Behavior Mining in h-index Ranking Game

Rustam Tagiew[1] and Dmitry I. Ignatov[2]

[1] ONTONOVATION, Dresden, Germany,
Alumni of TU Freiberg and Uni Bielefeld, Germany
rustam@tagiew.de
http://www.tagiew.de

[2] National Research University Higher School of Economics, Moscow, Russia
dignatov@hse.ru
https://www.hse.ru/en/staff/dima

Abstract. Academic rewards and honors are proven to correlate with h-index, although it was not the decision criterion for them till recent years. Once h-index becomes the rule-setting scientometric ranking measure in the zero-sum game for academic positions and research resources as suggested by its advocates, the rational behavior of competing academics is expected to converge towards its game-theoretic solution. This paper derives the game-theoretic solution, its evidence in scientometric data and discusses its consequences on the development of science. DBLP database of 07/2017 was used for mining. Additionally, the openly available scientometric datasets are introduced as a good alternative to commercial datasets of comparable size for public research in behavioral sciences.

Keywords: h-index, scientometrics, behavior mining, behavioral game theory, experimental economics, data science, social networks, research funding, R&D budget, innovation management

1 Introduction

One of the main pillars of modern day politics is to reward innovations in order to ensure competitiveness. The global expenditures on research and development lay over 1.6 trillions (10^{12}) dollars per year, whereof USA, China, Japan, and Europe constitute 78% [1]. Depending on the country, the Research&Development (R&D) percentage of Gross Domestic Product (GDP) varies from $< 0.9\%$ in developing countries and up to 4% for Israel and South Korea. Chinese R&D budget will overtake today's leader USA by early 2020's. The number of scientists and engineers per inhabitant varies in strong correlation with the R&D budget up to 7‰ for Finland. Distributing percents of GDP over permilles of population for R&D creates an above-average income in this branch – a strong incentive for competition.

The small group of top scientists and thought leaders is easily spotted. Residual average academic title holders are in contrast more challenging to rank for fair reward. The academic title 'doctor', once introduced by catholic church in the middle-ages, made its way through the centuries over witch hunting theologists by supervisor-student-links into more modern and secular disciplines and became the necessary precondition. The challenge of ranking academics is suggested to be solved by statistics of citations – the modern-day scietometrics. The scientometric measure h-index is proven to correlate with chances of winning the Nobel Prize, holding position at top universities and being accepted for research fellowships [2]. More precisely, ranking based on h-index is a good estimation of the chance for being rewarded for scientific publishing without necessarily being the obligatory criterion for research funding committees. h is calculated as the maximal number of publications with at least h citations [3].

The actual usage of h-index as a decision criterion is harshly criticized by successful scientists from diverse disciplines [4,5]. The main arguments are the alienation of scientific work from its purpose and the negligence its practical component beyond composing scientific prose. Every scientist has his own representation of publications in his field and his own view on ranking of the relevant research, which does not fit scientometric figures. Since the return to the less transparent, less exact and more time consuming alternative of manual content comparison is not desirable, h-index is advocated to be a yardstick for resource allocation beyond being a correlated indicator nevertheless:

> "*I think that considering the h-index should result in better decisions pertaining to hiring and promotion of scientists, granting of awards, election to membership in honorary societies and allocation of research resources by agencies that have to decide between different competing proposals. As long as this index is well used I think it should contribute positively to the progress of science and help reward those who contribute to such progress more fairly.*" J. E. Hirsch [6], his[2] $h = 56$

The efforts to improve the research funding practice are rather put into improvement of h-type measures [7,8,9] or into alternative scientometric measures also known as 'altmetrics' [10,11] than into the return to the ancient methods. The introduction of h-index ranking as a decision criterion for budget allocation is a solution, which creates new problems and requires further fixes (Verschlimmbesserung in German). All modifications of h-index like h_m, g, i_{10}, e, $\hbar$, w and others are kept out of the scope of this paper, since those are not well established yet.

This paper reviews the current status quo of h-index from the perspective of human behavior research. Its game-theoretical analysis is provided in Section 2. Section 3 lists data-based evidences for the game-theoretical model from literature and own experiments on Digital Bibliography & Library Project (DBLP) dataset from July 2017 [12]. Some of these evidences are derived in this paper for the first time. Sections 4 and 5 provide a data-driven reconceptualization of the narrative for scientific process. Section 6 discusses the usage of scientometric datasets for research in behavioral sciences. Section 7 concludes the paper.

