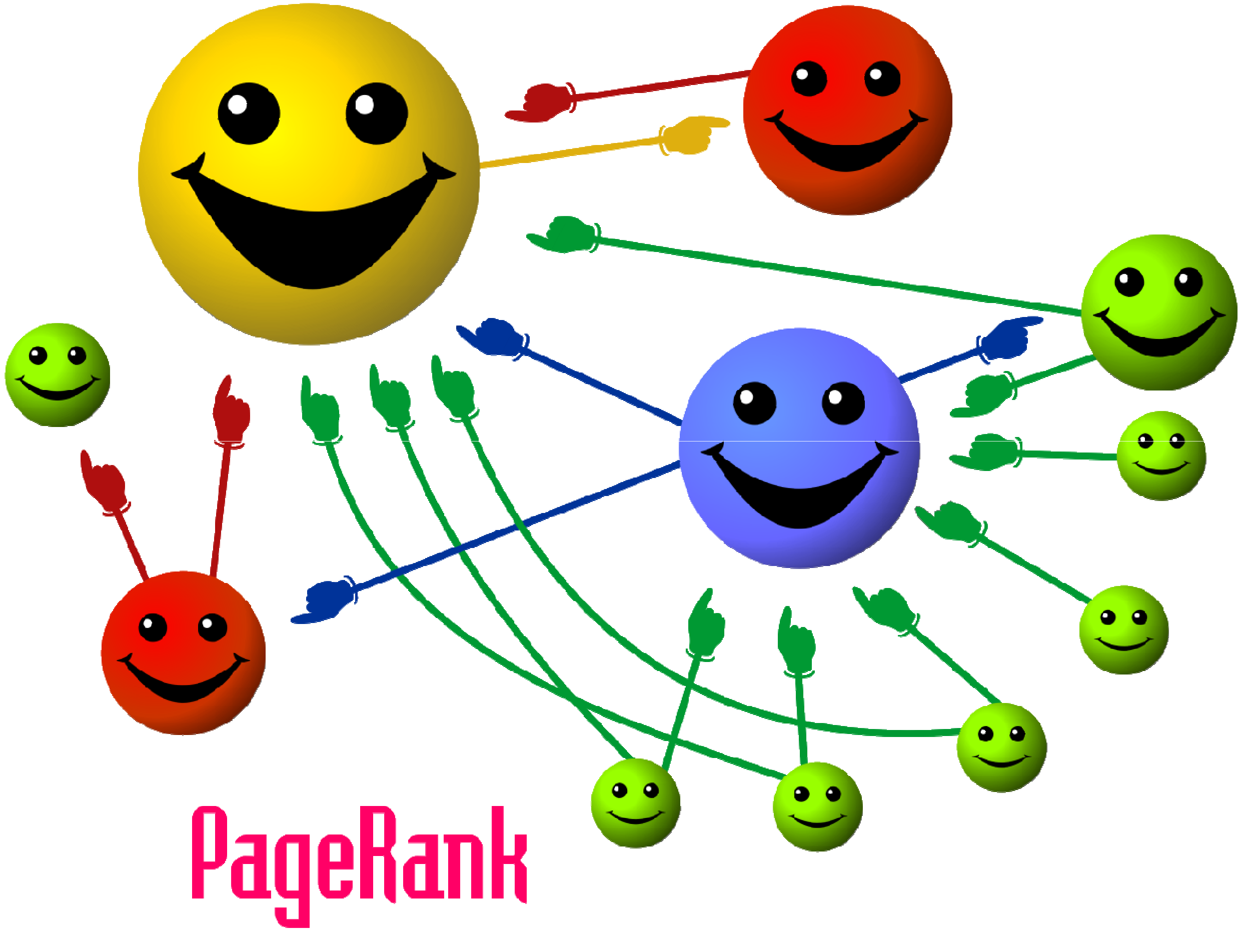
Google PageRank to the rescue of bibliometrics?



