



SOCIAL NETWORK ANALYSIS PROJECT

Prof. Shivraj Kanungo

Question of Interest

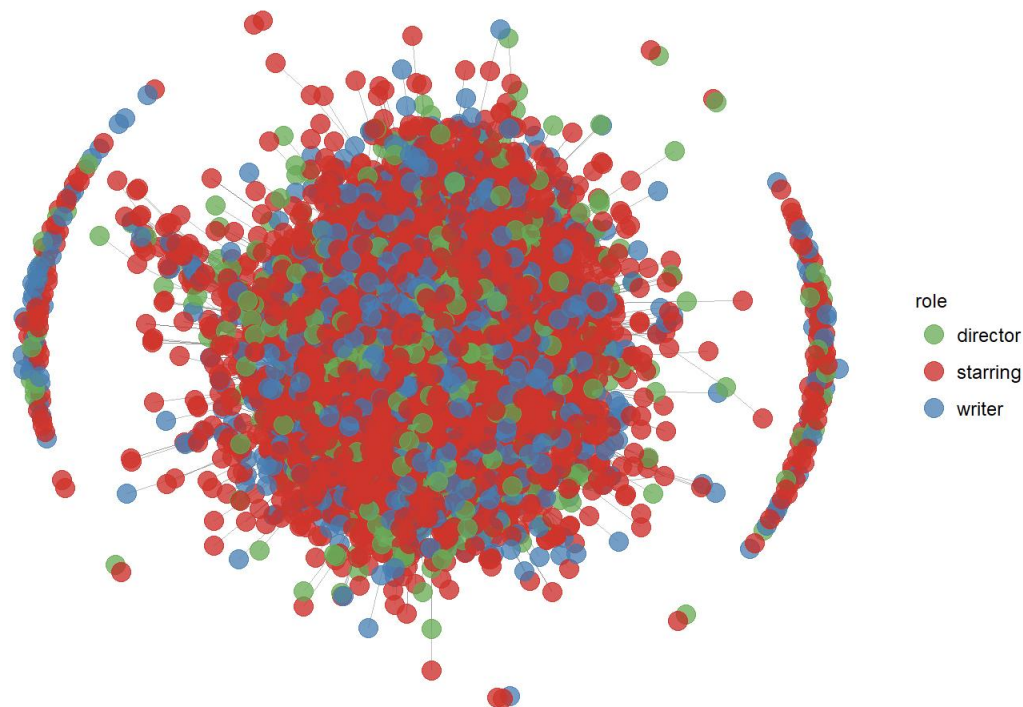
There is no doubt that the Oscar Academy Award Ceremony is one of the biggest events throughout the year. It once again drew the world's attention as the 90th Oscar Award Ceremony successfully landed in Los Angeles on March 4th. The focus of the topic is always about the Oscar Winning Actors and Actress and we are so amazed at how successful, famous and popular they are on and off stage. So the topic of movie actors also become the interest of our analysis.

People not only talk about these successful actors and their marvelous work in film industry, they are also very interested in their colorful personal life and their mysterious social network. How many peer actor friends do they have and are they also popular among them? How many important directors and writers are they connected with and does the relationship help them with their career?

These are all the questions we are super curious about. That's how our driving question of our analysis came out: How does actors' social network relate to their success?

To analyze this problem, we selected two groups of people in the film industry that have a potential huge influence on the actors: Directors who direct films and Writers who make the film scripts. We also created different actors group with different success levels and then compared their social performance in the different social relations.

The overall network with actors, directors and writers looks like below.



Data Source

We collected network data of actors-directors-writers from <https://aminer.org/data-sna#Movie> website. The source data consist of two sections, vertices and edges. The vertices section explains the node attributes and the edges section gives the information about ties between the nodes. Only one type of

Social Network Analysis

relation is used in the data set, which can be simplified as an undirected “knows” relationship, i.e. Actor X and Director Y know each other.

To download the dataset, click on the download link for movies dataset in the website mentioned above.

Additionally, we have used Wikipedia Oscar Best Actor/Actress List and IMDB Hollywood A-lister List to create a group of Oscar winners actors and a group of famous actors (not Oscar winners but are also popular and famous). The lists are not comprehensive list but serve our purpose of comparison.

