

AGE-DECAYING H-INDEX FOR SOCIAL NETWORKS OF CITATIONS

D. Katsaros^{1,2}, A. Sidiropoulos², Y. Manolopoulos²

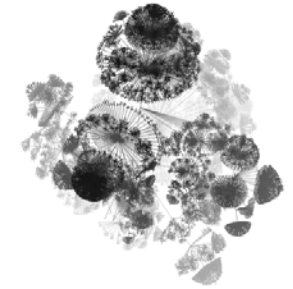


¹Dept. of Computer & Communication Engineering
University of Thessaly, Volos, Greece



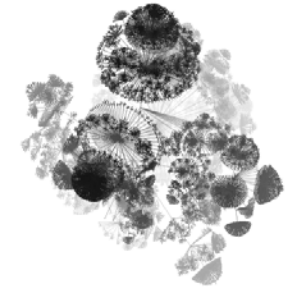
² Dept. of Informatics
Aristotle University, Thessaloniki, Greece

Presentation by:
Apostolos Papadopoulos²



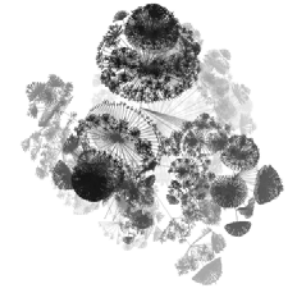
Methods for Ranking Scientists

- Evaluation of scientists by “experts”
 - e.g., surveys
- Citation Analysis
 - Task: Compute a score for the “objects”
- Hybrid method of previous two



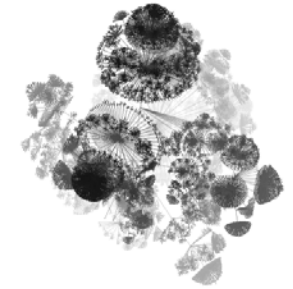
Drawbacks of various scientists ranking methods

- Not measure the importance of papers
- Affected by “big hits”
- Not measure productivity
- Need to set administrative parameters



H-index

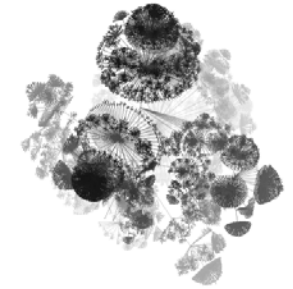
- Published by J.E. Hirsch in Oct. 2005
- Was known at least two years earlier
- Definition:
A researcher has *h-index* h if
 - h of his N_p articles have received at least h citations each
 - the rest $N_p - h$ articles have received no more than h citations each



H-index

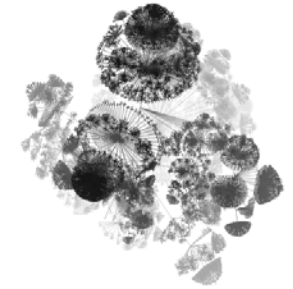
- Calculates the broadness of a researcher
 - ✓ Productivity
 - ✓ Impact

 - ✓ Not affected by “big hits”
 - ✓ Not affected by “noise”



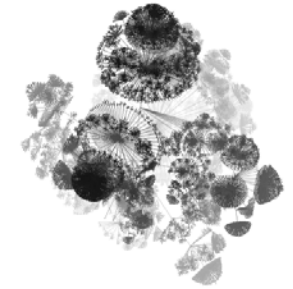
Index a

- $N_{c,tot} \geq h^2$
- Definition:
A researcher has index a if
$$N_{c,tot} = ah^2$$
- Second metric-index



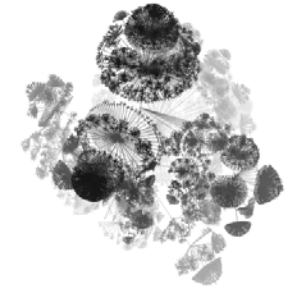
H-index drawbacks

- It is a growing function over time
- Does not show scientist's inactivity or retirement
- Scientists with short scientific life are out of competition



Contemporary H-index

- Definition (in *Scientometrics*, 72(2), 2007):
A researcher has *contemporary h-index* h_c if
 - h_c of his N_p articles have $S^c(i) \geq h_c$
 - the rest $N_p - h_c$ articles have $S^c(i) \leq h_c$
 - $S^c(i) = \gamma * (Y(now) - Y(i) + 1)^{-\delta} |C(i)|$
 - In our experiments: $\gamma=4$ and $\delta=1$
- An old article gradually loses its “value”
- Show how “active” a researcher is