This paper is a piece of interdisciplinary research. Combining knowledge and methods from game theory, behavioral economics and data science in order to understand human behavior is a direction, where such market leaders as Facebook, Microsoft and

[1] h-index is calculated by scopus.com and used here to value the context of different opinions

Google push into since recent years [13,14,15]. Also in academia, workshops and conferences are organized for the intersection of experimental economics and machine learning [16,17,18,19]. For the analysis of human behavior from web data, the term 'Behavior Mining' is suggested, whereby the knowledge from behavioral sciences is incorporated into the process [20].

2 Equilibria of h-index ranking game

> "I suggest that this index may provide a useful yardstick with which to compare, in an unbiased way, different individuals competing for the same resource when an important evaluation criterion is scientific achievement." J. E. Hirsch [3]

> "For a few years, there might be a great increase in scientific output; but, by going after the obvious, pretty soon science would dry out. Science would become something like a parlor game." L. Szilard [21, p. 1498], inventor of nuclear chain reaction, his $h = 10$

If achieving the highest h-index rank is the base for the payoff function of N rational competing individuals, we can derive the solution of this game in game-theoretic sense. A solution of a game is a prediction about the behavior of the players given the assumption of their rationality. Rationality means that a player maximizes his payoff considering what he knows. A solution to a game is a set of possible equilibria. Every equilibrium is a combination of players' behaviors, where no player can improve his payoff by deviating in solo action. h-index ranking game is zero sum – every player ranks as much up as much others rank down.

A definition of a game is often a simplified formalization of a real world strategic interaction. It consists of a number of participating players, their legal actions and a payoff function for every player. Let us assume for simplification that every player $i \in N$ produces one innovative publication $p_{r,i} \in P_i$ per round $r \in \mathbb{N}^+$, where $P_i = N \times \mathbb{N}^+ \times A_i \times C$. All publications are assumed to be of the same quality. The effects of different production speeds and qualities will be discussed in later sections. None of the players has a publications before round 1 as the game starts. $A_i \subset \wp(N)$ is the set of possible coauthors' sets for a publication, which are subset of all players including the player himself, i.e. $\forall O \in A_i : i \in O$. The set of cited publications of participating players is $C \subset \wp(P)$. C contains the cited publications from past rounds of participating players only and hereby makes the definition of P recursive and non-circular. C includes neither the publications written by researchers from outside nor concurrent publications nor future publications. Citations of a publication from P by researchers from outside are considered to be negligible or randomly and equally distributed. Every player i is allowed to create publications with only him in the (co)authors' set and no citations of his competitors' works. Since "there is no penalty to add authors to a paper" [9], every player is also allowed to add any other players as coauthors and cite any publications from previous rounds. A publication is assumed to have one decisive contributor. The rank is dependent on h and if two players have the same h, their rank among each other will be randomly chosen.

Only in the second round, players can get an $h > 0$. In the second round, it is not rational for any player to cite publications of his competitors from the first round, which do not include him as a coauthor. Surely, every player will cite outside researchers, which do not compete with him for a certain resource. Every player will achieve at least an h-index of 1 in the second round, since he will cite his publication from the first round. It is irrational to not cite own publications. If a player adds x randomly chosen players as coauthors to his publication in the first round, he will have one publication with $x+1$ citations in the second round and still end up with $h = 1$. All x randomly chosen players will have two publications in the second round – first one $x+1$ times cited and the second with only one citation. They also end up with $h = 1$ as well. Adding random coauthors in solo action does neither improve nor worsen ones position in the h-index ranking.

If a clique $q \in \wp(N)$ of $x+1$ players agree to add them all as coauthors to their publications in the first round, they will achieve $h = x+1$ in the second round and will rank higher than the rest $N \setminus q$ with $h = 1$. None from the clique q will improve his rank by defecting from the agreement in solo action, because this will only reduce the h-index of the whole clique by 1. Even if only one coauthor is excluded from $x+1$ publications, then in the second round this will result in having x publication with $x+1$ citations and one publication with x. The excluded player will not cite the publication, he was excluded from. Therefore any formed clique q is an equilibrium and the solution of the game is a set of multiple equilibria.

If every agreement for the round 1 is an equilibrium, then players will prefer to belong to a slim majority clique $smq_1, |smq_1| = (|N| \div 2) + 1$. If a clique is less than majority, then the rest might form a single clique with a higher h. If a clique is much bigger than slim majority, then the members will be randomly ranked on a longer list of places on the top. The members of the slim majority clique smq_1 from the first round will outperform the rest by at least $2 - (|N| \mod 2) \in \{1, 2\}$. The mechanism of making agreements is considered to depend on features of social networks and too extensive to be modeled game-theoretically in the this work. It will be referred to as collaborativeness.

