

A systematic literature review of the Prisoner’s Dilemma; collaboration and influence.

Nikoleta E. Glynatsi, Vincent A. Knight

1 Timeline

2 Analysing a large corpus of articles

The focus of the paper has been the academic publications on the topic the iterated prisoner’s dilemma. Whilst in Section 1 we covered several publications of specific interest and manually partitioned the literature in different sections, in the second part of this paper we analyse the publications using a large dataset of articles. The data collection process is covered in Section 2.1 and a preliminary analysis of the data is conducted in Section 2.2. In Section 2.3, we present the methodology which will be used to analyse the authors relationships. In summary, we will be using graph theoretical methods to ascertain the level of collaborative nature of the field and identify influence, relative to:

- Two other sub fields of game theory: auction games [11] and the price of anarchy [17].
- A temporal analysis.

Finally in Section 2.4, the results of the analysis are presented.

2.1 Data Collection

Academic articles are accessible through scholarly databases and collections of academic journals. Several databases and collections today offer access through an open application protocol interface (API). An API allows users to query directly a journal’s database and bypass the user interface side of the journal. Interacting with an API has two phases: requesting and receiving. The request phase includes composing a url with the details of the request. For example, [http://export.arxiv.org/api/query?search_query=abs:prisoner’sdilemma&max_results=1](http://export.arxiv.org/api/query?search_query=abs:prisoner'sdilemma&max_results=1) represents a request message. The first part of the request is the address of the API we are querying. In this example the address corresponds to the API of arXiv. The second part of the request contains the search arguments. In our example we are requesting for a single article that the word ‘prisoners dilemma’ exists within it’s title. The format of the request message is different from API to API. The receive phase includes receiving a number of raw metadata of articles that satisfied the request message. The raw metadata are commonly received in extensive markup language (xml) or Javascript object notation (json) formats [14]. Similarly to the request message, the structure of the received data differs from journal to journal.

The data collection is crucial to this study. To ensure that this study can be reproduced all code used to query the different APIs has been packaged as a Python library and is available online [12]. The software could be used for any type of projects similar to the one described here, documentation for it is available at: <http://arcas.readthedocs.io/en/latest/>. Project [12] allow us to collect articles from a list of APIs by specifying just a single keyword. Four prominent journals in the field and a pre print server were used as sources to collect data for this analysis. Those were PLOS, Nature, IEEE, Springer and arXiv.

A series of search terms were used to identify relevant articles. The terms used to collect the main data set were,

- “prisoner’s dilemma”,
- “prisoners dilemma”,
- “prisoners evolution”,
- “prisoner dilemma”,
- “prisoner game theory”

and articles for which any of these terms existed within the title, the abstract or the text are included in the analysis. More specifically, 23% of article considered here were included because any of the above terms existed within the abstract, 50% within the main text and 27% within the title. As will be described in Section 2.2, two other game theoretic sub fields were also considered in this work, auction games and the price of anarchy. For collecting data on these sub fields the search terms used were “auction game theory” and “price of anarchy”. The three data sets are archived and available at. Note that the latest data collection was perform on November 2018.

2.2 Preliminary Analysis

A summary of each of the three data sets used is presented in this section. The three data sets are:

- The main data set which contains articles on the prisoner’s dilemma.
- A secondary data set which contains article on auction games.
- A secondary data set which contains articles on the price of anarchy.

The main data set and the main focus of this analysis is [ref]. It consists of 3089 articles with unique titles. In case of duplicates the preprint version of an article (collected from arXiv) was dropped. Of these 3089 article, 89 have not been collected from the aforementioned APIs. These articles were of specific interest and manually added to the dataset throughout the writing of Section 1. A more detailed summary of the articles’ provenance is given by Table 1. Only 3% of the data set consists by articles that were manually added and 33% of the articles were collected from arXiv. The rest four journals have contributed 9%-21% percent of the articles.

provenance	# of Articles	Percentage
Manual	89	2.81
IEEE	295	9.31
PLOS	482	15.22
Springer	572	18.06
Nature	673	21.25
arXiv	1056	33.34

Table 1: Articles’ provenance for main data set.

The average number of publications was calculated for the entire dataset and for each provenance. The average number of publications is denoted as, $\mu_P = \frac{N_A}{N_Y}$, where N_A is the total number of articles and N_Y is the years of publication. The years of publication is calculated as the range between 2018 and the first published article, for each provenance, within the data. These averages are summarised in Table 2. Overall an average of 49 articles are published per year on the topic. The most significant contribution to this appears to be from arXiv with 16 articles per year, followed by Nature with 10 and Springer with 9.

Though the average publication offers insights about the publications of the fields, it remains a constant number. The data we are handling here is a time series between 1950, when the game was introduced, and 2018 (Figure 1). Two observations can be made from Figure 1.

Av. publication	
IEEE	5.0
PLOS	8.0
Springer	9.0
Nature	11.0
arXiv	16.0
Overall	49.0

Table 2: Average publication for main data set.

1. A steady increase to the number of publications since the 1980s and the introduction of computer tournaments.
2. A decrease in 2017-2018. This is due our data set being incomplete. Articles that have been written in 2017-2018 have either not being published or have are not retrievable by the APIs yet.