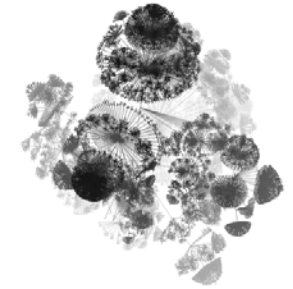


Trend H-index

- Definition (in *Scientometrics*, 72(2), 2007):
A researcher has *trend h-index* h_t if
 - h^t of his N_p articles have $S^t(i) \geq h$
 - the rest $N_p - h^t$ articles have $S^t(i) \leq h$

$$S^t(i) = \gamma^* \sum_{\forall x \in C(i)} (Y(now) - Y(i) + 1)^{-\delta}$$

- In our experiments: $\gamma=4$ and $\delta=1$
- An old citation gradually loses its “value”
- Shows how “trend” the work of a researcher is



Age Decaying H-index

- Definition:

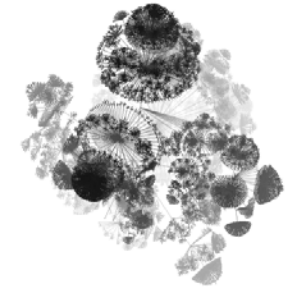
A researcher has *Age Decaying h-index* h_{ad} if

- h^{ad} of his N_p articles have $S^{ad}(i) \geq h^{ad}$
- the rest $N_p - h^{ad}$ articles have $S^{ad}(i) \leq h^{ad}$

$$S^{ad}(i) = \gamma^2 * (Y(now) - Y(i) + 1)^{-\delta_1} * \sum_{\forall x \in C(i)} (Y(now) - Y(i) + 1)^{-\delta_2}$$

- In our experiments: $\gamma=4$ and $\delta_1=\delta_2=1$

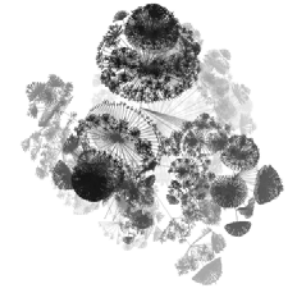
- An old article gradually loses its “value”
- An old citation gradually loses its “value”



The second metric “a” of h-index

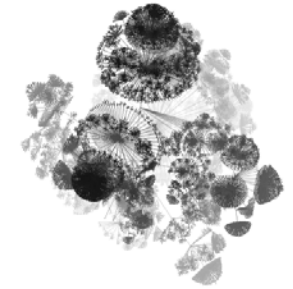
- A factor **a** is defined as:
 - $N_{c,tot} = ah^2$
- For h-index generalizations, instead of $N_{c,tot}$ we use the sum of each paper score $S_{ad}(i)$:

$$\sum_{\forall i \in P} S_{ad}(i) = a_{ad} h_{ad}^2$$



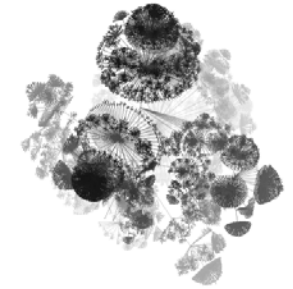
H-index generalizations

- Contemporary
- Trend
- Age Decaying
 - Scientists, journals, conferences or any other kind of semantic grouping of articles



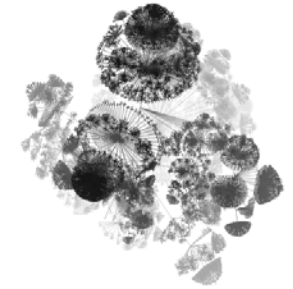
Experiments

- DBLP collection (<http://dblp.uni-trier.de/>)
 - Data timestamp: March 2006
- DBLP includes data for authors, journals and conferences
- Focuses in the DB area
- “Names Problem” is solved



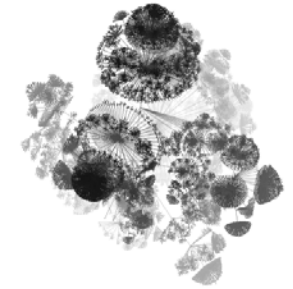
Experiments – *h-index*

Name	<i>h</i>	<i>a</i>	<i>N_{c,tot}</i>	<i>N_p</i>
1. Michael Stonebraker	24	3.78	2180	193
2. Jeffrey D. Ullman	23	3.37	1783	227
3. David J. DeWitt	22	3.91	1896	150
4. Philip A. Bernstein	20	3.39	1359	124
5. Won Kim	19	2.96	1071	143
6. Catriel Beerl	18	3.16	1024	93
7. Rakesh Agrawal	18	3.06	994	154
8. Umeshwar Dayal	18	2.81	913	130
9. Hector Garcia-Molina	17	3.60	1041	314
10. Yehoshua Sagiv	17	3.52	1020	121