If the game lasts more than two rounds, for every round r being a member of a slim majority clique smq_r will add at least 1 to h more than being a member of the rest. Sets smq_1 and smq_2 do not need to be the same. If a player of extraordinary collaborativeness manages to be the only one player, who was a member of all slim majority cliques in all rounds, he will be the indisputable winner of h-index ranking.

This game-theoretical analysis reveals following major characteristics of rational behavior for a successful player, if the allocation of resources correlates with h-index ranking or is even based on it:

1. Cite publications (co)authored by you.
2. Never cite those researchers that might be involved in the competition over the same resource with you.
3. Make an agreement for a coauthoring clique. This clique should establish a slim majority involved in the competition for a certain resource.

4. If possible, abandon worked-out coauthoring agreements, if formation of a new coauthoring agreement with new coauthors of lower h-index can establish a slim majority.

3 Evidence in data

> "We see here that in the *real* real world – when the chips are down, the payoff is not five dollars but a successful career, and people have time to understand the situation – the predictions of game theory fare quite well." R. J. Aumann [22], Nobel Prize winner, his $h = 22$

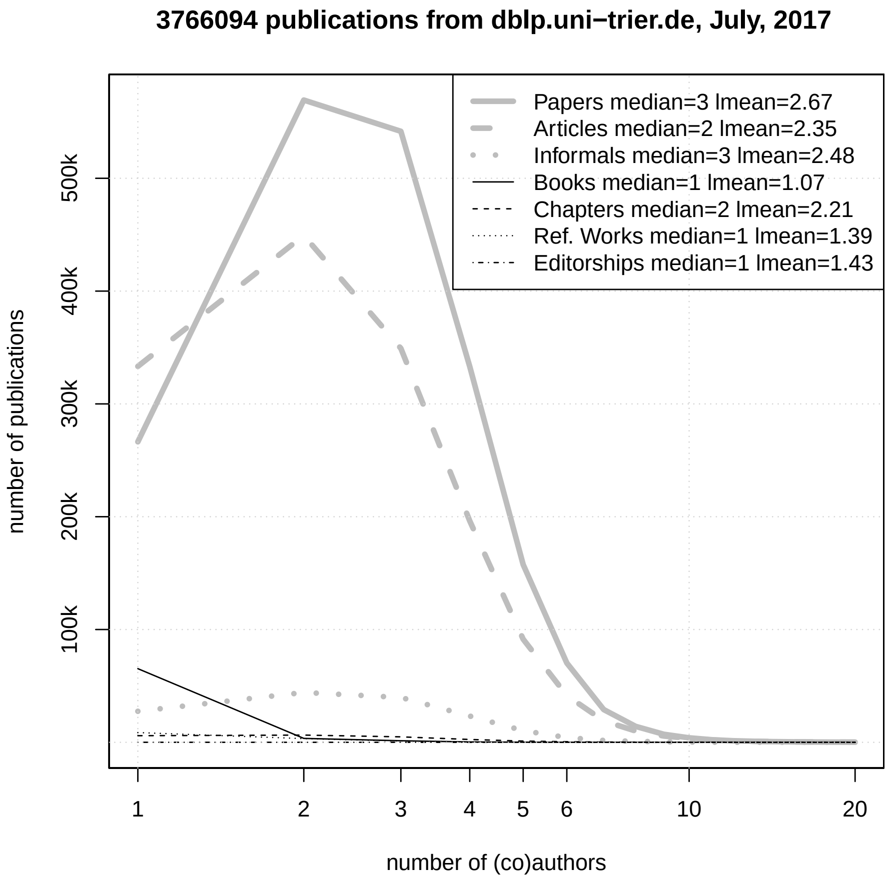
The 3.8M publications' dataset from DBLP computer science biography database shows that the distribution of the (co)authors' number per paper has the shape of a log-normal distribution (Fig.1). No limit seems to be set – one publication in DBLP has 267 coauthors. The medians of six out of seven types of records approximate the means of log-normal distributions. The observed proportions for $e^{\mu+\sigma}$ and $e^{\mu+2\sigma}$ upper bounds of (co)authors' number are close to the theoretically expected values of a log-normal distribution. 2.5 coauthors is today a typical value for informal publications, conference and journal papers in computer science. 3% of conference and journal papers and only 2.4% of informal publications have more than 6 authors. Informal publications target fast dissemination of ideas and have less of extra long coauthors' lists than conference and journal papers.

