The construction and application of impact model for social networks

Liguo Huang*, Lufang Mi

Department of Mathematics, Binzhou University, Huanghe Str. 5, Binzhou, China

Received 23 November 2014, www.cmnt.lv

Abstract

The co-authorship network of scientists represents a prototype of complex evolving networks. In addition, it offers one of the most extensive database to date on social networks. The focus of our work is to analyze influence and impact in research networks and other areas of society. Through the analysis over 18000 lines of raw data in the Erdos1.htm, we infer the structure of this network containing 9784 nodes and 17273 edges. Global metrics such as degree centrality, closeness centrality and betweenness centrality can be used to identify the influence. Based on the influence measurement model, we find that Alon Nogam, Harary Frank and Shelah Saharon have significant influence on the network. By analyzing the important works from its publication, we build another influence measure model, which includes the impact factor of researchers, publications and journals, to determine the papers' relative influence. The conclusion is that the second paper possesses significant influence on the network.

Keywords: Network influence; Degree Network; Density centrality; Social networks analysis method; Model

1 Introduction

In recent years, there has been a sharp increase in the number of collaborations between scholars. One of the techniques to determine influence of academic research is to build and measure properties of citation or co-author networks. Co-authoring a manuscript usually connotes a strong influential connection between researchers. One of the most famous academic co-authors was the 20th-century mathematician Paul Erdos who had over 500 co-authors and published over 1400 technical research papers. Since scientific collaborations are defined as "interactions taking place within a social context among two or more scientists that facilitate the sharing of meaning and completion of tasks with respect to a mutually shared, super-ordinated goal", those collaborations frequently emerge from, and are perpetuated through, social networks. Since social networks may span disciplinary, institutional, and national boundaries, it can influence collaboration in multiple ways [1]. Social network analysis has produced many results concerning social influence, social groupings, inequality, disease propagation, communication of information, and indeed almost every topic that has interested 20th century sociology[2-4].

The paper is organized as follows: on the second part, we build models and algorithm to get a co-author network of the approximately 510 researchers from the file Erdos1, who coauthored a paper with Erdos, but do not include Erdos. The properties of this network is also analyzed. On the third part, we measure to compare the significance of a research paper by analyzing the important works that follow from its publication and discuss how more thorough analyses of network, semantic, and text analyses of the message contents to help our models. After the above analysis, we understand that find an index to assess the likelihood of nodes is the key factor. At last, we discuss the science, understanding, utility of model influence and impact within networks and their application in other areas.

2 Influence measurement model

In this part, we construct a co-author network of the approximately 510 researchers from the file Erdos1, who coauthored a paper with Erdos, but do not include Erdos. We also analyze the properties of this network.

2.1 THEORETICAL SOURCES

A social network can be conceptualized as a set of individuals or groups, each of which has connections of some kind to some or all of the others. In the language of social network analysis, people or groups are called "actors" or "nodes" and connections are referred to as "ties" or "links". Both actors and ties can be defined in different ways depending on the questions of interest. An actor might be a single person, a team, or a company. A tie might be a friendship between two people, collaboration or common member between two teams, or a business relationship between companies [2-4]. In scientific collaborations' network actors (nodes) are authors and ties (links) are co-authorship relations among them. A tie exists between each two actors (scholars) if they have at least one co-authored paper.

* ***Corresponding author's*** e-mail: liguoh123@sina.com

2.1.1 Network density

Density describes the general level of linkage among the nodes in a network. The more nodes are connected to one another, the denser the network is. Density describes the general level of cohesion in a network ([5]. More specifically, density of a network is the proportion of exiting links to the maximum possible number of distinct links.

2.1.2 Network centralization

Another method used to understand networks and their participants is to evaluate the location of authors in the network. Therefore, to calculate a network centralization, the first step is to find all nodes[5].

To examine if a whole network has a centralized structure, centralization can be used. It refers to 'compactness' of a network. A network's centralization indicates how tightly the network is organized around its most central nodes. The general view is finding differences between most central nodes' centrality scores and others'. Then, centralization is calculated as the ratio of the sum of these differences to the maximum possible sum of differences. Therefore, to calculate a network centralization, the first step is to find all nodes measures and then find the whole network centralities measures. The important node centrality measures are:

Degree centrality

The degree centrality is simply the number of other nodes connected directly to a node. Necessarily, a central node is not physically in the centre of the network as degree of a node is calculated in terms of the number of its adjacent nodes.