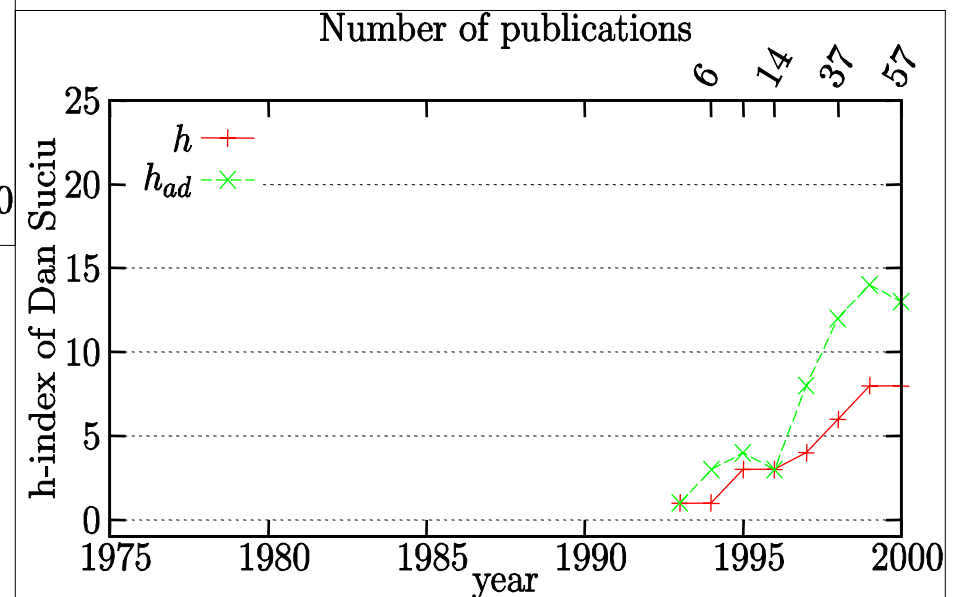
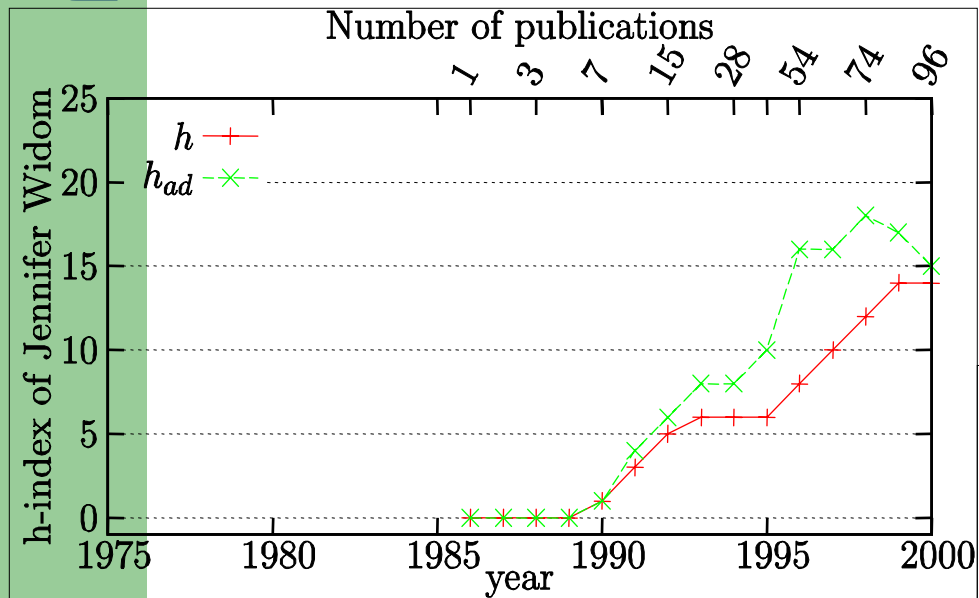


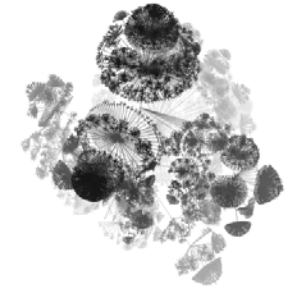
Experiments – Age Decaying h -index

Name	h_{ad}	a_{ad}	<i>Pos</i> <i>by h</i>	<i>Pos</i> <i>by h_c</i>	<i>Pos</i> <i>by h_t</i>
1.Rakesh Agrawal	16	3.28	7	4	4
2.Jennifer Widom	16	3.19	23	6	5
3.Serge Abiteboul	16	3.08	13	5	6
4.Dan Suciu	16	2.79	100	22	25
5.Alon Y. Levy	15	2.85	69	21	13
6.Jeffrey D. Ullman	14	4.18	2	2	2
7.David J. DeWitt	14	3.41	3	1	1
8.Jeffrey F. Naughton	14	2.95	18	10	9
9.Hector Garcia-Molina	13	4.07	9	8	7
10.Christos Faloutsos	13	2.62	16	11	8