JL Foulley (IMAG) in collaboration with Gilles Celeux (INRIA-Select, Paris-Orsay),
Simon Grah (M2, Data Science, X Telecom, Paris) & Julie Josse (CMAP, Paris-Saclay),

Content

- Context
- Journal Impact Factor
- Google PageRank
- Journal Influence Indicators derived from P-Rank
- Example of Statistical Journals
- Related works and extensions
- Discussion

Context

- **Great demand for « objective » measures** of productivity, relevance and value of scientific works from the different levels of society

- Historically measure of **journal influence** came first

- **Journal Impact Factor** based on average number of citations per article published in peer reviewed WOS publications during the 2 previous years

- **Objections against JIF** (& bibliometric methods)

Context/Objections

- **Epistemological**
  - ➢Peer review & expertise examination should prevail
  - ➢Quality of research is not based on a citation window
- **Ethical**
  - ➢Misuse, abuse and manipulation
- **Socio-political**
  - ➢JCR reshaped scientist behaviour, evaluation procedures, budget allocation and international science policy
  - ➢Indicators powered by one (or two) private companies
- **Technical**
  - ➢Two-year time window
  - ➢Self citations
  - ➢Negative and positive appreciations joined together
- see Archambault & Larivière (2009), Seglen (1992,1997)

Data Bases & Citation Metrics

- **Journal of Citation Reports** (Thomson-Reuters)
  - 11365 journals, 234 disciplines
  - Impact Factor (IF)
  - EigenFactor (EF)
  - Article Influence (AI)
- **Scopus Metrics** (Elsevier)
  - 21500 peer-reviewed (4200 open access), 116000 books
  - SCImago Journal Rank (SJR)
  - Impact per Publication (IPP)
- **Google Scholar** Metrics
  - H-index
- **Altmetrics**
  - Altmetric Attention Score

  (News, Blogs, Twitter, Sina Weibo, Facebook, YouTube, LinkedIn)

Journal Impact Factor (JIF)

- **Origins** of Journal Impact Factor: Archambault & Larivière (2009)
  - ➢Pioneer work of **Gross and Gross** (1927) to help US Librarians in journal selection
    - ➢« to measure the desirability of purchasing a particular journal »
    - ➢tabulate the references to a particular journal 'The Journal of the American Chemical Society'
  - ➢Hack (1936), Raisig (1960) ratio of no of quotes by no of articles published
  - ➢Martyn & Gilchrist (1967) two years citation window
- **Eugene Garfield**, Founder of the Institute of Scientific Information (ISI),
  - ➢First published JIF in 1975 and promoted it
  - ➢Flagship of Thomson-Reuters via Web of Science & Journal of Citation Reports (JCR)

JIF/Continued

- Defined as
  - IF(A)=N/D
  - N=No of cites in 2015 to items published by A in the 2 previous yrs (ie 2014,2013)
  - D=No of articles (citable items) published by A in the 2 previous years
- Example
  - Biometrika

Cites in 2015 to items published in:	2014 =47	Number of items published in:	2014 =77
	2013 =127		2013 =77
	Sum: 174		Sum: 154

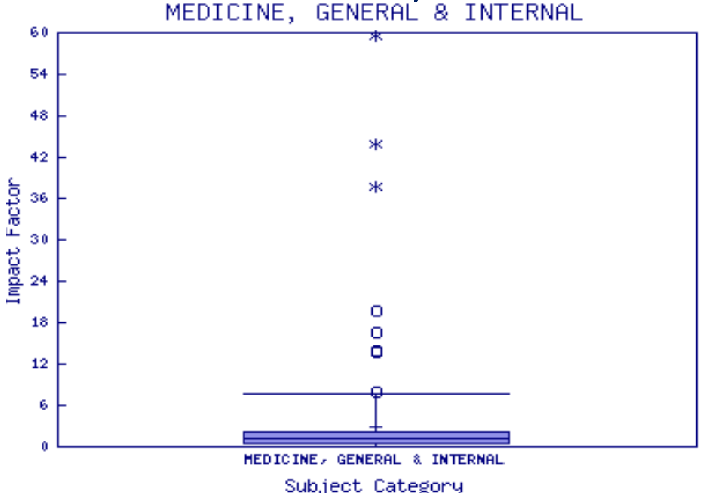
Calculation: $\frac{\text{Cites to recent items}}{\text{Number of recent items}} = \frac{174}{154} = \mathbf{1.130}$