These observations can be confirmed by studying the time series. Using [6], an exponential distribution is fitted to the data from 1980-2016. The perfect fitting proves that since 1980 there has been an exponential increase to the number of publications till 2016 (Figure 2). The fitted model can also be used to project the behaviour of the field for the next 5 years. The forecasted periods are plotted in Figure 3 and their exact values are given by Table 3. Though the time series has indicated a slight decrease we can see that the model forecasts that the number of publications will keep increasing, thus indicating that the field of the iterated prisoner’s dilemma still attracts academic attention.

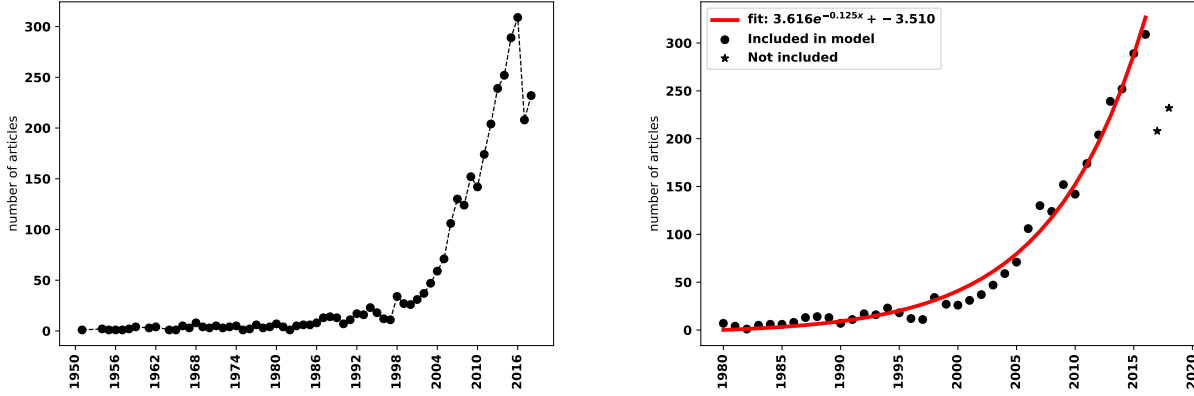


Figure 1: Line plot; # of articles published on the PD Figure 2: Scatter plot; # of articles published on the PD 1980-2019.

Forecast	
2017	371.0
2018	421.0
2019	478.0
2020	542.0
2021	615.0

Table 3: Forecasting the number of publications over the next 10 years.

Moreover, two sub fields of game theory have been chosen for this work; auction game and the price of anarchy.

- Auction theory is a branch of economics which deals with how people act in auction markets and researches the properties of auction markets. Game theory is being used for years to study auctions and the behaviour of the bidders [18]. The earliest entry in our data set [ref] goes back to 1974 (Figure 4).

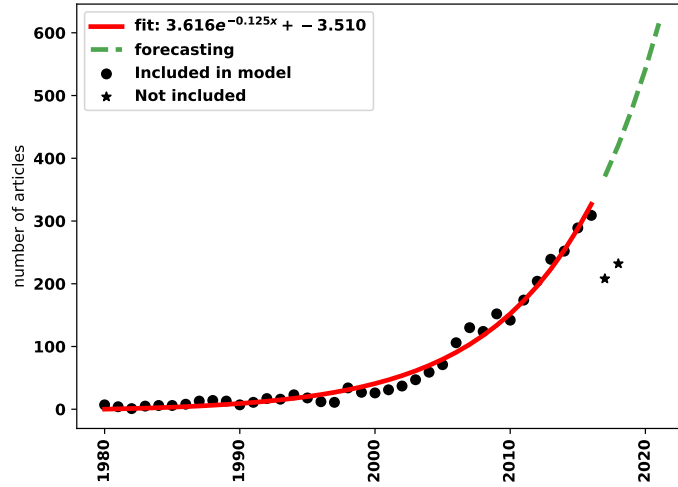


Figure 3: Forecast for 2017-2022

- Price of Anarchy is a concept in economics and game theory which measures how the efficiency of a system degrades due to selfish behaviour of its agents. There is a variety of such measures however the price of anarchy has attracted a lot of attention since its informal introduction in 1999 by [7]. The first entry in the data set [ref] is a year later in 2000 (Figure 5).

A summary of both data sets collected on both topics, in comparison to that of [ref], is given by Table 4.

The iterated prisoner’s dilemma and auction theory are very well studied topics that have been publicising for decades. A large number of articles have been collected for both topics, 3089 and 3444 respectively. Though, auction games have a larger number of articles, the iterated prisoner’s dilemma has almost 300 more authors. Auction games have an overall average publication of 93 articles compared to the prisoner’s dilemma with 49. The 50% of articles for [ref] have been collected from the pre print server arXiv and no articles have been published in PLOS.

Compared to these two topics the price of anarchy is a fairly recent one. Only a total of 747 articles have been collected, however it has a large number of 1229 authors. Meaning that on average each paper had had at least two authors. It has an overall average publication of 39 articles and the biggest contribution has been made from Springer. Note that no article have been added manually for the data sets for the two extra sub fields.

	Num. Articles	Num. Authors	Manual (%)	PLOS (%)	Nature (%)	Springer (%)	IEEE (%)	arXiv (%)	Av. Publication
Prisoner’s Dilemma	3089	5092	2.88	15.6	21.79	18.52	9.55	34.19	49.0
Auction Games	3444	4770	-	-	5.89	37.63	7.46	51.36	93.0
Price of Anarchy	746	1227	-	1.74	24.66	38.07	30.70	8.85	41.0

Table 4: Measures of all three data sets.