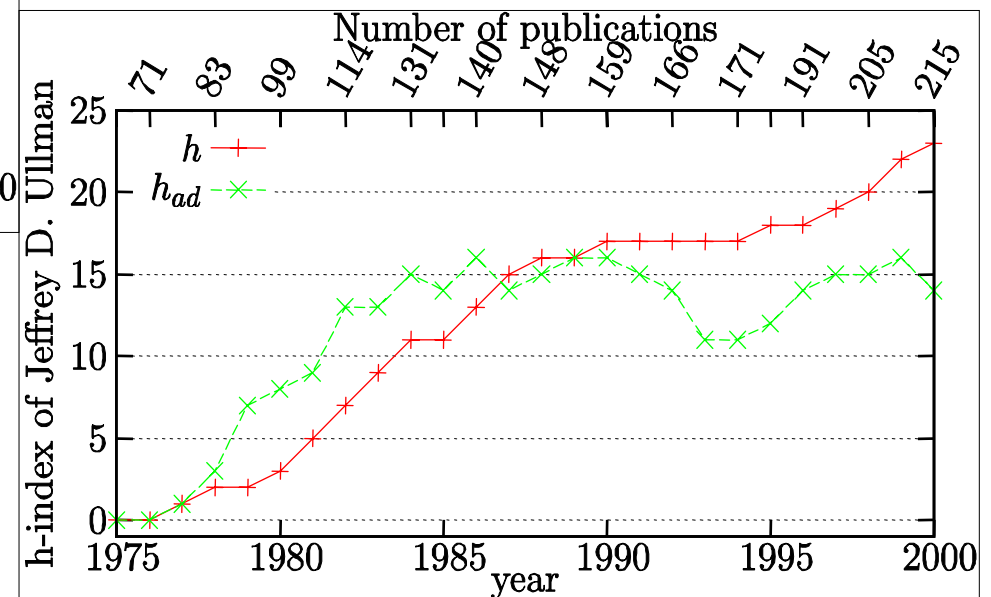
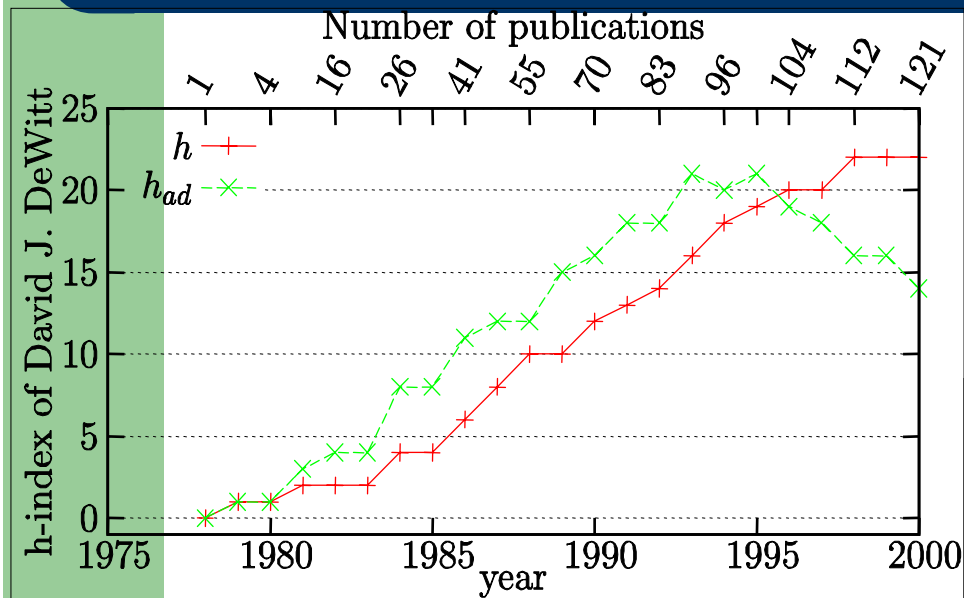