Factors influencing JIF/Discipline

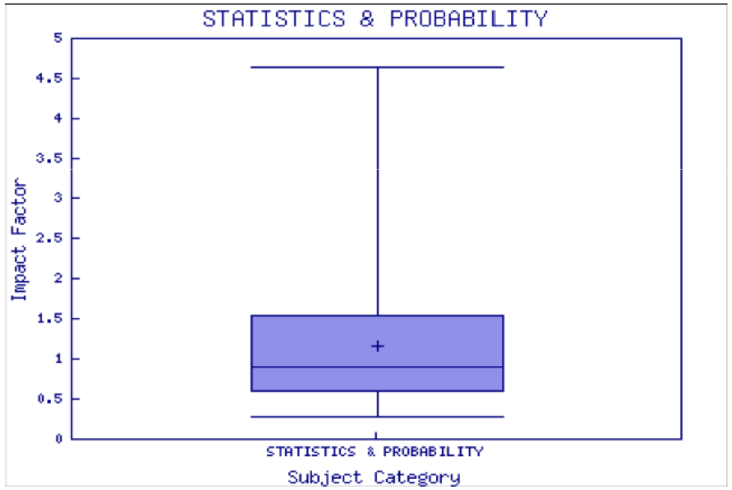
- Field dependent criterion

  - ➢ Medicine vs Probability & Statistics

4.105=178923/43598



1.128=19034/16872



Factors influencing JIF/Type of paper

- Reviews vs Research articles

Nature Review Immunology	*Immunity*
39.416=4927/125	**24.082**=7971/331
Statistical Science	*Biometrics*
2.213=166/75	**1.360**=287/211

- Theoretical vs Applied

JRSS-B	*JRSS-C*
4.222=304/72	**1.354**=107/79

Factors influencing JIF/Self Citations

- Included in IF

JRSS-B

IF=**4.222**=304/72 IF'=**4.097**=295/72

Self Citations=9/304=**3%** Total: 111/17404=**0.6**%

Scientometrics

IF=**2.084**=1236/593 IF'=**1.278**=758/593

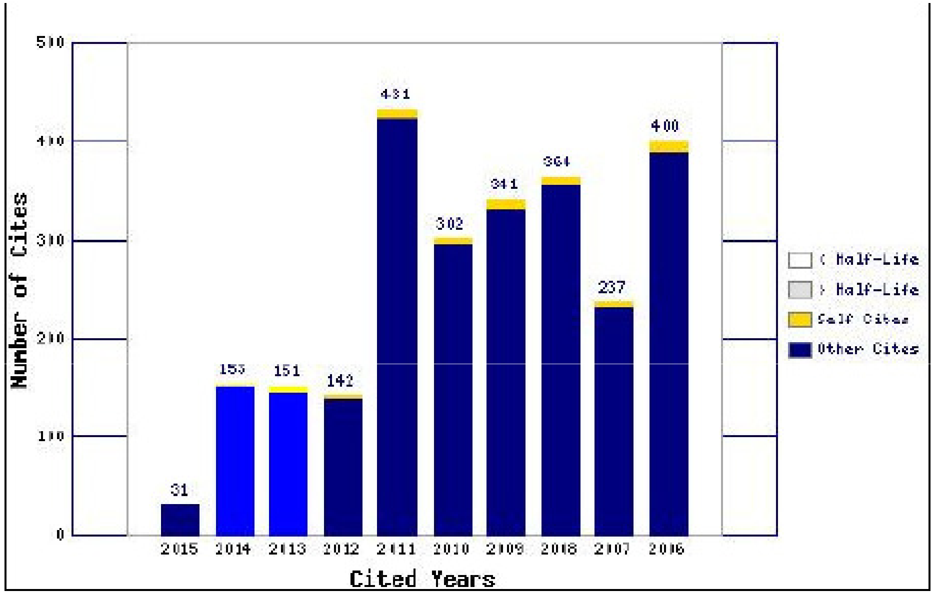
Self Citatins=478/1236=**38**% Total: 1963/6436=**30**%

- Coercition

« To ensure that a submitted manuscript meets sufficient interest of the readership it is expected that articles recently published in the Journal of.....are cited to a reasonable extent »

Factors influencing JIF/Time windows

JRSS-B*:* IF2=(153+151)/(38+34)=304/72=**4.222**



Distribution of citations received by the journal according to years of publication