2.3 Methodology

As discussed in [19], bibliometrics or the statistical analysis of published works (originally described by [15]) have been used to support historical assumptions about the development of fields [16], identify connections between scientific growth and policy changes [3], and investigate the collaborative structure of an interdisciplinary field [10]. Most academic research is undertaken in the form of collaborative effort and as [8] points out, it is rationale that two or more people have the potential to do better as a group than individually. Collaboration in groups has a long tradition in experimental sciences and it has been proven to be productive according to [4]. The number of collaborations can be very different between research fields and understanding how collaborative a field is, it is not always an easy task. Several studies tend to consider academic citations as a measure for these things. A blog post published in Nature [13] argues that depending on citations can often be misleading because the true number of citations can not be known. Citations can be missed due to

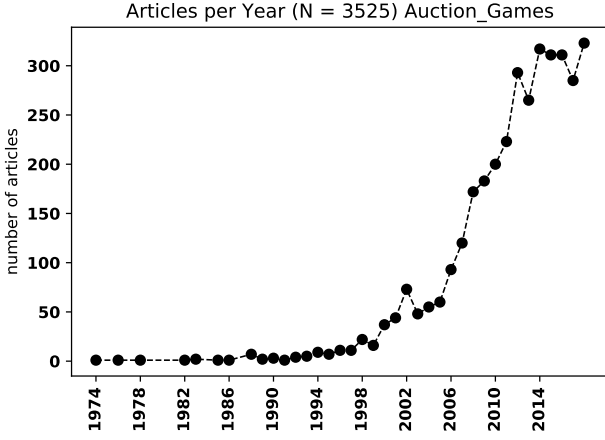


Figure 4: Line plot; # articles published on auction games 1974-2018.

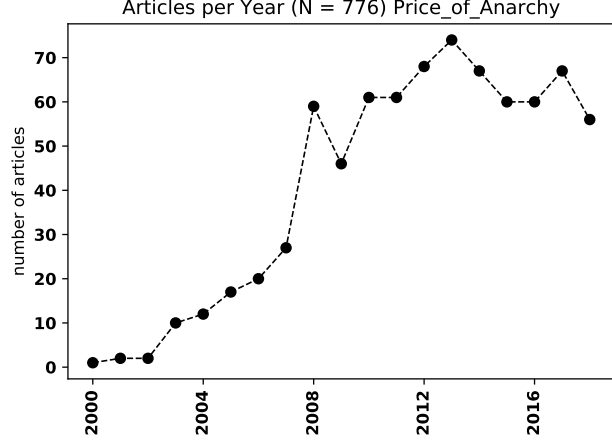


Figure 5: Line plot; # articles published on the price of anarchy.

data entry errors, academics are influenced by many more papers than they actually cite and several of the citations are superficial.

A more recent approach to measure collaborative behaviour is to use the co authorship network, as described in [10]. Using this approach has many advantages as several graph theoretic measures can be used as proxies to explain authors relationship. In [10], they analyse the development of the field “evolution of cooperation” using this approach. The topic “evolution of cooperation” is a multidisciplinary field which also includes a large number of publications on the prisoner’s dilemma. In this paper we build upon the work done by [10] and extend their methodology. Though in [10], they considered a data set from a single source, Web of Science, our data have been collected from five different sources. Moreover, the collaborative results of our analysis are compared to those of two different sub fields. Co authorship networks have also been used in [19] for classifying topics of an interdisciplinary field. This was done using centrality measures, which will be covered below, here we use centrality measures in order to understand the influence an author can have and can receive by being part of an academic group.

The relationship between the authors within a field will be modelled as a graph G with a set V_G of nodes and E_G of edges. The set V_G represents the authors and an edge connects two authors if and only if those authors have written together. The co authorship network is constructed using the main data set described in Section 2.2 and the open source package Networkx [5]. The prisoner’s dilemma network is denoted as G_1 where the number of unique authors $|V(G_1)|$ is 5092 and $|E(G_1)| = 9882$. Note that the names of all authors names were formatted as their last name and first initial (i.e. Martin A. Nowak to M.Nowak). This was done to avoid errors such as Martin A. Nowak and Martin Nowak, being treated as a different person. Networkx will also be used the following section to conduct our analysis.

Collaborativeness, will be analysed using measures such as, isolated nodes, connected components, clustering coefficient, modularity and average degree. These measures allow us to understand the number of connections authors can have and how strongly connected these people are. The number of isolated nodes is the number of nodes that are not connected to another node, thus the number of authors that have published alone. The average degree denotes the average number of neighbours for each nodes, i.e. the average number of collaborations between the authors.

A connected component is a maximal set of nodes such that each pair of nodes is connected by a path. We are interested in the number of connected components but also the size of the largest connected component in the network. The size of largest connected component represents the scale of the central cluster of the entire network, as it will be discussed in the analysis section. Clustering coefficient and modularity are also calculated. Clustering coefficient defined by,

$$\text{clustering coeff.} = \frac{3 \times (\text{number of triangle on the graph})}{\text{number of connected triples of nodes}}$$

is a local measure of the degree to which nodes in a graph tend to cluster together in a clique. It is precisely the probability

that the collaborators of an author also write together. In comparison, modularity is a global measure designed to measure the strength of division of a network into modules. A high value of modularity corresponds to a structure where authors mainly write in groups and interact less with the rest of the network. We will be using the Louvain method described in [2] to calculate modularity.