Closeness centrality

Freeman proposed closeness in terms of the distance among various nodes. In unconnected networks, every node is at an infinite distance from at least one node, and closeness centrality of all nodes would be 0. Thus, in order to solve this problem to consider all nodes, Freeman proposed to calculate closeness of a node. [6-7]

Betweenness centrality

Freeman also proposed a concept of centrality which measures the number of times a particular node lies between the various other nodes in the network[6]. Betweenness centrality is defined more precisely as the number of shortest paths (between all pairs of nodes) that pass through a given node[1].

The giant component: In small networks (few nodes and connections), all individuals belong to a small group of collaborations or communications. As all connected to one another by paths of intermediate acquaintances [2-4], it is important to realize that a collaboration network is usually fragmented in many clusters (components). One of the reasons for this is that in every field there are scientists that do not collaborate at all, that is they are single authors of all papers on which their names appear.

2.2 COMPLEX NETWORK MODEL

We build the complex network model as follows:

$$C_{ij} = \sum_j (P_{ij} + \sum_k P_{ik} P_{kj})^2 \text{ for } k \neq \text{i,j}, \quad (1)$$

where P_{ij} is the proportion of the number of cooperation in the number of all people, i is the number of cooperation. j is all the number of all people, C_{ij} represents the network constraint.

Structure of the network: To reveal the relationship between collaboration relationship parameter (CRP) and the network structure, we researched the networks ability to withstand the removal of links. The definition of CRP is given as

$$O_{ij} = (\sum_{k \in G} W_{ijk}) / (M_i + M_j - W_{ij} - \sum_{k \in G} W_{ijk}), \quad (2)$$

where G is the set of nodes that are neighbors of either i or j. M is the numerical of a node. M_i represents the numerical of node i. W_{ij} is the number of collaborate works between nodes i and j. $W_{ijk}=min\{W_{ik},W_{jk}\}a_{ij}$. O_{ij} is collaboration relationship parameter of denotes.

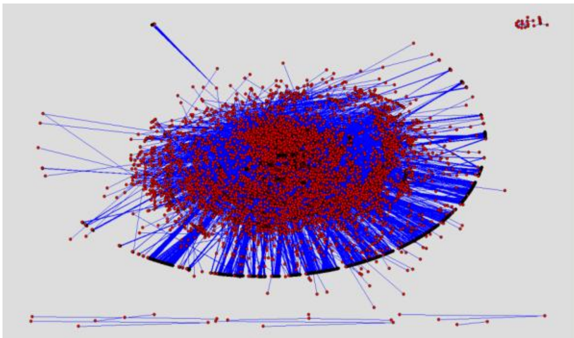

FIGURE 1 The co-author network from the Erdos1 data

2.3 THE CONCLUSION OF COMPLEX NETWORKS MODEL

Through complex networks model, we can obtain a series of data. We can easily find out there are 9784 nodes which we can regard as co-authors. There are 17273 edges which we can regard as cooperations. Alon, Nogam has most significant influence within the network. And there 7072 co-authors connects with him. Therefore it is in the heart of the network. From Table 1, we can conclude that 31 authors' cooperation is over 100, 77 authors' cooperation is between 50 and 100, in addition, 9870 authors' cooperation is below 50. Among them, there are 7 nodes only cooperated with Erdos. They are Rircere, Ceorg Johann Fried, Hans Busolini, Donald Terence Anning, Normann Feldheim, Ervin Oblath, Richard and Colbourn, Charle Sjoseph. Further more, there are 7065 co-authors having only one co-author.