Our Network Analysis Approach

The first part of our analysis is based on comparing three groups of actors (group 1- Oscars winners, group 2 - famous actors who are not Oscar winners but are also popular and famous and group 3 - all actors) in terms of different types centralities/importance in three different networks (actor-actor, actor-director and actor-writer networks). This comparative analysis among three groups will help determine whether actors’ sociability in the networks is actually one of the main reasons for their success in Oscar or popularity.

The second part of our analysis is focused on finding the most sociable directors and writers in the network and how many of these directors or writers are known by the actors from group 1, 2 and 3. The motive is to find out how each group differs in forming ties with most sociable or important directors and writers in the network.

Results

Firstly, we calculated the degree centrality, closeness centrality and betweenness centrality of all the actors from the network that consist of only actor-actor ties. We compared the descriptive statistics of different centralities among the group. Mean values are given in the below table to illustrate the difference between these groups.

Group	Average Degree Centrality	Average Closeness Centrality	Average Betweenness Centrality
Oscar Winners	37.8	6.644e-07	156447
Famous Actors	31.7	6.640e-07	103444
All actors in the network	6.9	6.444e-07	12518

To find out the importance/sociability of an actor in his/her social network we gave equal weight to all the three types of centralities and we also assigned an overall rank to each actor.

The results clearly show that ranks of actors from group 1 and group 2 are better than that of group 3. Oscar winners have higher average degree centrality, higher average betweenness centrality, and higher average closeness centrality compared to famous actors, and famous actors all have higher centrality compared to average centrality of all the actors. Please refer to below tables for more details on the results generated in r console.

Social Network Analysis

	Name	deg_centr_actors	clo_centr_actors	bet_centr_actors
A Martinez	: 1	Min. : 1.000	Min. :1.793e-08	Min. : 0
A. J. Cook (actress)	: 1	1st Qu.: 2.000	1st Qu.:6.603e-07	1st Qu.: 0
Ádgar RamÁrez	: 1	Median : 4.000	Median :6.618e-07	Median : 1551
Aaliyah	: 1	Mean : 6.912	Mean :6.444e-07	Mean : 12518
Aaron Ashmore	: 1	3rd Qu.: 8.000	3rd Qu.:6.628e-07	3rd Qu.: 9240
Aaron Eckhart	: 1	Max. :204.000	Max. :6.661e-07	Max. :1106394
(Other)	:7463			

	Name	deg_centr_actors	clo_centr_actors	bet_centr_actors
Adrien Brody	: 1	Min. :10.00	Min. :6.626e-07	Min. : 6354
Brad Pitt	: 1	1st Qu.:23.75	1st Qu.:6.638e-07	1st Qu.: 63425
Casey Affleck	: 1	Median :32.00	Median :6.646e-07	Median : 86951
Cate Blanchett	: 1	Mean :37.80	Mean :6.644e-07	Mean :156447
Charlize Theron	: 1	3rd Qu.:46.75	3rd Qu.:6.649e-07	3rd Qu.:183550
Christian Bale	: 1	Max. :93.00	Max. :6.661e-07	Max. :750236
(Other)	:14			

	Name	deg_centr_actors	clo_centr_actors	bet_centr_actors
Ian McKellen	:1	Min. : 10.0	Min. :6.631e-07	Min. : 9690
James Franco	:1	1st Qu.: 17.5	1st Qu.:6.637e-07	1st Qu.: 35204
Jennifer Aniston	:1	Median : 25.0	Median :6.640e-07	Median : 51905
Jessica Alba	:1	Mean : 31.7	Mean :6.640e-07	Mean :103444
Keira Knightley	:1	3rd Qu.: 32.5	3rd Qu.:6.642e-07	3rd Qu.: 84905
Marilyn Monroe	:1	Max. :102.0	Max. :6.655e-07	Max. :510702
(Other)	:4			

Descriptive statistics of different Centralities in actor-actor network

actors_vertex_id	Name	deg_centr_actors	clo_centr_actors	bet_centr_actors	deg_score	clo_score	bet_score	OverallScore	rank
4621	John Wayne	204	6.657719e-07	1106394.21	7469	7468	7469	22406	1
11031	Meryl Streep	93	6.660796e-07	750235.47	7463	7469	7468	22400	2
11316	Elvis Presley	115	6.656349e-07	672745.69	7466	7464	7467	22397	3
6938	Frank Sinatra	140	6.656243e-07	624572.70	7468	7463	7466	22397	4
13275	Bette Davis	120	6.656628e-07	575130.95	7467	7465	7464	22396	5
6963	Katharine Hepburn	95	6.656899e-07	491385.41	7464	7467	7462	22393	6
11046	Jack Nicholson	87	6.656788e-07	624297.54	7458	7466	7465	22389	7
12386	Marilyn Monroe	102	6.654963e-07	510702.39	7465	7457	7463	22385	8
10058	Dean Martin	92	6.655871e-07	385859.62	7462	7461	7455	22378	9
8647	Tom Hanks	80	6.655485e-07	474851.20	7453	7459	7459	22371	10