In the time before the introduction of h-index, the distribution of (co)authors per paper showed the same shape of a log-normal distribution in other disciplines too as a study on all papers indexed in the 1980-2000 annual volumes of the Science Citation Index (SCI) of the Institute for Scientific Information [23]. The median amount of coauthors increased from 2 to 3 between 1980 and 1998. It converges towards the solution of h-index ranking game – the size of cliques grows. The number of citations for a paper grows close to linear with the number of its (co)authors – the slope of this relationship became steeper from 1980 to 1998.

Fig. 2 shows the development of cumulative contributions per author per year. Since the share of each coauthor in a paper is not recorded in the database, it is derived by simply dividing 1 (a paper) by the number of coauthors. If an author (co)authored more than one paper in a year, these shares are added. For instance, being a coauthor of 2 2-coauthors papers will result a cumulative contribution of 1. 96% of DBLP records from 1980-2016 are used for this calculation. One can see in the graph that the approximately log-normal distribution of the recorded authors' cumulative share drifts towards 0. The game-theoretically predicted prolongation of (co)authors' list enables the incorporation of a growing number of scientists with far lower output into the scientific process.

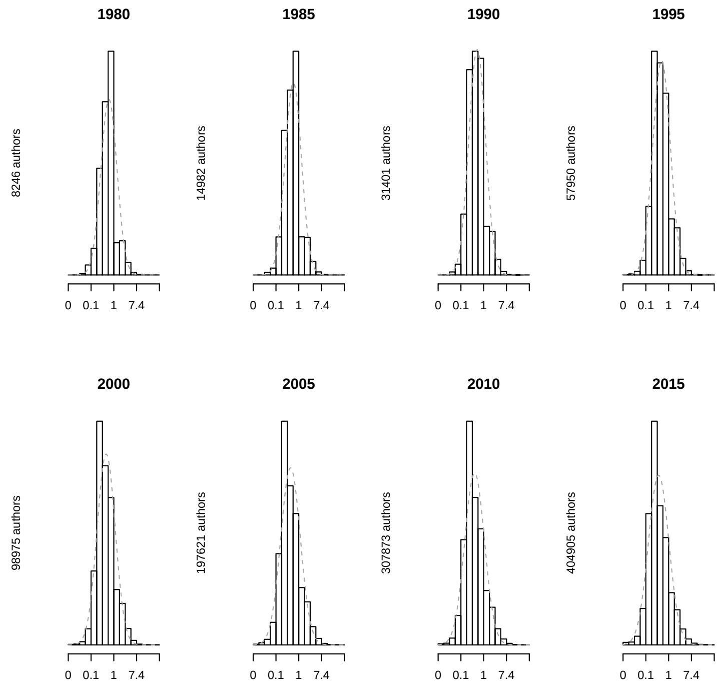
A policy close to the game-theoretic solution for h-index implemented by one large scientific institute, the Collider Detector at Fermilab (CDF)[3] since 1998. It enforces the addition of all its scientists and engineers as coauthors to all of its publications. Employees are added to the CDF authors' list after one year of full-time work and removed after a year since the date they left. This list contains typically over 300 authors. A

[3] `www-cdf.fnal.gov`



Type of publication	#(Co)authors $\leq e^{\mu+\sigma}$	$e^{\mu+\sigma}$	#(Co)authors $\leq e^{\mu+2\sigma}$	$e^{\mu+2\sigma}$
Expected proportion	84.13%		97.73%	
Conference **Papers**	85.53 %	4	96.92 %	6
Journal **Articles**	88.15 %	4	96.97 %	6
Informal **Publications**	88.05 %	4	94.81 %	5
Books and Theses	97.28 %	2	99.25 %	3
Book **Chapters**	89.52 %	4	94.6 %	5
Reference Works	95.53 %	3	98.31 %	4
Editorships	97.1 %	3	100 %	4

Fig. 1. (Co)authors per paper in computer science. x-axis has a logarithmic scale. The means of the log-normal distributions are denoted as ‘lmean’. Key data for $+\sigma$ and $+2\sigma$ upper bounds is organized in the table. Corresponding types from graph and table are boldly indicated.