TABLE 1 The cooperation degree data of co-author network

number	name	degree	number	name	degree
10	ALON,NOGAM.	412	500	WORMALD,NICHOLASCHARLES	130
183	HARARY,FRANK*	359	197	HELL,PAVOL	126
84	COLBOURN,CHARLESJOSEPH	225	257	KOSTOCHKA,ALEXANDRV.	126
415	SHELAH,SAHARON	224	193	HEDETNIEMI,STEPHENTRAVIS	125
161	GRAHAM,RONALDLEWIS	220	20	BABAI,LASZLO	121
474	TUZA,ZSOLT	216	350	POMERANCE,CARLBERNARD	120
43	BOLLOBAS,BELA	184	279	LINIAL,NATHAN	119
494	WEST,DOUGLASBRENT	178	498	WINKLER,PETERMANN	114
372	RODL,VOJTECH	172	457	SZEMEREDI,ENDRE	113
245	KLEITMAN,DANIELJ.	164	154	GODDARD,WAYNEDEAN	112
283	LOVASZ,LASZLO	161	303	MCKAY,BRENDANDAMIEN	111
76	CHUNG,FANRONGKING(GRAHAM)	157	285	LUCA,FLORIAN	110
330	ODLYZKO,ANDREWMICHAEL	154	208	HOFFMAN,ALANJEROME	108
389	SALAMON,PETER	151	225	JANSON,SVANTE	108
336	PACH,JANOS	150	440	STINSON,DOUGLASROBERT	106
144	FUREDI,ZOLTAN	145	286	LUCZAK,TOMASZ	105
60	CAMERON,PETERJ.	143	16	ARONOV,BORIS	104
322	NESETRIL,JAROSLAV	143	302	MCELIECE,ROBERTJAMES	101
388	SAKS,MICHAELEZRA	143	106	DIACONIS,PERSIW.	100
434	SPENCER,JOELHAROLD	142	126	FAUDREE,RALPHJASPER,JR.	100
67	CHARTRAND,GARYTHEODORE	133	411	SHALLIT,JEFFREYOUTLAW	100
215	HSU,DERBIAUFRANK	130	……	……	……

TABLE 2 The centers data of the network

number	Name	centers	number	name	centers
10	ALON,NOGAM.	20038.71204	157	GOLOMB,MICHAEL*	23.25531915
183	HARARY,FRANK*	9832.996583	158	GOODMAN,ADOLPHW.*	19.69976661
415	SHELAH,SAHARON	1192.318366	398	SCHERK,PETER*	18.73834756
231	KAC,MARK*	334.9875802	123	EVANS,ANTHONYB.	17.42691384
298	MAULDIN,RICHARDDANIEL	333.3608437	272	LEHNER,JOSEPH	17.41749264
389	SALAMON,PETER	302.5540927	235	KARAMATA,JOVAN*	17.31034483
256	KOREN,ISRAEL	288.4007407	352	PRACHAR,KARL*	15.72727273
75	CHUI,CHARLESKAM-TAI	268.2554733	18	ASHBACHER,CHARLESD.	14
50	BRENNER,JOELLEE*	247.74846	299	MAXSEIN,THOMAS*	13.01568627
319	MURTY,MARUTIRAMPEDAPROLU	159.604793	226	JARNIK,VOJTECH*	11.59003322
58	CACCETTA,LOUIS	150.8287174	502	ZAKS,ABRAHAM	10.92712551
468	TOTIK,VILMOS	136.3565579	375	ROSENBLOOM,PAULCHARLES*	10.56856187
167	GRUNBAUM,BRANKO	128.4501522	506	ZHANG,ZHENXIANG	10.43619048
259	KRANTZ,STEVENGEORGE	121.7998212	484	VAZSONYI,ANDREW*(WEISZFELD,ENDRE)	10
200	HENRIKSEN,MELVIN*	88.85357241	387	SAIAS,ERIC	9.57243745
21	BABU,GUTTIJOGESH	84.58976147	189	HARTTER,ERICH	8
202	HERZOG,MARCEL	84.39426375	110	DOWKER,YAELNAIM	6
295	MARCUS,SOLOMON	71.20766035	32	BENKOSKI,STANLEYJ.	5.136363636
99	DEHEUVELS,PAUL	65.89270624	406	SEGAL,SANFORDLEONARD*	5.070588235
42	BOES,DUANECHARLES	64.39841312	135	FOWLER,THOMASGEORGE	5.052173913
122	ERNE,MARCEL	52.2917183	7	ALAOGLU,LEONIDAS*	4
459	TARSKI,ALFRED*	48.55499448	81	CLARKSON,JAMESANDREW*	4
166	GRUBER,PETERMANFRED	47.18751227	112	DUDLEY,UNDERWOOD	4
326	NEY,PETERE.	45.46016677	332	OFFORD,ALBERTCYRIL*	3
422	SILVERMAN,RUTH	39.53460443	479	VAN	1
507	ZIV,ABRAHAM	26	9387	Aarts,EmileH.L.	1
433	SPECKER,ERNSTP.	23.31578947	……	……	……