Furthermore, the second part of the analysis focuses on the study of influence. Networks are commonly dominated by one person who controls information flow and people that receive a great amount of information due to their position. In this paper we aim to understand two things, (1) which people control the flow; as in which people influence the field the most and (2) which are the authors that gain the most from the influence of the field. To measure these concepts we will be using graph theoretic metrics, more specifically centrality measures. Centrality measures are often used to understand different aspects of social networks [9]. In order to achieve that two centrality measures that have been chosen were closeness and betweenness centrality.

1. In networks some nodes have a short distance to other nodes and consequently are able to spread information on the network very effectively. A representative of this idea is **closeness centrality**, where a person is seen as centrally involved in the network if it requires only few intermediaries for contacting others and thus is structurally relatively independent. Here, we define this as influence. Authors with a high value of closeness centrality, are the authors that spread scientific knowledge easier on the network and we say that they have high influence.
2. Another centrality measure is the **betweenness centrality**, where the determination of an author's centrality is based on the quotient of the number of all shortest paths between nodes in the network that include the regarded node and the number of all shortest paths in the network. In betweenness centrality the position of the node matters. Nodes with a higher value of betweenness centrality are located in positions that a lot of information pass through them, we define this as the gain from the influence, thus these authors gain the most from their networks.

In the next section we will be using all the metrics discussed here to provide insights on the field.

2.4 Analysis of co authorship network

As mentioned previously, G_1 denotes the co authorship network of iterated prisoner's dilemma. The open source software Gephi [1] has been used to plot the networks of this work, more specifically G_1 is given by Figure. It is evident that our network is disjoint, which is only natural as many authors write academic articles on their own. More specifically, a total of 157 authors, have had single author publications, which corresponds to the 0.033 (%) of authors in G_1 . There are a total of 1027 connected components and the largest one has a size of 1457. The largest connected component is shown in Figure. The network as a clustering coefficient of 0.685, thus you are 68% likely to write with a collaborators co author. Overall the networks have an average degree of 4.194, meaning that the average publication on the field has 4 authors. The distribution of the degrees, Figure 6, indicates that though the average is 4 there are authors with far more connections, the largest one being around 30.

How does these compare to other fields and more specifically to other fields of game theory? A summary of the two graphs, which be denoted as G_2 for auction games and G_3 for the price of anarchy, are given by Figure. A summary of metrics and for all three co authorship networks is given by Table 5, and the following remarks can be made:

	# Connected Components	# Edges	# Isolated	# Nodes	% Isolated	Av. Degree	Clustering	Largest cc	Modularity
Prisoner's Dilemma	1027	9882	157	4713	0.033	4.194	0.685	1457	0.930
Auction Games	947	7740	210	4567	0.046	3.390	0.595	2076	0.892
Price of Anarchy	193	1910	4	1072	0.004	3.563	0.705	421	0.947

Table 5: Network metrics for G_1, G_2, G_3 .

- Comparing to another well studied topic (G_2), the co authorship network G_1 appears to be more modular. This is due the high values of modularity, connected components and clustering coefficient. Authors in G_1 tend to write more in teams (modularity .931 > .892), separated from the main cluster and it's more likely for them to create smaller clusters of 3 (clustering .685 > .595). On the other hand, G_2 has main cluster of bigger size (2079 > 1457), suggesting a more chained community than G_1 .

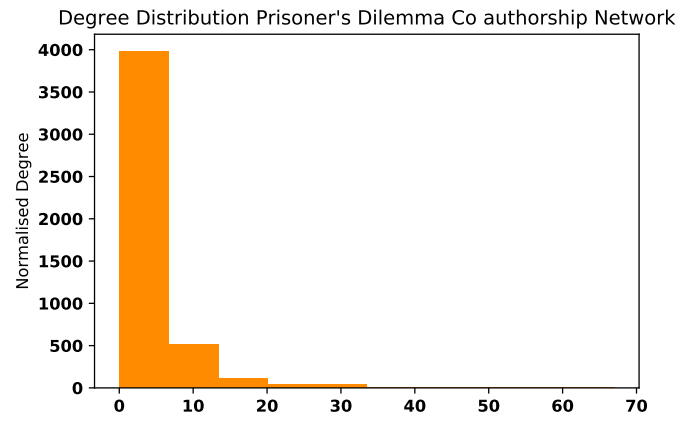


Figure 6: Degree distribution for network G_1 .

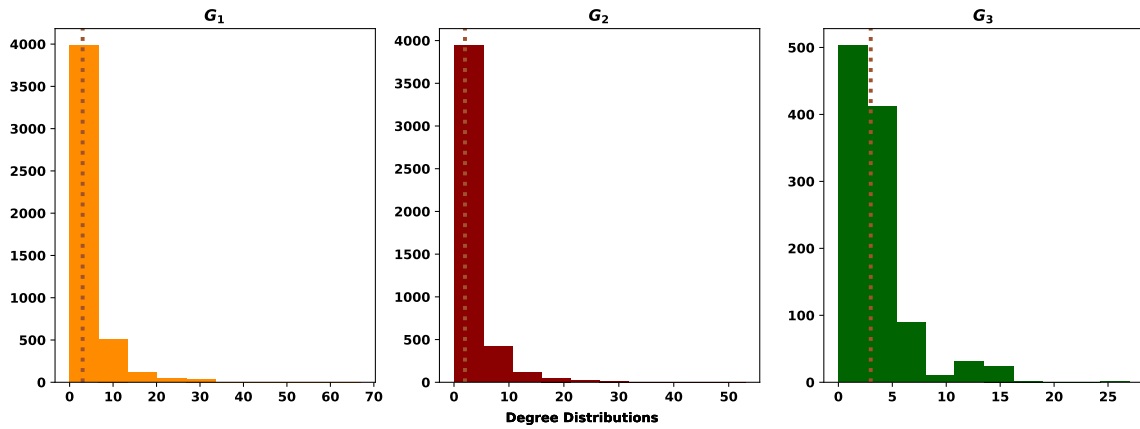


Figure 7: Degree distribution for networks G_1G_3 .