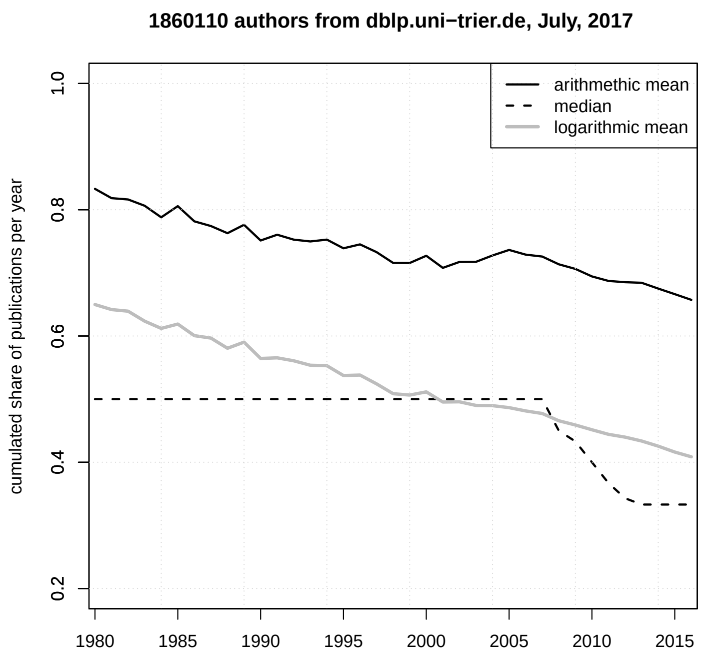


Fig. 2. Drift of average cumulative share of annually (co)authored papers during last decades. In the top graph, normal distributions are fitted to logarithmicly binned histogram of chosen years - dashed grey line. Bottom graph shows the development of average activity of contributing (co)authors.

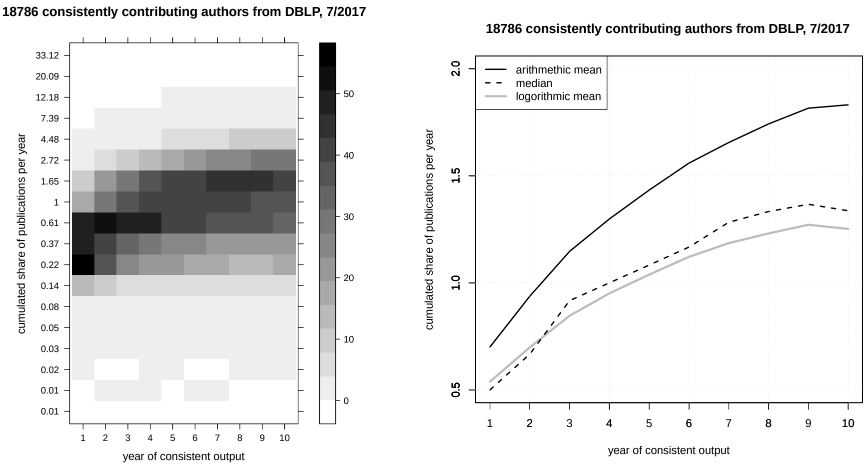


Fig. 3. Share of annually (co)authored papers of authors, who coauthored at least one paper a year since their first paper in the following 10 years. Top graph shows the distribution density in %. Bottom graph shows the growth of average productivity.

study conducted on a dataset of 189k publications [7] showed that the number of coauthors is strongly correlated with h as suggested by game-theoretic analysis. Every field has its typical average coauthor number. Mathematics has a big proportion of single-author paper, therefore mathematicians achieve lower h than others.

The correlations between scientometric measures and graph measures of coauthors' social network were calculated in a study on a dataset of 1809 authors from information management and systems schools of 5 US universities [24]. The highest correlation of 0.861 was observed between the number of publications and average tie strength, which is the number of joint publications. With growing number of publications, the cliques of coauthors stabilize – frequent reestablishment of a new clique might cost more than nothing. Alternatively, authors with a big number of publications might follow different goals than dominating h-index ranking game. Eigenvector centrality, which increases with the number of connected nodes and their connections, is not correlated with the number publications at all. In this context of clique stabilization being strongly correlated with number of papers, h-index is rather correlated with average tie strength at 0.660 than with eigenvector centrality at 0.042. h-index moderately correlates with average tie strength, because authors with a higher h have more stable cliques.

h-index does not suffer from dilution of innovation into multiple papers, since the growth of citations is strongly correlated with growth of papers per single research project. This is showed in a study on a dataset of 96 BIF grant applicants [25]. Quality of publications is thus not important for h-ranking game as assumed. Therefore, the speed is expected to raise. The number of scientific publications grows exponentially in many disciplines [26,27]. The share of publications available online grows as well. At the same time, the period of time for a publication to loose at chance to be cited anymore shortens – the quality of literature review is diminishing. Also the relationship between impact factor and citations weakens since late 90s [28]. This means that the quality of peer-review becomes less important and a potentially best-cited paper can be published in a journal with a less rigorous peer-review process[4].