In order to discover the key co-authors in the network, we have to analyze the network and discover the highly focused nodes. From Figure 1, it shows that there are 20 cores and 27 nodes having significant influence within the network. The detailed information is shown in table 2. From table 2, we can also conclude that Alon Nogam has published important works or connects important researchers within Erdos1. Betweenness centrality measures the degree of resource control for the actors. Closeness centrality is used to depict the node degree of difficult to reach the other nodes. We can get each Betweenness centrality and closeness centrality of each node by pajek software which is widely used into social network analysis. We can get the data of Betweenness centrality and closeness centrality(see table 3 and table 4). For example, Harary, Frank's betweenness centrality is 0.072516329, Alon, Nogam's closeness centrality is 0.315626758. Which means that he has the most close cooperation with others.

Therefore it is in the heart of the network. Rubinstein, Reuveny's betweenness centrality is 0 and Tucker, Bessie's closeness centrality is 0.000200582. Which means that he has the most loose cooperation with others. Therefore it is in the edge of the network.

TABLE 3 The betweenness centralization data of co-author network

number	name	betweenness	number	name	betweenness
10	ALON,NOGAM.	0.107308184	432	SOS,VERATURAN	0.040487444
183	HARARY,FRANK*	0.09543728	336	PACH,JANOS	0.039108326
161	GRAHAM,RONALDLEWIS	0.069437734	144	FUREDI,ZOLTAN	0.039014166
443	STRAUS,ERNSTGABOR*	0.052829108	245	KLEITMAN,DANIELJ.	0.037869718
330	ODLYZKO,ANDREWMICHAEL	0.048869222	254	KOMLOS,JANOS	0.03715828
474	TUZA,ZSOLT	0.048658208	379	RUBEL,LEEALBERT*	0.036755442
350	POMERANCE,CARLBERNARD	0.045307229	76	CHUNG,FANRONGKING(GRAHAM)	0.036135103
415	SHELAH,SAHARON	0.043915394	84	COLBOURN,CHARLESJOSEPH	0.035351801
434	SPENCER,JOELHAROLD	0.043777011	500	WORMALD,NICHOLASCHARLES	0.030716679
372	RODL,VOJTECH	0.042465495	322	NESETRIL,JAROSLAV	0.030351319
43	BOLLOBAS,BELA	0.041436572	383	RUZSA,IMREZ.	0.030309405
494	WEST,DOUGLASBRENT	0.040614125	389	SALAMON,PETER	0.030105252

Table 3 depicts that 7799 nodes' betweenness centrality is 0, 44 nodes' betweenness centrality is between 100 and 412, 391 authors' betweenness centrality is between 10 and 100, 2302 authors' betweenness centrality is between 2 and 10, and 7048 authors' betweenness centrality is 1.

TABLE 4 Closeness centralization of data

number	name	closeness	number	name	closeness
10	ALON,NOGAM.	0.333247776	……	……	……
161	GRAHAM,RONALDLEWIS	0.327355357	110	DOWKER,YAELNAIM	0.000408831
43	BOLLOBAS,BELA	0.320848525	7	ALAOGLU,LEONIDAS*	0.000306623
144	FUREDI,ZOLTAN	0.320848525	81	CLARKSON,JAMESANDREW*	0.000306623
434	SPENCER,JOELHAROLD	0.318970737	112	DUDLEY,UNDERWOOD	0.000306623
372	RODL,VOJTECH	0.316835168	4598	Deinert,Erhard	0.000292022
183	HARARY,FRANK*	0.314881236	4599	Hofmeister,GerdRolf	0.000292022
474	TUZA,ZSOLT	0.311774521	4600	Stohr,Alfred	0.000292022
283	LOVASZ,LASZLO	0.310328396	4601	Zollner,Joachim	0.000292022
432	SOS,VERATURAN	0.308512514	3292	Auslander,Joseph	0.000245298
322	NESETRIL,JAROSLAV	0.308492893	3293	Friedlander,FriedrichGerard	0.000245298
245	KLEITMAN,DANIELJ.	0.307252043	3294	Lederer,George	0.000245298
330	ODLYZKO,ANDREWMICHAEL	0.307222852	479	VAN	0.000204415
76	CHUNG,FANRONGKING(GRAHAM)	0.307076982	695	Birkhoff,Garrett	0.000204415
457	SZEMEREDI,ENDRE	0.30509721	696	Giese,JohnH.	0.000204415
500	WORMALD,NICHOLASCHARLES	0.304886263	2616	Adams,C.Raymond	0.000204415
336	PACH,JANOS	0.304379267	2617	Randels,W.C.	0.000204415
137	FRANKL,PETER	0.303969168	3306	Pankratz,Alan	0.000204415
443	STRAUS,ERNSTGABOR*	0.301391077	3307	Tucker,Bessie	0.000204415
20	BABAI,LASZLO	0.300058025	9387	Aarts,EmileH.L.	0.000204415