Ranks of all actors

actors_vertex_id	Name	deg_centr_actors	clo_centr_actors	bet_centr_actors	deg_score	clo_score	bet_score	OverallScore	rank
11031	Meryl Streep	93	6.660796e-07	750235.472	7463	7469	7468	22400	2
8647	Tom Hanks	80	6.655485e-07	474851.201	7453	7459	7459	22371	10
2965	Robin Williams	56	6.652422e-07	287625.442	7412	7447	7442	22301	27
11989	Brad Pitt	53	6.649878e-07	210284.486	7400	7416	7415	22231	52
2506	Nicole Kidman	49	6.650820e-07	188680.325	7389	7429	7400	22218	57
2007	Leonardo DiCaprio	44	6.648976e-07	164055.674	7368	7400	7387	22155	73
6830	Julia Roberts	46	6.647893e-07	181840.040	7375	7379	7397	22151	77
7767	Natalie Portman	41	6.645454e-07	147964.385	7351	7318	7369	22038	108
8813	Kate Winslet	31	6.646996e-07	120539.357	7246	7360	7337	21943	131

Ranks of Oscar winners

Social Network Analysis

actors_vertex_id	Name	deg_centr_actors	clo_centr_actors	bet_centr_actors	deg_score	clo_score	bet_score	OverallScore	rank
12386	Marilyn Monroe	102	6.654963e-07	510702.385	7465	7457	7463	22385	8
3376	Ian McKellen	40	6.646130e-07	157312.611	7344	7334	7379	22057	103
8051	Jennifer Aniston	33	6.642320e-07	79499.264	7278	7189	7235	21702	211
1556	Scarlett Johansson	31	6.640759e-07	86707.233	7255	7111	7262	21628	234
150	Jessica Alba	25	6.640168e-07	59252.689	7141	7070	7124	21335	323
14439	Orlando Bloom	25	6.640662e-07	44288.492	7120	7102	6996	21218	351
5341	Keira Knightley	25	6.637854e-07	44556.617	7135	6869	6999	21003	424
3791	James Franco	15	6.631423e-07	32175.130	6672	6113	6813	19598	808
14197	Mark Ruffalo	10	6.637347e-07	10257.764	6000	6826	5722	18548	1132
11403	Ryan Gosling	11	6.633292e-07	9690.422	6194	6383	5667	18244	1214

Ranks of famous actors

Then we calculated the degree centrality of all actors from the network that consist of only actor-director ties. We saw similar results as we saw in case of actor-actor network. The result that Oscar winners, on average know more directors than famous actors and famous actors on average know more directors than the average of all the actors. Mean values are given in the below table to illustrate the difference between these groups.

Group	Mean Degree Centrality
Oscar Winners	11.3
Famous Actors	8.50
All actors in the network	2.85

Please refer to below tables for more details on the the results generated in r console.

```

      Name      Role degree_ActDir_only rank
Ãdgar RamÃrez: 1 director: 0 Min. : 1.00 Min. : 1
Aaliyah      : 1 movie : 0 1st Qu.: 1.00 1st Qu.:1182
Aaron Ashmore : 1 starring:4727 Median : 2.00 Median :2364
Aaron Eckhart : 1 writer : 0 Mean : 2.85 Mean :2364
Aaron Yoo     : 1 3rd Qu.: 3.00 3rd Qu.:3546
Aasif Mandvi  : 1 Max. :42.00 Max. :4727
(Other)      :4721

```

```

      Name      Role degree_ActDir_only rank
Adrien Brody : 1 director: 0 Min. : 2.00 Min. : 11.00
Brad Pitt    : 1 movie : 0 1st Qu.: 7.00 1st Qu.: 88.75
Casey Affleck : 1 starring:20 Median :11.00 Median : 162.00
Cate Blanchett : 1 writer : 0 Mean :11.30 Mean : 336.15
Charlize Theron: 1 3rd Qu.:13.25 3rd Qu.: 367.00
Christian Bale : 1 Max. :26.00 Max. :2152.00
(Other)      :14