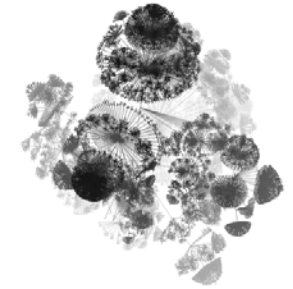
Experiments – h-index for scientists



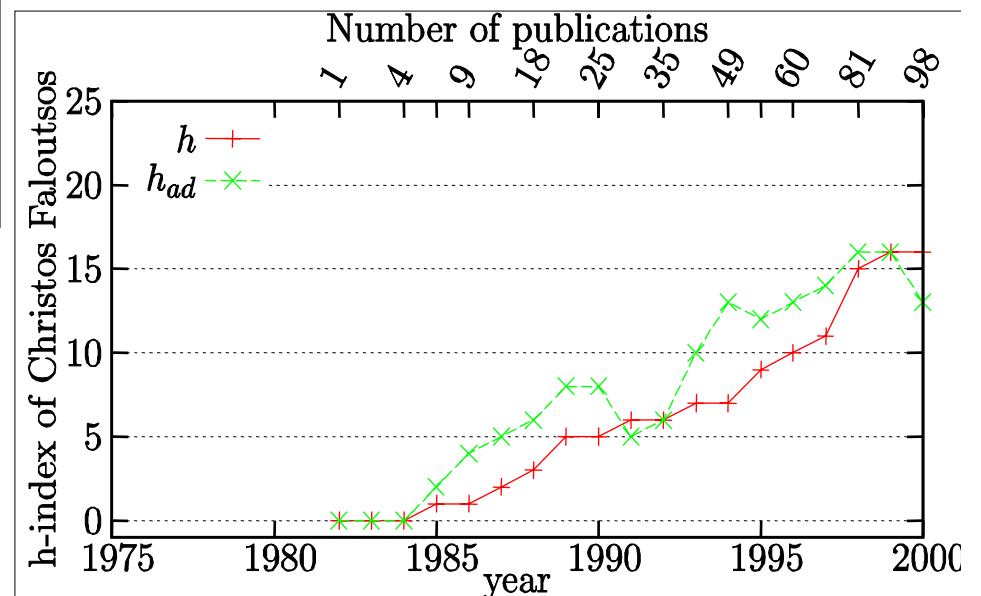
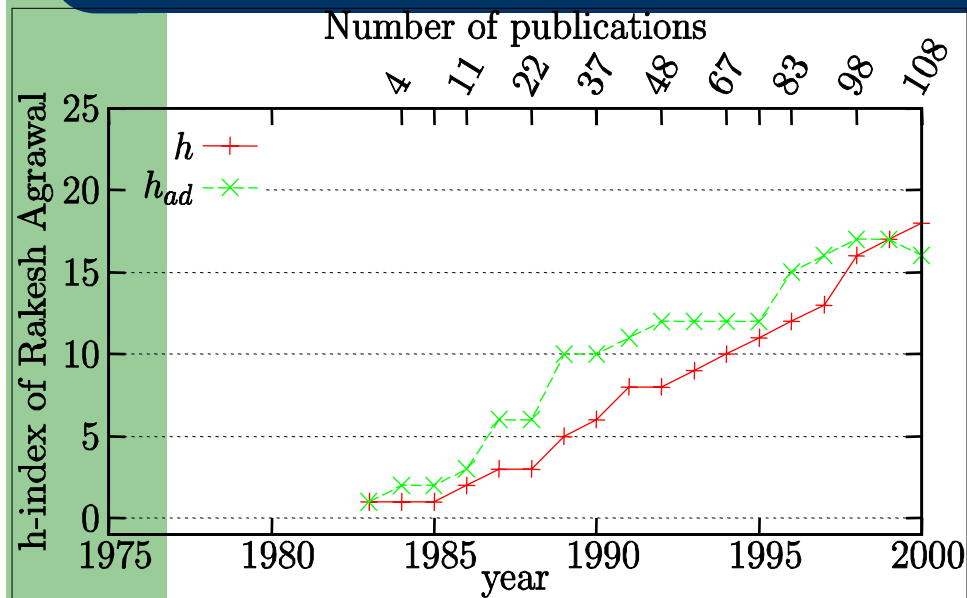