The table 4 shows that 20 authors' closeness centrality is over 0.3, 6541 authors' closeness centrality is between 0.2 and 0.3, 3182 authors' closeness centrality is between 0.1 and 0.2, 41 authors' closeness centrality is 0.2-0.3, and 3182 authors' closeness centrality is blow 0.1.

3 Citation coefficient model

The evaluation of scientific paper is one of the important components in the science and technology evaluation. At present, scientific papers are frequently evaluated on the basis of the journal Impact Factor (IF) in scientific research evaluation. However, scientific papers are sometimes overestimated or underestimated. In order to explore this phenomenon, two bibliometric indicators including the journal IF and the article citation frequency are analyzed, the indicators is also studied.

In this paper, the origin of the journal IF is explored and the calculating method of the journal IF are studied. The elements affecting the journal IF such as citation motives, citation habits, database selection, editing, subject character and journal character are summarized and analyzed. The amending method about journal IF is also compared [8]. The citing and cited articles make up citation networks, and the prospect of the application of citation analysis in the field of science and technology evaluation is outlined in this paper.

In this paper, we study the relationship between the journal IF and the article citation frequency, and taking an academy. The result shows that the journal IF reflects the form value of an article and the article citation frequency reflects the academic value of an article. In the macro-aspect, the evaluation results of these two bibliometric indicators are consistent with each other. In the microaspect, the evaluation results of these two bibliometric indicators are not completely consistent with each other. On the whole, the journal IF and the article citation frequency are complemented in science and technology evaluation. Both the journal IF and the article citation frequency should be used comprehensively for the overall evaluation of scientific papers. We should not only attach importance to the journal's quality, but also pay more attention to the article citation frequency, the citation journals, and the methods of qualitative evaluation such as peer review are supplemented with the methods of quantitative evaluation.

Journal articles influence evaluation as every Chinese institution of higher learning, scientific research institutes and government agencies to evaluate a unit or individual work important basis. It has the vital significance to correctly evaluate the efficacy of the scientific research work, create a good scientific research atmosphere, inspire researcher's academic creative potential and improve their ability of the constant innovation of academic research [9]. Library information science has been committed to evaluate the influence of journal articles through literature metrology research. We get which is one of the most influential authors through the influence of the literature on comparative study on the evaluation method of measurement. For some comprehensive or field of studying, studying areas are broad so that reference rate is quite high. Some papers easily have higher influence, such as biological and chemical class papers. Although influence factors to some extent can characterize the quality of its academic strengths and weaknesses, the impact factors and the academic quality is a nonlinear proportional relationship.

3.1 THEORETICAL SOURCES

3.1.1 Paper influence evaluation methods

Paper articles influence evaluation method based on the literature metrology in general can be divided into two categories: the first kind is based on the citation evaluation index, the method mainly includes papers were cited. The second type is based on the evaluation index of papers in papers and its derivative index, the method mainly includes the paper journal impact factor, the influence of the paper the factor scores of average and so on[10].

3.1.2 Paper influence evaluation and periodical evaluation

The academic level has two main factors. Journal of the principle and policy is the first main factor .The other is the level of the journal. This article mainly discusses the relationship and influence of papers. After the journal paper are published, if the level is higher, the influence is bigger, and it may be referenced by other scholars, so that the paper cited situation at this time can measure the quality of the paper.

Impact factor is not one of the most objective evaluation criteria of paper influence. Generally the higher impact factor has the greater influence. For some comprehensive or field of studying, studying areas are broad so that reference rate is quite high.