For the plot in Fig. 3, all the authors from 1980-2016 are taken, who (co)authored at least one paper in the 10 following years after their first paper. The curve shows that the average productivity of a scientist roughly doubles within ten years. This can also be interpreted as growth of dilution of invitations into multiple papers. The curve shows pattern saturation too – there might be limits to productivity or the competing individuals change their objectives. Fig.4 shows *k*-means centroids of productivity trajectories, where the *k* is set to 4. The productivity of top 1-2% of authors show a linear growth over 50% per years. Meanwhile, roughly the half of authors shows no productivity growth at all. The publication shares of the non-growing half add up to one paper a year.

4 *h*-index rank measures collaborativeness

The game-theoretic solution reveals that collaborative academics are preferred by the *h*-index ranking, while academics writing single-author papers loose. If an academic is several times more productive, he will still achieve the same result as those, who achieve being added as coauthors same amount of times. An academic with a few high quality publications like the Fields Medal nominee Grigori Perelman would also loose in *h*-index ranking game. Hirsch identified as the major "short-coming" of original *h*-index definition "its inability to discriminate between authors that publish alone or in small cliques versus those authors whose papers have usually many coauthors" [6]. This paper assumes a neutral position seeing it rather as a feature than "short-coming".

If the fast *h*-index growth naturally results as a by-product from certain type of behavior, then this behavior is key to success in acquiring budget shares. According to the critics [4,5], the rational behavior in *h*-index ranking game deviates strongly from the natural behavior and should not be honored by budget shares. Could the natural behavior and the rational behavior in *h*-index ranking game be the same, since the result is the same? Academics might follow the game-theoretic solution unconsciously – they only need to have a bias for self-citations, for non-mentioning competitors and for participation in rapidly changing slim majority coauthorships. The Bonzi and Snyder's survey [29] conducted in early 90-s they studied scientists' perception of both self-citation and citation to others surveying 51 self-citing authors in several natural science disciplines. They found that scientists did not testify for any substantial difference in self-citation motivation. Thus main motivation was scientific and one of their respondents argued: "If you are a major contributor, it's difficult to avoid citing yourself." The unconscious bias for self-citation improves *h*, which correlates with success. Mastering to always be a part of dynamic slim majorities seems intuitively to be correlated with success also beyond academic world. Conscious rational behavior in *h*-index ranking game would require the same skill.

[4] For instance, the paper "The Conceptual Penis as a Social Construct" arguing that penises cause climate change could be published in a peer-reviewed journal