```

```

      Name      Role degree_ActDir_only rank
Ian McKellen :1 director: 0 Min. : 3.00 Min. : 5.0
James Franco :1 movie : 0 1st Qu.: 5.25 1st Qu.: 272.8
Jennifer Aniston:1 starring:10 Median : 6.00 Median : 459.5
Jessica Alba :1 writer : 0 Mean : 8.50 Mean : 587.2
Keira Knightley :1 3rd Qu.: 8.25 3rd Qu.: 574.0
Marilyn Monroe :1 Max. :31.00 Max. :1544.0
(Other)      :4

```

Descriptive statistics of different Centralities in actor-director network

Social Network Analysis

We performed similar analysis on actor-writer network and the results were as expected. Oscar winners, on average know more writers than famous actors and famous actors on average know more writer than the average of all the actors. Mean values are given in the below table to illustrate the difference between these groups.

Group	Mean Degree Centrality
Oscar Winners	14.55
Famous Actors	8.3
All actors in the network	3.05

Please refer to the below tables for more details on the results for the results generated in r console.

Name	Role	degree_ActWri_only	rank
A. J. Langer : 1	director: 0	Min. : 1.000	Min. : 1
Ádgar RamÁrez: 1	movie : 0	1st Qu.: 1.000	1st Qu.:1282
Aaliyah : 1	starring:5124	Median : 2.000	Median :2562
Aaron Eckhart : 1	writer : 0	Mean : 3.057	Mean :2562
Aaron Norris : 1		3rd Qu.: 4.000	3rd Qu.:3843
Aasif Mandvi : 1		Max. :59.000	Max. :5124
(Other) :5118			

Name	Role	degree_ActWri_only	rank
Adrien Brody : 1	director: 0	Min. : 3.00	Min. : 4.0
Brad Pitt : 1	movie : 0	1st Qu.: 7.75	1st Qu.: 79.0
Casey Affleck : 1	starring:20	Median :12.00	Median : 161.0
Cate Blanchett : 1	writer : 0	Mean :14.55	Mean : 289.6
Charlize Theron: 1		3rd Qu.:15.00	3rd Qu.: 377.0
Christian Bale : 1		Max. :37.00	Max. :1597.0
(Other) :14			

Name	Role	degree_ActWri_only	rank
Ian McKellen :1	director: 0	Min. : 2.0	Min. : 11.0
James Franco :1	movie : 0	1st Qu.: 4.0	1st Qu.: 347.5
Jennifer Aniston:1	starring:10	Median : 6.0	Median : 671.0
Jessica Alba :1	writer : 0	Mean : 8.3	Mean : 747.0
Keira Knightley :1		3rd Qu.: 8.0	3rd Qu.:1052.2
Marilyn Monroe :1		Max. :30.0	Max. :1874.0
(Other) :4			

Descriptive statistics of different Centralities in actor-writer network

In the second part of our analysis, we initially identified most sociable/important directors and writers in the network. To identify most sociable/important directors and writers, we calculated the degree centrality, closeness centrality and betweenness centrality of all the directors and writers in whole network and assigned ranks to them by giving equal weight to all the centralities. An ordered (based on rank) list of top 20 important directors and writers is shown on the below screenshot.