- In the more recent topic price of anarchy (G_3) there are hardly any people that have published a paper alone. There is already a small community that is chained in the main cluster of 421 authors. The network is the most modular of all three (modularity= 0.948). The value of modularity and the percentage of isolated authors, suggest that authors are either in the main cluster or in smaller connected components with a high clustering probability (clustering coeff.= 0.7).
- Shown in Figure 7 are the degree distributions of all three networks. It has been statistically tested, and none of the distributions are normally distributed. More specifically, all three distributions are very skewed to the left side. Though the average degree is near 4 the medians for $G_1 - G_3$ are 3, 2 and 3. Based Kruskal Wallis test $p < 0.05$ there is significant difference in the medians. In G_1 and G_2 there are cases of high degree (> 20) but this could be an affect of the size of the data, networks and subsequently the size of the main clusters.

In order to study the growth collaborative behaviour we construct the network cumulatively, with a year (of publication) interval. There are a total of 64 subgraphs, for 64 periods starting from 1954. All the collaborative metrics have been calculated for each period and they are given in Table 6.

Similar to the results of [10], we can observe that a rapid growth to network over time. The number of nodes, the number of connected components and the degree all increase over time. Note that since the beginning there has been a high value of modularity to G_1 .

In Figure 8 the normalised number of nodes which is calculated by dividing by the total number of nodes in G_1 (5092). A steep increase is occurring around 2000. This was briefly comment upon in [10]. No compare whether this is occurring only in G_1 we plot the normalised number of nodes for all three graphs in Figure 9, which argues with that. There is a perfect match in the growth of the networks over time since the 2000. This is evidence that something else shocked the academic community around that time.

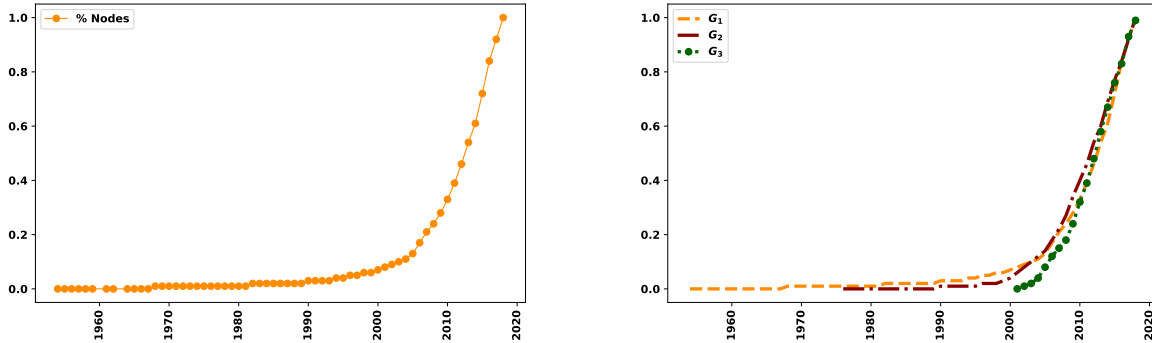


Figure 8: Normalized # Nodes over time for G_1 . Figure 9: Normalized # Nodes over time for $G_1 - G_3$.

The average degree for $G_1 - G_3$ has been reported to be 4. The change of the average degree over time differs for the three graphs, as shown in Figure 10. In G_1 the first publications were single author ones. There is a steep increase to a degree of 1 just before the 1970s and the average degree is steadily increasing since then. A similar trend appears in G_2 as well, in G_3 however there has been a sharp increase since the beginning of the field, with a minor decrease.

The next results were presenting here are on the centrality measures. For G_1 the most central author based on closeness and betweenness are given by Tables 11 and 12 respectively. All centrality measures can have values between 0 and 1. As shown in Table 12, the betweenness centrality of the top authors is rather small. The author that gains the most from the network is M. Perc, who's work we briefly discussed in Section. Perc is connected to 59 nodes, and he the second most influent person in G_1 . The values of closeness centrality appear to be higher than those of betweenness.

The distributions for both centralities are plotted in Figures 13 and 14. Several remarks are made from the centralities' distributions. Regarding between centrality, none of the three distribution is normally distributed. The medians of the distributions are compared and found to be statistically different based on a Kruskal Wallis test. That is because authors in G_3 gain more from being in the network than others in G_1 and G_2 . Even so, the distributions are skewed to the left around 0. That implies that in all three networks, authors do not gain much from the influence of their fields.

