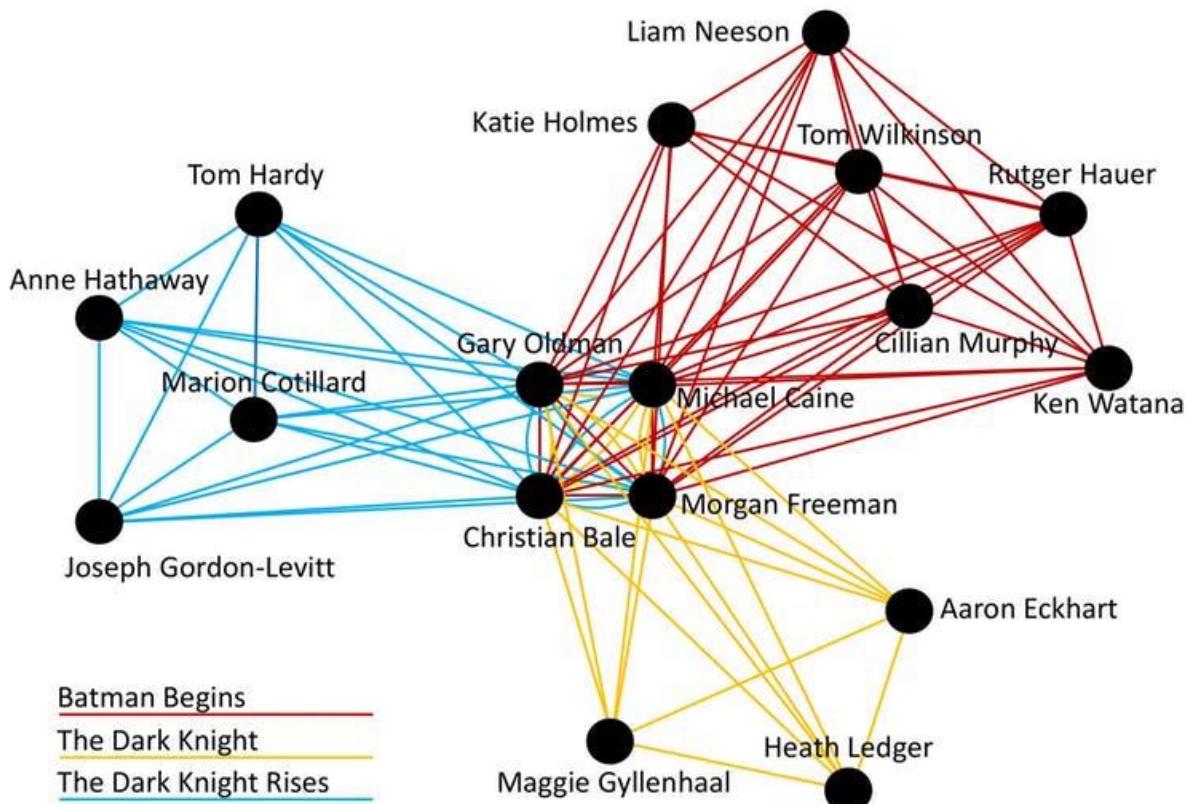
Reader in Mathematics, Cardiff University



Robert De Niro is one of the top middle men in the world of cinema. [Denis Makarenko/Shutterstock](#)

Social network analysis is a branch of data science that allows the investigation of how people connect and interact. It can help to reveal patterns in voting preferences, aid the understanding of how ideas spread, and even help to model the spread of diseases. It can also be used to find out who is at the centre of the movie universe.

To give an example, the picture below shows the social network of actors appearing in Christopher Nolan's three Batman movies. Each node in this network corresponds to an individual actor. Connections between actors indicate an appearance in a movie together and different coloured connections represent different movies. The network clearly shows that Christian Bale is very central, having appeared in all three movies and with all other actors. The same is also true of Gary Oldman, Michael Caine and Morgan Freeman. In contrast, the remaining actors are rather peripheral and have fewer connections to others.



The same principles can be used to [generate a social network](#) of all actors from all movies. Luckily, the information is available through [IMDB](#), which contains cast lists of around 160,000 different movies. This leads to a large social network of around 400,000 actors and nearly 10 million different connections. From here it is possible to determine the most “important” or “central” actors in this network .

### Most connected

One way of measuring importance is to simply count the number of movies that an actor has appeared in.



Bollywood actress Sukumari. Wikimedia, [CC BY-SA](#)

The top spots are occupied by actors from Indian cinema – with the great Bollywood actress Sukumari winning the competition, having made 703 credited movie appearances. The next actors are then [[Jagathy Sreekumar](#)], [Adoor Bhasi](#), [Brahmanandam](#), [Manorama](#), [Sankaradi](#) and [Prem Nazir](#), who each appear in more than 500 movies. The only non-Bollywood actor breaking the top 20 is Oliver Hardy, who has 373 films to his name. Then, if you were to look for any of this year's acting Oscar winners, you would have to go all the way down to the bottom to find Brad Pitt, who has made a total of 56 movie appearances.

Another simple statistic from this network is the number of different people that each actor has appeared with. In this case, the top ranked actors again come from the world of Bollywood, with actors such as [Nassar](#), Sukumari and Manorama all having appeared with more than 2,500 people. Other prominent names include Christopher Lee and John Carradine at just over 2,000, and John Wayne on 1,900. In comparison, the highest of this year's Oscar winners is again Brad Pitt on a lowly 735, while Renée Zellweger and Joaquin Phoenix both have disappointing statistics of less than 500.

### [The middlemen](#)

A trait of the above two measures is that they highlight actors that happen to have appeared in many movies that also feature large casts. Other methods of finding central actors are therefore also needed. One such measure is “[betweenness centrality](#)”, which considers the number of shortest paths in a network that pass through a particular node. This helps to detect the “middlemen” that serve as a links between different parts of a network. A useful analogy can be drawn with road networks of cities. Short driving routes across a city will often pass through the same locations in the road network. As a result, the nodes at these locations are considered more “central” and important to the network.

In social networks we can do similar things by looking at the shortest paths between actors and then identifying actors that always seem to appear on these paths. According to these measures, the top five most central actors are now Christopher Lee, Michael Caine, Harvey Keitel, Christopher Plummer and Robert De Niro. By comparison, none of this year's recipients of acting Oscars appears in the top 1,000 of this list.

It is noticeable that the top actors in these lists have all had very long acting careers. Their being middlemen then is natural since long careers bring about more acting opportunities,

helping to improve an actor's network connectivity. So how do today's Oscar winners compare? Right now it seems that they have some way to go. But we really should be fair and give them a few more decades before making a final judgement.