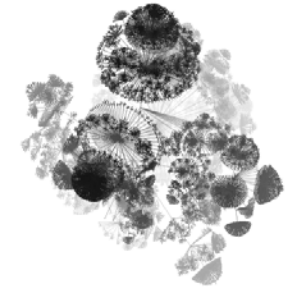
Experiments – h-index for scientists





Experiments – h-index for scientists

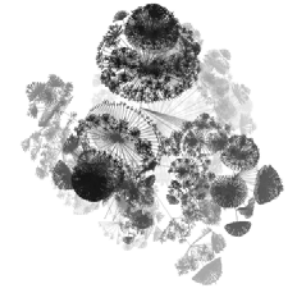




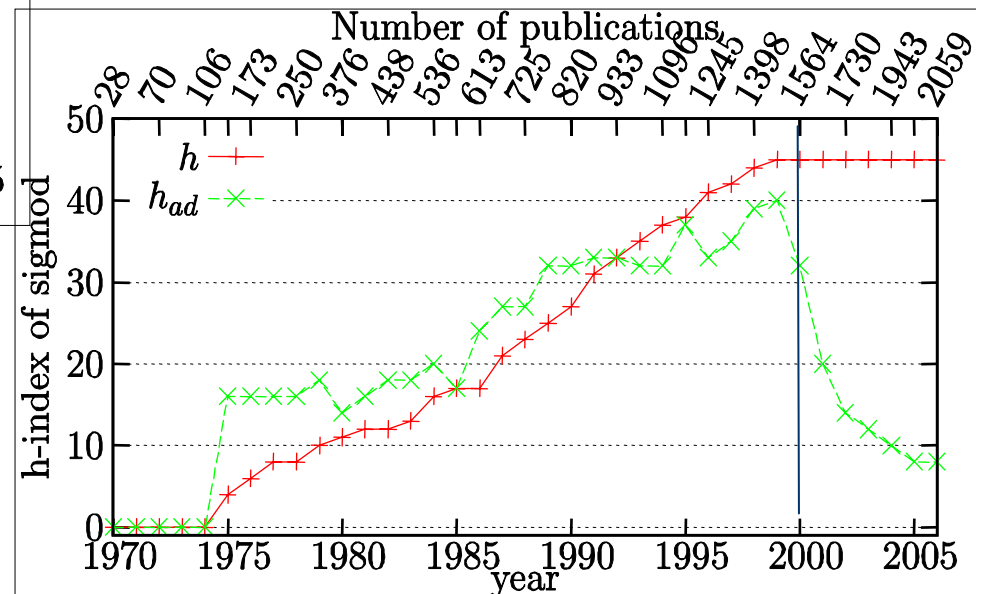
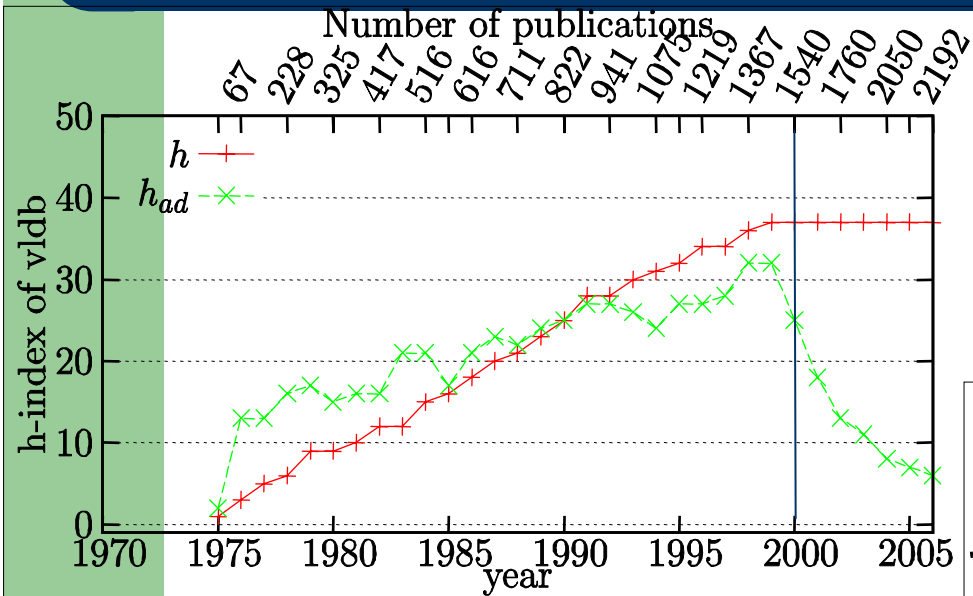
Experiments – H-index for conferences

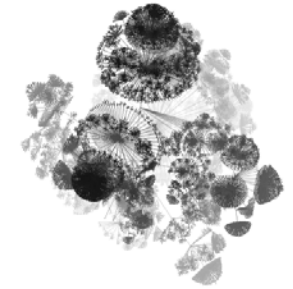
Name	h	a	$N_{c,tot}$	N_p
1.sigmod	45	6.05	12261	2059
2.vldb	37	7.10	9729	2192
3.pods	26	5.74	3883	776
4.icde	22	6.83	3307	1970
5.er	16	5.80	1486	1338
6.edbt	13	3.89	658	434
7.eds	12	3.65	527	101
8.adbt	12	2.86	412	42
9.icdt	11	4.79	580	313
10.oodbs	11	3.96	480	122

Name	h_{ad}	a_c	$N_{c,tot}$	N_p	h
1.sigmod	32	5.85	12261	2059	45
2.vldb	25	6.94	9729	2192	37
3.pods	20	5.32	3883	776	26
4.icde	17	8.01	3307	1970	22
5.icdt	12	5.27	580	313	11
6.kdd	11	4.08	243	1074	6
7.edbt	11	3.92	658	434	13
8.webdb	9	2.69	31	163	3
9.cikm	8	4.18	211	1030	7
10.ssdbm	8	3.71	321	609	7