Social Network Analysis

Name	Role	deg_centr_directors	clo_centr_directors	bet_centr_directors	deg_score	clo_score	bet_score	OverallScore	rank
Steven Spielberg	director	298	8.113513e-08	8272421.7	2377	2377	2377	7131	1
Alfred Hitchcock	director	290	8.112335e-08	6157076.6	2376	2375	2376	7127	2
Martin Scorsese	director	197	8.112750e-08	5048441.7	2373	2376	2375	7124	3
Clint Eastwood	director	203	8.111485e-08	3235796.6	2374	2373	2372	7119	4
Laurence Olivier	director	215	8.111094e-08	3656896.3	2375	2369	2373	7117	5
Robert Altman	director	159	8.111801e-08	3127804.6	2369	2374	2370	7113	6
Charles Laughton	director	163	8.110630e-08	4235589.1	2370	2367	2374	7111	7
John Ford	director	172	8.110470e-08	3185211.8	2372	2364	2371	7107	8
Sean Connery	director	147	8.110949e-08	2398828.9	2367	2368	2367	7102	9
Paul Newman	director	146	8.111256e-08	2148545.1	2365	2370	2365	7100	10
Sidney Lumet	director	169	8.110387e-08	2082606.9	2371	2363	2362	7096	11
Al Pacino	director	151	8.110111e-08	2035527.0	2368	2362	2359	7089	12
Jack Lemmon	director	136	8.110563e-08	2062762.7	2361	2366	2361	7088	13
Sam Peckinpah	director	126	8.111303e-08	2049266.1	2356	2371	2360	7087	14
Mel Gibson	director	139	8.109956e-08	2355345.0	2362	2356	2366	7084	15
Johnny Depp	director	146	8.109777e-08	2406087.8	2364	2351	2368	7083	16
David Lean	director	96	8.111420e-08	2588514.7	2337	2372	2369	7078	17
Dennis Hopper	director	115	8.110530e-08	2132525.4	2348	2365	2364	7077	18
James Garner	director	124	8.110007e-08	1972406.8	2355	2359	2357	7071	19
Nicolas Cage	director	122	8.110109e-08	1766130.0	2353	2361	2354	7068	20

List of important/sociable directors in the whole network

Name	Role	deg_centr_writers	clo_centr_writers	bet_centr_writers	deg_score	clo_score	bet_score	OverallScore	rank
Orson Welles	writer	274	8.015470e-08	7904072.7	3724	3724	3724	11172	1
Woody Allen	writer	221	8.014511e-08	3866106.4	3723	3721	3722	11166	2
Francis Ford Coppola	writer	174	8.014595e-08	3831535.8	3722	3722	3721	11165	3
Billy Wilder	writer	144	8.014064e-08	3206407.7	3718	3720	3719	11157	4
Terry Southern	writer	129	8.015074e-08	4244324.1	3710	3723	3723	11156	5
Stephen King	writer	173	8.013418e-08	3612691.9	3721	3715	3720	11156	6
Stanley Kubrick	writer	144	8.013603e-08	2444218.8	3717	3718	3714	11149	7
Tim Burton	writer	161	8.012946e-08	2537175.7	3720	3712	3715	11147	8
Barbra Streisand	writer	140	8.013425e-08	2392583.2	3715	3716	3713	11144	9
Steve Martin	writer	131	8.013386e-08	2618683.2	3711	3714	3717	11142	10
Roger Ebert	writer	118	8.013498e-08	2841339.5	3704	3717	3718	11139	11
Quentin Tarantino	writer	149	8.012326e-08	2585045.7	3719	3703	3716	11138	12
Neil Simon	writer	141	8.012641e-08	2071886.2	3716	3707	3709	11132	13
Arnold Schwarzenegger	writer	128	8.012797e-08	2348294.3	3708	3711	3712	11131	14
David Bowie	writer	112	8.012796e-08	1744849.8	3700	3710	3703	11113	15
Gene Hackman	writer	138	8.012066e-08	1806011.2	3713	3694	3705	11112	16
Peter Bogdanovich	writer	89	8.013773e-08	1992039.4	3684	3719	3708	11111	17
Walt Disney	writer	135	8.011720e-08	2300928.7	3712	3686	3711	11109	18
Oliver Stone	writer	127	8.012107e-08	1817675.5	3707	3696	3706	11109	19
George Lucas	writer	111	8.012693e-08	1711149.8	3698	3709	3702	11109	20

List of important/sociable writer in the whole network

Social Network Analysis

Finally, we calculated the average important directors and writers each group knows. As we can see from the below tables that Oscar winners know more important directors and writers than other groups. This result proves that Oscar winners and famous actors not only know more people in the network, they also know more important people, which may relate to their success in the film industry.

Groups	Mean
Oscar Winners	4.9
Famous Actors	3.3
All the Actors	1.6

Groups	Mean
Oscar Winners	4.1
Famous Actors	2.1
All the Actors	1.6

Conclusion

- ❖ Oscar winners are more sociable among actors, writers and directors and they know more big names than famous actors.
- ❖ Famous actors are more sociable among actors, writers and directors and they know more big names comparing to the average of the average of the whole actor population.
- ❖ An actor's success is highly related to the number of people in the film industry they know and the number of important people in the industry they know.