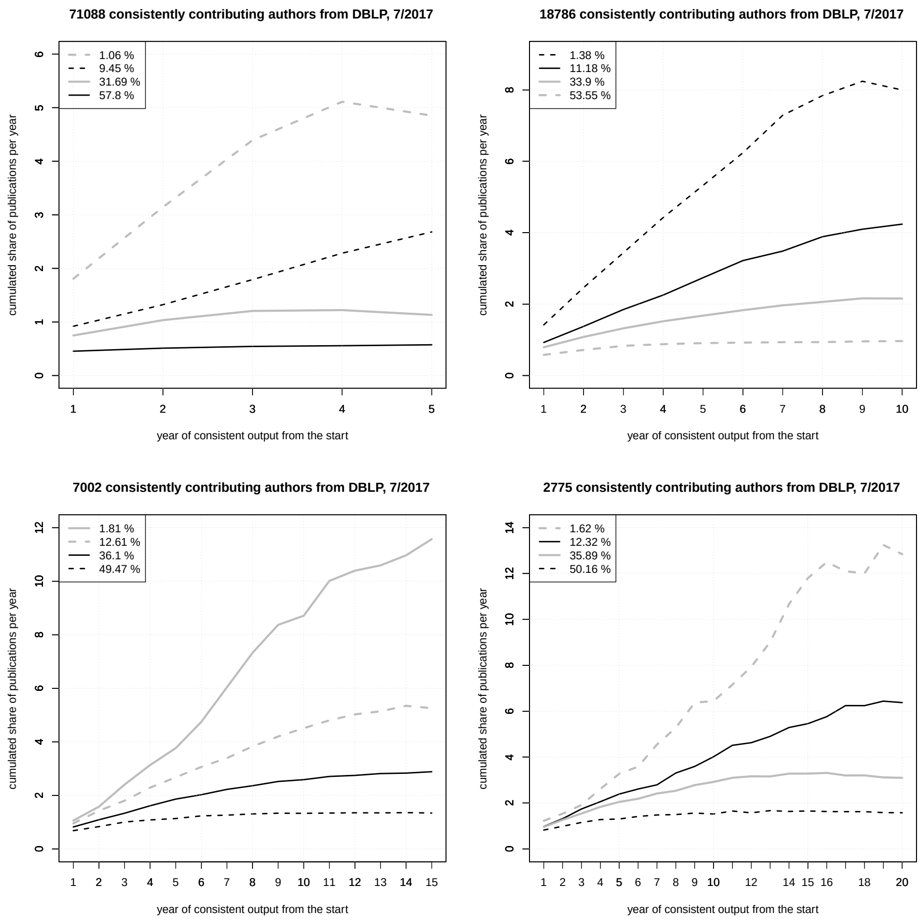


Fig. 4. Share of annually (co)authored papers of authors, who coauthored at least one paper a year since their first paper in the subsequent 5,10,15, and 20 years. The four graphs show the *k*-means clustering results on maximumly available consistently contributing authors.

5 Division of labor in science

A scientific community is obviously a collaborative network, which needs socially active members to exist. In the case of naturally grown h-index, it represents the degree of collaborativeness of an author inside of a scientific community. Scientific publications network is a mirror of the real social network. h-index ranking rewards its socially active members the most. Socially active members are also the rule-setter in a community and therefore will advocate the status quo of h-index ranking, since they profit from it the most. Is the core of scientific achievement to create a community around a certain topic, which urges to create innovations?

Even if the academics don't develop own ideas for next publications, they will at least actively adopt and develop ideas from non-scientific sources. They will become an idea hungry community, which is eager to publish, to coauthor and to cite everything as a consequence of rationality in h-index ranking game. The ground-truth is that the real originator of an innovation is not always among the authors of the publication exposing it and also not among those, who will be rewarded for this academic achievement. Like in patent affairs [30, e.g.], where most patents are owned by non-individuals, there is a division of labor between those, who originate ideas, and those, who promote them into real life. Ideas might appear in different heads simultaneously – it is an honorable scientific achievement to effectively place those ideas into the scientific community using social skills. On the other hand, tools should provided to better honor scientific beyond composing publications. For instance, data citation [31] is useful feature for this goal.

6 Scientometric datasets for behavior sciences

The bottleneck of public research in behavior mining is the limited access to large datasets, which are mostly held by for-profit companies. Preparation and release of large, authentic and recent datasets tend to contradict the commercial interest. Even if several studies on commercial datasets are put into public domain, their datasets might not be available for reproducibility of results. While the sizes of datasets from non-profit social networks are about 100k participants [32], non-profit scientometric databases like Citeseerx [33] and DBLP offer datasets with millions of participants. The game-theoretic solution of h-index ranking game, which is presented in this, can be used as base for the tailored hypothesis space in data mining.

7 Conclusion

This paper introduced game-theoretic perspective into scientometrics. The leading measure – the h-index ranking established a reward system, which prefers socially active academics and therefore furthers the labor division in science. The evidences for the convergence towards the game-theoretic have been found in the data of DBLP database and in results of related work.

Acknowledgments The research was supported by the Russian Science Foundation under grant 17-11-01294 and performed at National Research University Higher School of Economics, Russia. We thank the people behind DBLP for providing free full access to their database.