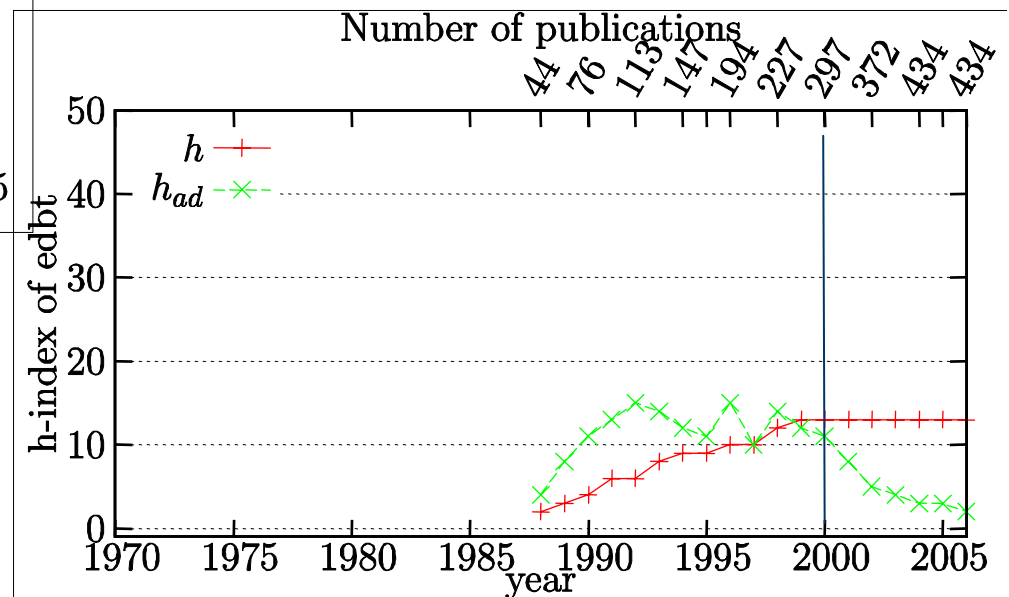
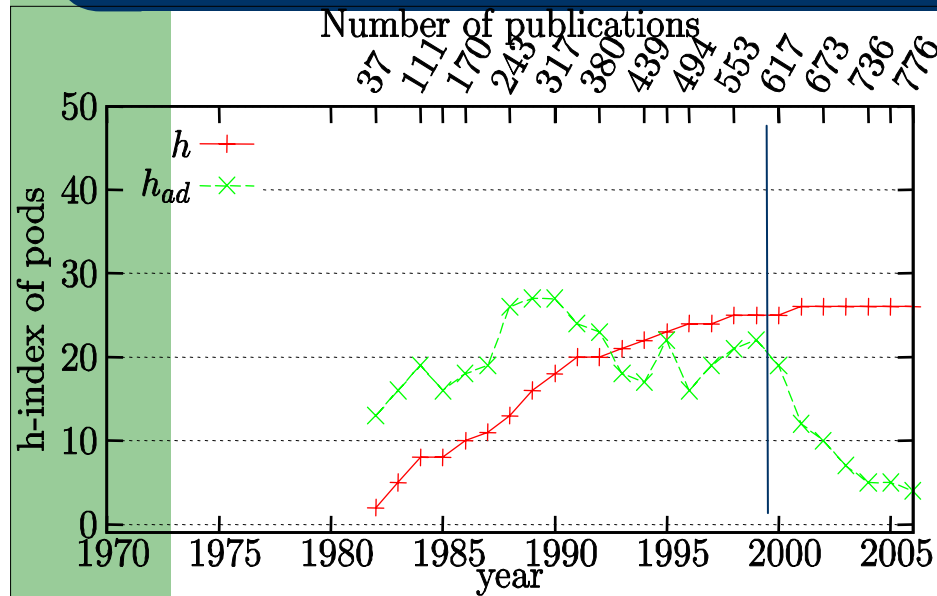