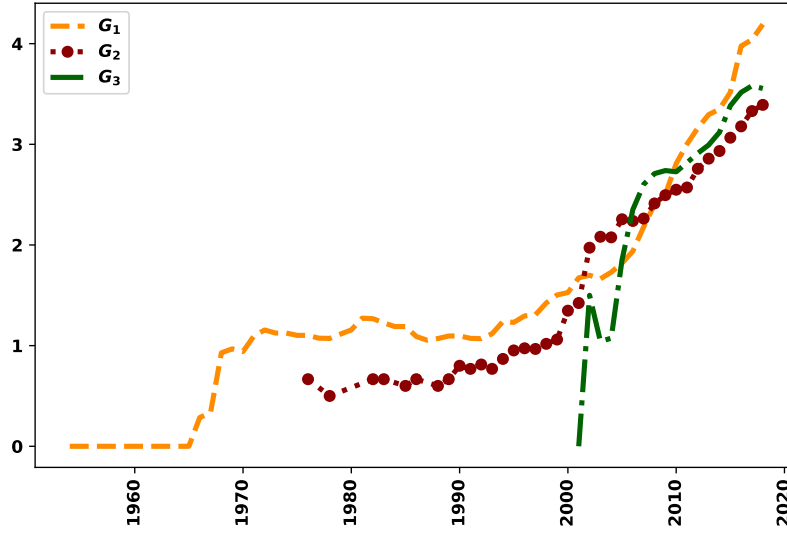


Figure 10: Degree distribution for network G_1 .

	Name	Closeness
1	L. Wang	0.096421
2	M. Perc	0.095338
3	Y. Zhang	0.094736
4	Z. Wang	0.094260
5	Y. Chen	0.090542
6	J. Wang	0.089248
7	X. Wang	0.088720
8	Y. Liu	0.088546
9	J. Zhang	0.088181
10	L. Zhang	0.087923

	Name	Betweenness
1	M. Perc	0.018903
2	Z. Wang	0.015962
3	L. Wang	0.014842
4	Y. Zhang	0.013178
5	M. Nowak	0.011588
6	H. Wang	0.008221
7	Y. Chen	0.008070
8	Y. Li	0.007993
9	Y. Moreno	0.007132
10	N. Masuda	0.006087

Figure 11: Authors that gain the most influence in G_1 .

Figure 12: Ten most influenced authors in G_1 .

	# Connected Components	# Isolated	# Nodes	% Isolated	Av. Degree	Clustering	Largest cc	Modularity
1954 - 1950	3	3	3	1.00	0.00	0.00	1	-
1954 - 1955	2	2	2	1.00	0.00	0.00	1	-
1955 - 1956	3	3	3	1.00	0.00	0.00	1	-
1956 - 1957	4	4	4	1.00	0.00	0.00	1	-
1957 - 1958	6	6	6	1.00	0.00	0.00	1	-
1958 - 1959	7	7	7	1.00	0.00	0.00	1	-
1959 - 1961	7	7	7	1.00	0.00	0.00	1	-
1961 - 1962	8	8	8	1.00	0.00	0.00	1	-
1962 - 1964	9	9	9	1.00	0.00	0.00	1	-
1964 - 1965	10	10	10	1.00	0.00	0.00	1	-
1965 - 1966	12	10	14	0.71	0.29	0.00	2	0.5
1966 - 1967	15	12	18	0.67	0.33	0.00	2	0.666667
1967 - 1968	19	13	28	0.46	0.93	0.16	5	0.591716
1968 - 1969	21	15	31	0.48	0.97	0.16	6	0.533333
1969 - 1970	23	16	34	0.47	0.94	0.15	6	0.585938
1970 - 1971	26	16	42	0.38	1.10	0.26	6	0.763705
1971 - 1972	27	16	45	0.36	1.16	0.31	6	0.801775
1972 - 1973	29	17	48	0.35	1.12	0.29	6	0.814815
1973 - 1974	29	17	48	0.35	1.12	0.29	6	0.814815
1974 - 1975	30	18	49	0.37	1.10	0.29	6	0.814815
1975 - 1976	30	18	49	0.37	1.10	0.29	6	0.814815
1976 - 1977	33	19	54	0.35	1.07	0.26	6	0.837099
1977 - 1978	34	19	56	0.34	1.07	0.25	6	0.846667
1978 - 1979	36	21	61	0.34	1.11	0.25	6	0.854671
1979 - 1980	37	22	64	0.34	1.16	0.28	6	0.866326
1980 - 1981	37	22	66	0.33	1.27	0.29	6	0.85941
1981 - 1982	40	24	71	0.34	1.27	0.32	6	0.873086
1982 - 1983	43	26	75	0.35	1.23	0.30	6	0.878072
1983 - 1984	46	28	79	0.35	1.19	0.28	6	0.882752
1984 - 1985	46	28	79	0.35	1.19	0.28	6	0.882752
1985 - 1986	54	35	88	0.40	1.09	0.26	6	0.887153
1986 - 1987	58	38	93	0.41	1.05	0.24	6	0.891295
1987 - 1988	61	39	99	0.39	1.07	0.26	6	0.903524
1988 - 1989	70	43	115	0.37	1.10	0.26	6	0.921643
1989 - 1990	75	45	124	0.36	1.10	0.26	6	0.930363
1990 - 1991	84	50	138	0.36	1.07	0.26	6	0.939007
1991 - 1992	87	52	144	0.36	1.07	0.25	6	0.942486
1992 - 1993	95	54	161	0.34	1.12	0.26	6	0.951852
1993 - 1994	110	62	192	0.32	1.24	0.33	6	0.95996
1994 - 1995	118	64	208	0.31	1.23	0.31	6	0.964111
1995 - 1996	127	68	227	0.30	1.30	0.32	6	0.966357
1996 - 1997	137	71	248	0.29	1.31	0.34	6	0.970812
1997 - 1998	138	55	274	0.20	1.42	0.36	6	0.976778
1998 - 1999	139	49	290	0.17	1.50	0.38	6	0.979126
1999 - 2000	156	54	330	0.16	1.53	0.40	6	0.981765
2000 - 2001	157	42	356	0.12	1.67	0.42	9	0.980204
2001 - 2002	178	48	405	0.12	1.70	0.43	9	0.982271
2002 - 2003	208	54	470	0.11	1.66	0.43	9	0.985315
2003 - 2004	230	54	539	0.10	1.73	0.44	10	0.986425
2004 - 2005	256	54	629	0.09	1.82	0.47	20	0.985488
2005 - 2006	309	62	790	0.08	1.94	0.50	22	0.988459
2006 - 2007	361	72	973	0.07	2.18	0.53	26	0.986283
2007 - 2008	404	82	1151	0.07	2.40	0.55	40	0.984604
2008 - 2009	451	94	1324	0.07	2.50	0.55	70	0.979801
2009 - 2010	503	106	1541	0.07	2.81	0.56	195	0.965391
2010 - 2011	570	116	1848	0.06	3.00	0.59	259	0.96551
2011 - 2012	637	120	2190	0.05	3.16	0.62	377	0.96145
2012 - 2013	700	131	2532	0.05	3.30	0.63	498	0.956573
2013 - 2014	769	139	2869	0.05	3.35	0.64	651	0.948369
2014 - 2015	857	148	3395	0.04	3.52	0.65	845	0.942306
2015 - 2016	935	155	3961	0.04	3.98	0.67	1116	0.940605
2016 - 2017	978	157	4321	0.04	4.04	0.68	1253	0.936951
2017 - 2018	1029	157	4706	0.03	4.19	0.68	1456	0.92802