References

1. Battelle Memorial Institute: 2014 global R&D funding forecast (2013)
2. Bornmann, L., Daniel, H.D.: What do we know about the h index? Journal of the American Society for Information Science and Technology **58**(9) (2007) 1381–1385
3. Hirsch, E.J.: An index to quantify an individual's scientific research output. Proc. Nat. Acad. Sci. **46** (2005)
4. Lawrence, P.A.: The mismeasurement of science. Current Biology **17**(15) (2007) 583–585
5. Şengör, A.M.C.: How scientometry is killing science. GSA Today (2014) 44–45
6. Hirsch, J., Buela-Casal, G.: The meaning of the h-index. International Journal of Clinical and Health Psychology **14**(2) (2014) 161–164
7. Batista, P.D., Campiteli, M.G., Kinouchi, O.: Is it possible to compare researchers with different scientific interests? Scientometrics **68**(1) (2006) 179–189
8. Egghe, L.: The hirsch index and related impact measures. Annual review of information science and technology **44**(1) (2010) 65–114
9. Hirsch, J.E.: An index to quantify an individualâĂŹs scientific research output that takes into account the effect of multiple coauthorship. Scientometrics **85**(3) (2010) 741–754
10. Yu, M.C., Wu, Y.C.J., Alhalabi, W., Kao, H.Y., Wu, W.H.: Researchgate: An effective altmetric indicator for active researchers? Computers in human behavior **55** (2016) 1001–1006
11. Tagiew, R.: Research project: Text engineering tool for ontological scientometry. CoRR **abs/1601.01887** (2016)
12. Ley, M.: DBLP - some lessons learned. PVLDB **2**(2) (2009) 1493–1500
13. Bailey, M.: What does an economist at facebook do? quora.com/What-does-an-economist-at-Facebook-do (2014)
14. Microsoft Inc.: Microsoft research new york city. research.microsoft.com/en-us/labs/newyork/ (2014)
15. Varian, H.R.: Big data: New tricks for econometrics. people.ischool.berkeley.edu/ hal/Papers/2013/ml.pdf (2013)
16. Tagiew, R., Ignatov, D.I., Neznanov, A.A., Poelmans, J., eds.: First International Workshop on Experimental Economics and Machine Learning, KU-Leuven (2012) ceur-ws.org/Vol-870/, at ICFCA.
17. Sunstein, C., Zittrain, J.: Social media and behavioral economics. today.law.harvard.edu/social-media-and-behavioral-economics-conference/ (2013)
18. Tagiew, R., Ignatov, D.I., Amroush, F., eds.: Second International Workshop on Experimental Economics and Machine Learning, IEEE Computer Society (2013) dx.doi.org/10.1109/ICDMW.2013.178, at ICDM.
19. Tagiew, R., Ignatov, D.I., Hilbert, A., Delhibabu, R., eds.: Proceedings of the Third Workshop on Experimental Economics and Machine Learning co-located with the 13th International Conference on Concept Lattices and Their Applications (CLA 2016), Moscow, Russia, July 18, 2016. Volume 1627 of CEUR Workshop Proceedings., CEUR-WS.org (2016)
20. Chen, Z.: From data mining to behavior mining. International Journal of Information Technology and Decision Making **5**(4) (2006) 703–712
21. United States. Congress. Senate: Hearings. Number Bd. 6. U.S. Government Printing Office (1961)

22. Roth, A., Sotomayor, M.: Two-Sided Matching: A Study in Game-Theoretic Modeling and Analysis. Econometric Society Monographs. Cambridge University Press (1992)
23. Persson, O., Glänzel, W., Danell, R.: Inflationary bibliometric values: The role of scientific collaboration and the need for relative indicators in evaluative studies. Scientometrics **60**(3) (2004) 421–432
24. Abbasi, A., Altmann, J., Hossain, L.: Identifying the effects of co-authorship networks on the performance of scholars: A correlation and regression analysis of performance measures and social network analysis measures. Journal of Informetrics **5**(4) (2011) 594–607
25. Bornmann, L., Daniel, H.D.: Multiple publication on a single research study: does it pay? the influence of number of research articles on total citation counts in biomedicine. Journal of the Association for Information Science and Technology **58**(8) (2007) 1100–1107
26. Parolo, P.D.B., Kumar, R., Ghosh, R., Huberman, B.A., Kaski, K.: Attention decay in science. (2015)
27. Evans, J.A.: Electronic publication and the narrowing of science and scholarship. Science (New York, N.Y.) **321**(5887) (2008) 395–399
28. Lozano, G.A., LariviÃÍre, V., Gingras, Y.: The weakening relationship between the impact factor and papers' citations in the digital age. Journal of the Association for Information Science & Technology **63**(11) (2012) 2140–2145
29. Bonzi, S., Snyder, H.W.: Motivations for citation: A comparison of self citation and citation to others. Scientometrics **21**(2) (Jun 1991) 245–254
30. Sterzi, V.: Patent quality and ownership: An analysis of uk faculty patenting. Research Policy **42**(2) (2013) 564–576
31. Group, D.C.S.: Joint Declaration of Data Citation Principles. force11.org/datacitation (2014)
32. Tagiew, R., Ignatov, D.I., Delhibabu, R.: Economics of internet-based hospitality exchange. In: Proceedings of the 2015 IEEE / WIC / ACM International Conference on Web Intelligence and Intelligent Agent Technology (WI-IAT) - Volume 01. WI-IAT '15, Washington, DC, USA, IEEE Computer Society (2015) 493–498
33. Wu, J., Williams, K., hsuan Chen, H., Khabsa, M., Caragea, C., Ororbia, A., Jordan, D., Giles, C.L.: CiteSeerX : AI in a Digital Library Search Engine. (2014) 2930–2937