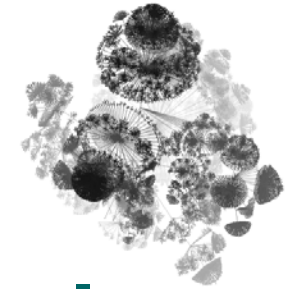
Experiments – H-index for conferences





Experiments – H-index for conferences

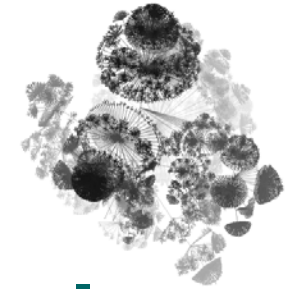




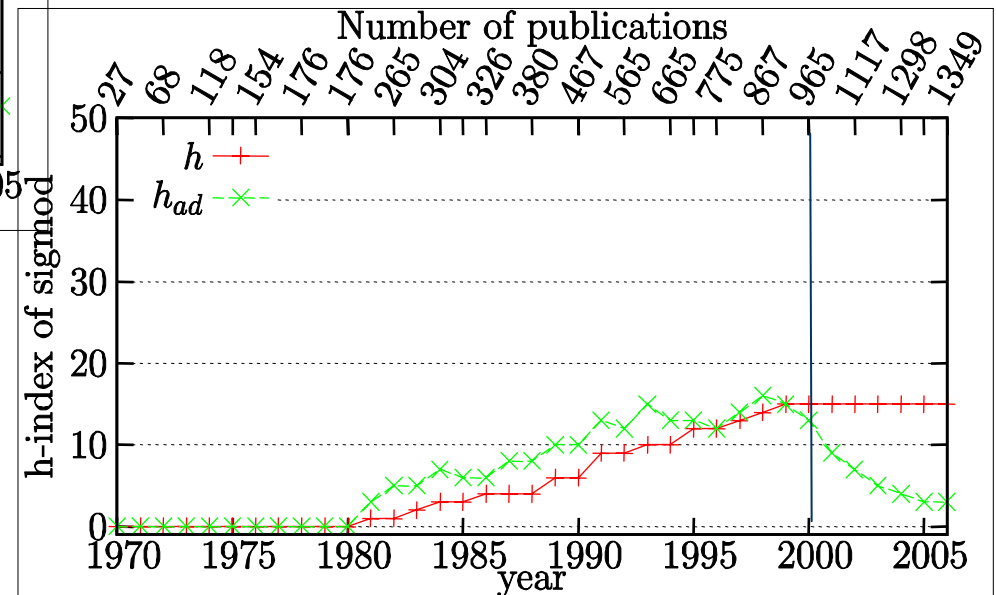
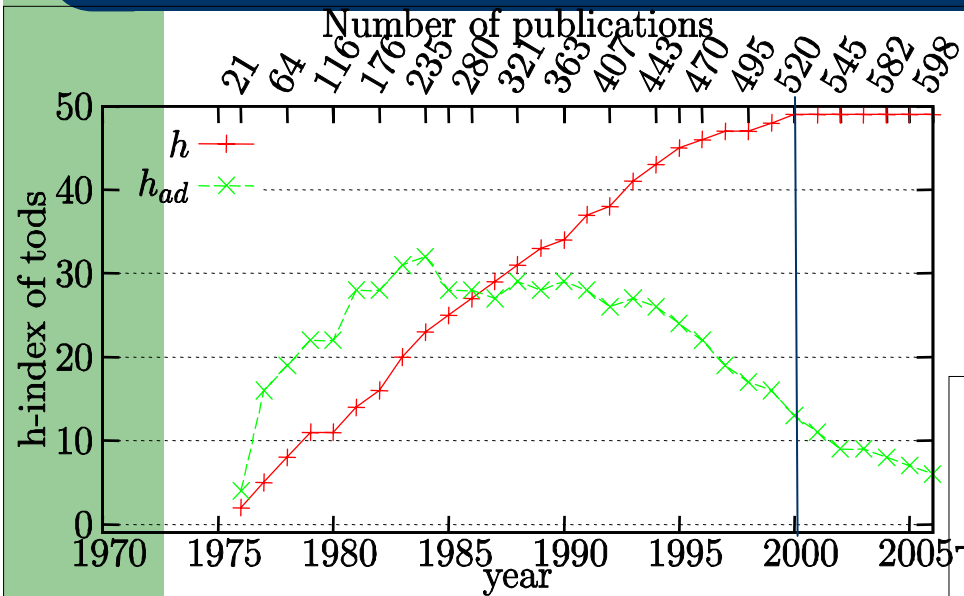
Experiments – H-index for journals

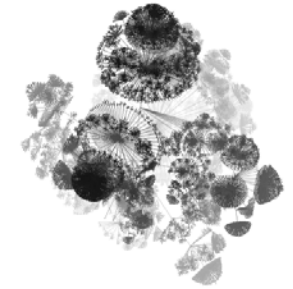
Name	h	a	$N_{c,tot}$	N_p
1.tods	49	3.88	9329	598
2.tkde	18	4.69	1520	1388
3.is	16	4.71	1208	934
4.sigmod	15	5.07	1142	1349
5.tois	13	4.37	740	378
6.debu	11	7.13	863	877
7.vldb	9	5.03	408	281
8.ipl	8	6.06	388	4939
9.dke	6	8.77	316	773
10.dpd	6	5.25	189	238

Name	h_{ad}	a_{ad}	$N_{c,tot}$	N_p	h
1.tods	13	7.71	9329	598	49
2.sigmod	13	4.94	1142	1349	15
3.tkde	12	5.77	1520	1388	18
4.debu	12	3.49	863	877	11
5.vldb	12	2.82	408	281	9
6.dpd	7	3.82	189	238	6
7.is	6	7.51	1208	934	16
8.jiis	6	5.67	156	318	6
9.tois	5	7.14	740	378	13
10.dke	5	6.52	316	773	6



Experiments – H-index for journals

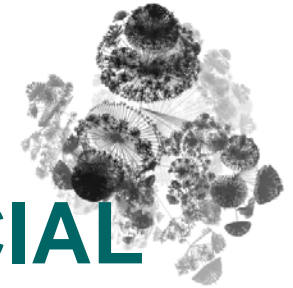




Conclusion

- Evaluation of scientists based on citation analysis
- Evaluation of publication forums based on citation analysis
- H-index shortcomings:
 - Active – inactive scientists
 - Significant works in the past – not any more significant
- H-index generalizations on the time dimension

AGE DECAYING H-INDEX FOR SOCIAL NETWORKS OF CITATIONS



D. Katsaros, A. Sidiropoulos, Y. Manolopoulos

Thank you for your attention!

Presenter: Apostolos Papadopoulos