3.2 THE CITATION COEFFICIENT MODE

We introduce the citation coefficient model. The model formula is as follows:

$$C = ((N + \sum_{i=1}^{m} a_i / b_i) / N_j) \cdot (IF / JIF), \quad (3)$$

where C represents the paper citation strength. N represents the numbers of paper referenced by others. IF represents the paper impact factor. JIF represents the average impact factor of academic paper. m represents the numbers of paper self-priming article. N_j represents the numbers of paper referenced by others on average each year. a_i represents the number of quotation i cited by others. b_i represents the number of quotation[11].

$$L(p) = \{1/[0.5n(n+1)]\} \sum_{i \geq j} d(p_{ij}), \quad (4)$$

where L stands for the average path length. And it said any two nodes in the network see average shortest path length. We can depict the connection between network nodes from overall length, which is suitable from node i to node j. d_{ij} is geometric distance. p is the evolution of probability [12].

The meaning of citation coefficient: By using the relative indexes, the paper citation coefficient can effectively avoid discipline unfair phenomenon in the evaluation of the paper.

3.3 THE CONCLUSION OF CITATION COEFFICIENT MODEL

By retrieving the ISI JCR database collection of data and using paper citation coefficient model to the set of foundational papers in the emerging field of network science which is composed with 16 papers, we obtained the paper cited coefficient. At the same time, we can also get the betweenness of 16 papers, which due to the influence of the acquisition data appears inevitably. By the analysis of some degree of betweenness, it avoids data losing. We can get no.2 paper citation coefficient up to 152.0235, followed by no.14, 104.8957, and no.4, 77.62738 by observing the table 5. The higher the centricity, the node is bigger. We can see that the 11th papers node degree is highest, followed by 4. Considering that, paper 4 is the most influential in the network science. See Figure 2.

We get the network average path length, which is 5.258. The average path length is shortest path between any two nodes in the network. It reflects the properties of the network. Small average path length Erdos1 network has a typical the Small-World phenomenon.

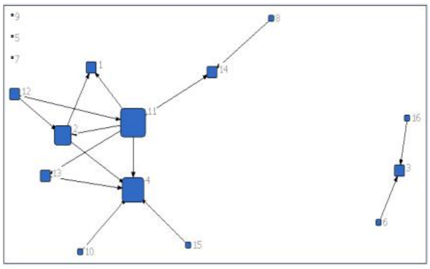

FIGURE 2 The paper's relative influence in the network

TABLE 5 The evaluation of the paper's influence

no.	I F	cite	year	ave. cite/y	JIF	IF/JIF	The art Num.	C
2	44.982	13250	2002	1104.167	0.960	46.856	215.000	152.024
14	38.597	21688	1998	1355.500	1.462	24.400	27.000	104.896
4	31.027	18843	1999	1256.200	1.473	21.064	20.000	77.627
11	5.952	10616	2003	965.091	0.766	7.770	16.000	22.077
8	38.597	1246	2000	89.000	1.252	30.828	10.000	9.267
10	9.737	2748	2001	211.385	0.965	10.090	30.000	6.614
12	28.43	680	2007	97.143	1.473	19.301	95.000	6.258
13	31.027	835	2002	69.583	1.331	23.311	10.000	5.695
15	42.761	33	2011	11.000	1.020	41.923	117.000	3.120
6	3.381	763	2000	54.500	1.069	3.163	41.000	0.634
16	3.381	498	1996	27.667	1.046	3.232	68.000	0.397
9	2.313	1532	2001	117.846	2.883	0.802	19.000	0.308
1	0.322	4531	1959	82.382	0.559	0.576	8.000	0.162
3	0.875)	1956	1987	72.444	2.101	0.416	15.000	0.105
7	4.375	16	1979	0.457	1.892	2.312	7.000	0.101
5	0.424	25	2006	3.125	2.000	0.212	16.000	0.011

4 Application and extension

4.1 MEASURE OF INFLUENCE FOR UNIVERSITY, DEPARTMENT, OR A JOURNAL

By the analysis of the above, we know that the model can be extended to determine the role or influence measure of an individual network researcher. The data need to be collected by the network or some data collecting web-site. As a university, which the page views high and school BBS good comment is an influential school in the network science. Similarly, if a department has high school page views, faculties BBS good comment, we will call this faculties having a big influence in the network science. And in terms of a journal, if it is high referenced and downloaded we will call that the journal is a great influence in the network science. According to a series of impact factor and the model, we analyze their specific effects.