Table 6: Collaborativeness metrics for cumulative graphs.

On the other hand, closeness distributions have more variation. The following observations are made from the distributions:

- Neither are normally distributed and there is a significant difference between the medians of all three distributions, with G_3 having a larger median.
- There are clusters from all three networks for which a number of authors have a closeness centrality greater than 0.02. The authors in these clusters were explored but no pattern was found behind their publications. The provenance and the year of publication were checked.
- The authors in these clusters, are the authors which are in the main clusters of their relative networks. Thus, the people that influence the field the most are the most central authors in the main cluster of the co authorship network of a field.
- In Figure 15, the distributions are plotted in a violin plot. The network with the highest value of closeness centrality is the plot with the largest cluster.
- Both G_2 and G_3 have more people influencing the field compared to G_1 .

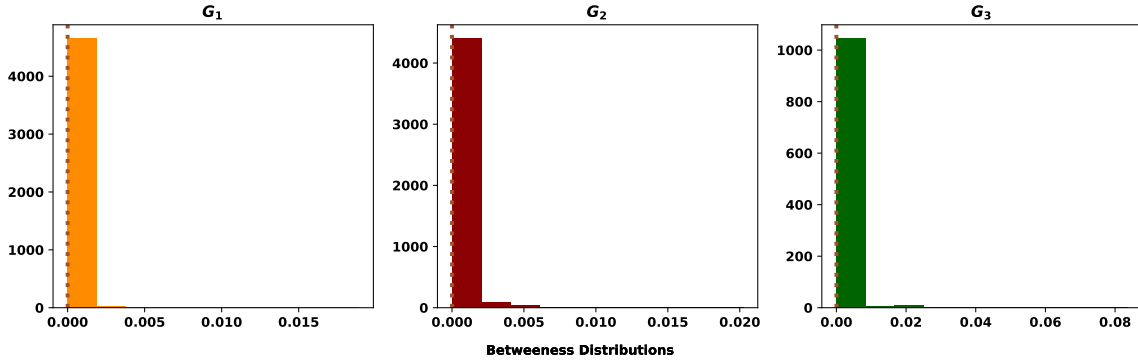


Figure 13: Betweenness centrality distributions $G_1 - G_3$.

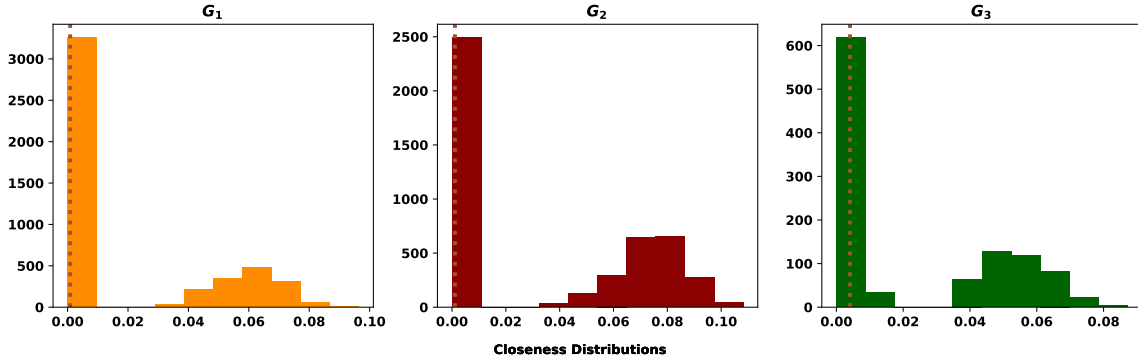


Figure 14: Closeness centrality distributions $G_1 - G_3$.

These results can be extended to the cumulative graphs. The distributions for both centralities have been plotted over time. In Figure 16, it is seen that betweenness has always been low in the network over time. Though there have been several outliers. This could be more evidence that in academic networks, betweenness is very low, the gain of being well connected to several authors in your fields.

Closeness centrality over time has changed more. There are several periods that there have been people with a big influence in the field. This seems not to be decreasing over the last period. The centrality

2.5 Conclusion

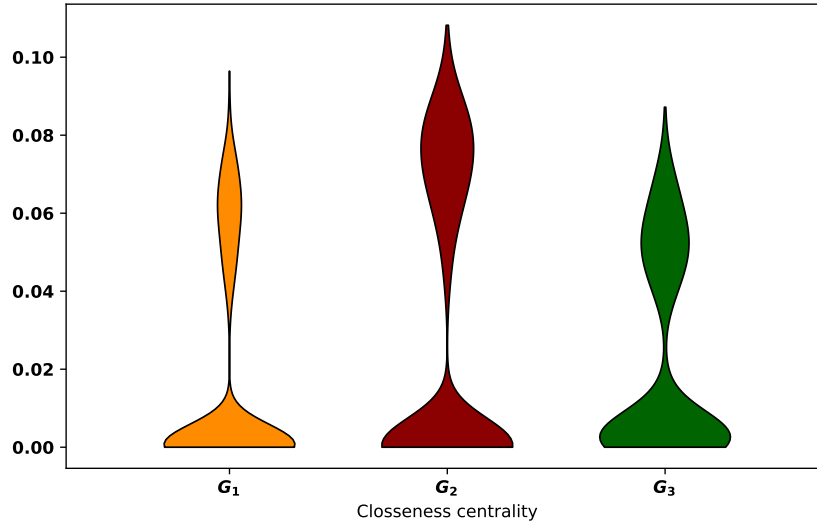


Figure 15: Violin plots of closeness centralities.

References

- [1] Mathieu Bastian, Sebastien Heymann, and Mathieu Jacomy. Gephi: An open source software for exploring and manipulating networks, 2009.
- [2] Vincent D Blondel, Jean-Loup Guillaume, Renaud Lambiotte, and Etienne Lefebvre. Fast unfolding of communities in large networks. *Journal of statistical mechanics: theory and experiment*, 2008(10):P10008, 2008.
- [3] Raymundo das Neves Machado, Benjamín Vargas-Quesada, and Jacqueline Leta. Intellectual structure in stem cell research: exploring brazilian scientific articles from 2001 to 2010. *Scientometrics*, 106(2):525–537, 2016.
- [4] Henry Etzkowitz. Individual investigators and their research groups. *Minerva*, 30(1):28–50, 1992.
- [5] Aric A. Hagberg, Daniel A. Schult, and Pieter J. Swart. Exploring network structure, dynamics, and function using NetworkX. In *Proceedings of the 7th Python in Science Conference (SciPy2008)*, pages 11–15, Pasadena, CA USA, August 2008.
- [6] Eric Jones, Travis Oliphant, Pearu Peterson, et al. SciPy: Open source scientific tools for Python, 2001–. [Online; accessed [today]].
- [7] Elias Koutsoupias and Christos Papadimitriou. Worst-case equilibria. In *Proceedings of the 16th Annual Conference on Theoretical Aspects of Computer Science, STACS’99*, pages 404–413, Berlin, Heidelberg, 1999. Springer-Verlag.
- [8] Svein Kyvik and Ingvild Reymert. Research collaboration in groups and networks: differences across academic fields. *Scientometrics*, 113(2):951–967, 2017.
- [9] Andrea Landherr, Bettina Friedl, and Julia Heidemann. A critical review of centrality measures in social networks. *Business & Information Systems Engineering*, 2(6):371–385, Dec 2010.
- [10] Peng Liu and Haoxiang Xia. Structure and evolution of co-authorship network in an interdisciplinary research field. *Scientometrics*, 103(1):101–134, Apr 2015.
- [11] Flavio M Menezes and Paulo Klinger Monteiro. *An introduction to auction theory*. OUP Oxford, 2005.
- [12] Nikoleta and Vince Knight. Nikoleta-v3/arcas: Arcas v 0.0.4, December 2017.
- [13] Richard Van Noorden. The science thats never been cited. <https://www.nature.com/articles/d41586-017-08404-0>, 2017.

- [14] Nurzhan Nurseitov, Michael Paulson, Randall Reynolds, and Clemente Izurieta. Comparison of json and xml data interchange formats: a case study. *Caine*, 2009:157–162, 2009.
- [15] Alan Pritchard et al. Statistical bibliography or bibliometrics. *Journal of documentation*, 25(4):348–349, 1969.
- [16] Dhruv Raina and BM Gupta. Four aspects of the institutionalization of physics research in india (1990–1950): Substantiating the claims of histortical sociology through bibliometrics. *Scientometrics*, 42(1):17–40, 1998.
- [17] Tim Roughgarden. *Selfish routing and the price of anarchy*, volume 174. MIT press Cambridge, 2005.
- [18] Martin Shubik. The dollar auction game: a paradox in noncooperative behavior and escalation. *Journal of Conflict Resolution*, 15(1):109–111, 1971.
- [19] Mason Youngblood and David Lahti. A bibliometric analysis of the interdisciplinary field of cultural evolution. *Palgrave Communications*, 4(1):120, 2018.