4.2 ON A COMPLETELY DIFFERENT SET OF NETWORK INFLUENCE DATA

This algorithm can be implemented to different network data, for instance, influential songwriters, music bands, performers, movie actors, directors, movies, TV shows, columnists, journalists, newspapers, magazines, novelists, novels, bloggers, tweeters, or any data set we care to analyze. We may wish to restrict the network to a specific genre or geographic location or predetermined size. As long as we have the relevant data, through our model can calculate a specific network influence of each node.

We collected two different types of data in IMDb. The two types of data are movies' background information and users' data. Movies' background data include the category of the movie, actors, directors, box office, and duration. The users' data include rating, review and rating for other users' comments.

The categories of movies will be treated as nodes in the community. If the two movies both reviewed by a user, then we will create an edge between the two nodes.

The collected raw data will be preprocessed to transform into analyze-able information. Basically, data will be pre-processed into two different types of information including film node and user node.

Therefore, we can easily predict the box office from the comments of key users. If a movie can't receive enough positive comments and reviews, then it usually can't achieve idea box office. In another word, we can observe the comments and reviews from the key users to roughly predict the box office of a movie. The movie theater can also use this information to make decision about the duration of a movie.

4.3 THE SCIENCE, UNDERSTANDING AND UTILITY

Our models can play an important role in discussing the science, understanding and utility of modeling within networks. They could help individuals, organizations, nations, and society use influence methodology to improve relationships, conduct business, and make wise decisions. For instance, at the individual level, we could use all the cooperations' network effects data to make up the form by referring to our models and using pajek software. Make the network data diagram and find out the most influential partners to cooperate. So we can boost our mathematical influence as rapidly as possible. Or we can use our models and results to help decide on a graduate school or thesis advisor to select for our future academic work.

5 Conclusion

The Influence measurement model used in networks to measure influence and impact is based on the complex networks system theory. This takes some skilled data extraction and modeling efforts to obtain the correct set of nodes and their links. In the model combining qualitative analysis and quantitative description, which we can obtain who has significant influence within the network. We also build another type of influence measure to compare the significance of a research paper by analyzing the important works that follow from its publication. This method by the impact through articles factor and as references times decision. Our research method can be applied to other areas, for instance, university, department, or a journal in network science. Using our models and results can help to decide on a graduate school or thesis advisor to select for your future academic work.

Acknowledgments

This work was supported by NNSFC(11401041), the Shandong Province Science Project in Statistics of China(2014Y207), the Shandong Province University Research and Development Foundation of China(J12L59) and by Binzhou University Science Foundation of China (2013Y02, BZXYL1303).

References

[1] Borgatti, S 1995 Centrality and AIDS Connections 18(1), 112-14
[2] Newman, M E J 2001a Physical Review E 64(1), 16131
[3] Newman, M E J 2001b Physical Review E 64(1), 16132
[4] Newman, M E J 2001c Academy of Sciences of the USA 98(2), 404
[5] Scott, J 1991 Social network analysis: A handbook. London: Sage
[6] Freeman, L C 1979 Social Networks 1(3), 215-39
[7] Freeman, L C 1980 Quality and Quantity 14(4), 585-92
[8] Seglen P O 1997 Bri. Med. J 314,498-502
[9] Wang X N, He M, Guo J J and Han D Y 2004 Science and technology management 19,15-18
[10] Alireza A, Liaquat H, Shahadat U and Kim J R R 2011 Scientometrics 89(2),687–710
[11] Zhong W Y and Chen Y P 2011 Information science 29(5) 706-12 (in Chinese)
[12] Xiong, H H 2012 South China University of Technology: Guangzhou (in Chinese)

Authors

Liguo Huang, 1981.01, China

Current position, grades: lecturer in Binzhou university of China, MS degree
University studies: MS degree in operations research and control theory from the College of Information Science and Engineering, Shandong University of science and Technology, China in June 2006.
Scientific interest: data mining, complex system modeling and optimization method
Publications: 5
Experience: lecturer of the Department of Mathematics in Binzhou University, Shandong, China.

Lufang Mi, 1975.03, China

Current position, grades: associate professor in Binzhou university of China, doctor
University studies: PhD degree in basic mathematics from the school of Mathematical Sciences, Fudan University, China in July 2013.
Scientific interest: systems analysis, stability theory and Hamilton dynamical system
Publications: 11
Experience: associate professor of the Department of Mathematics in Binzhou University, Shandong, China.