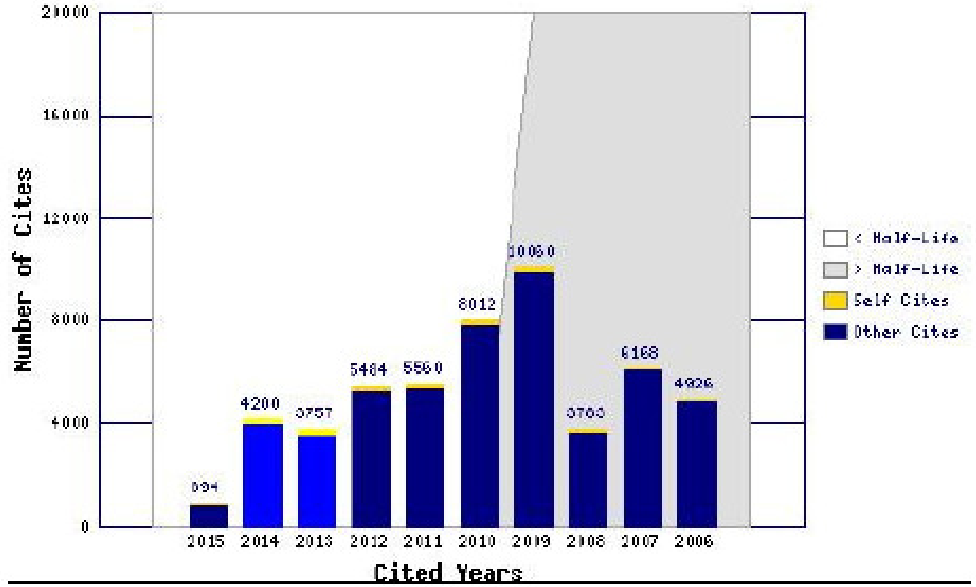
Cited Year	2015	2014	2013	2012	2011	2010	2009	2008	2007	2006	2005-all
# Cites from 2015	31	153	151	142	431	302	341	364	237	400	14852
Cumulative %	0.18	1.06	1.92	2.74	5.22	6.95	8.91	11.00	12.36	14.66	100

IF5=(153+151+142+431+302)/(38+34+32+29+29)=1179/162=**7.278**

Biometrika: IF2=1.130 IF5=2.016

Factors influencing JIF/Time windows

BIOINFORMATICS*:* IF2=(4200+3757)/(731+649)=7957/1380=**5.766**



Distribution of citations received by the journal according to years of publication

Cited Year	2015	2014	2013	2012	2011	2010	2009	2008	2007	2006	2005-all
# Cites from 2015	241	1811	2238	2202	1815	1694	1496	1126	907	750	5008
Cumulative %	1,25	10,64	22,24	33,66	43,07	51,85	59,61	65,44	70,15	74,04	100

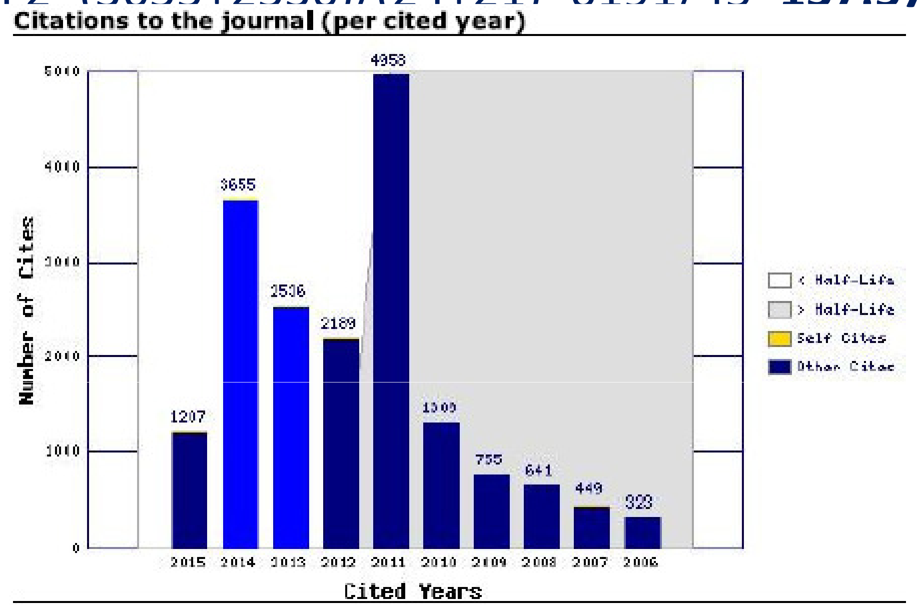
IF5=(4200+3757+….+8012)/(731+649+….+710)=27013/3515=**7.685**

PNAS: IF2=9.423 IF5=**10.285**

Factors influencing JIF/Time windows

CA- A Cancer Journal for Clinicians:

IF2=(3655+2536)/(24+21)=6191/45=**137.578**



Distribution of citations received by the journal according to years of publication

Cited Year	2015	2014	2013	2012	2011	2010	2009	2008	2007	2006	2005-all
# Cites from 2015	1207	3655	2536	2189	4958	1309	755	641	449	323	2466
Cumulative %	5.89	23.73	36.11	46.79	70.99	77.38	81.07	84.20	86.39	87.96	100

IF5=(3655+…+1309)/(24+…+18)=14647/101=**145.020**

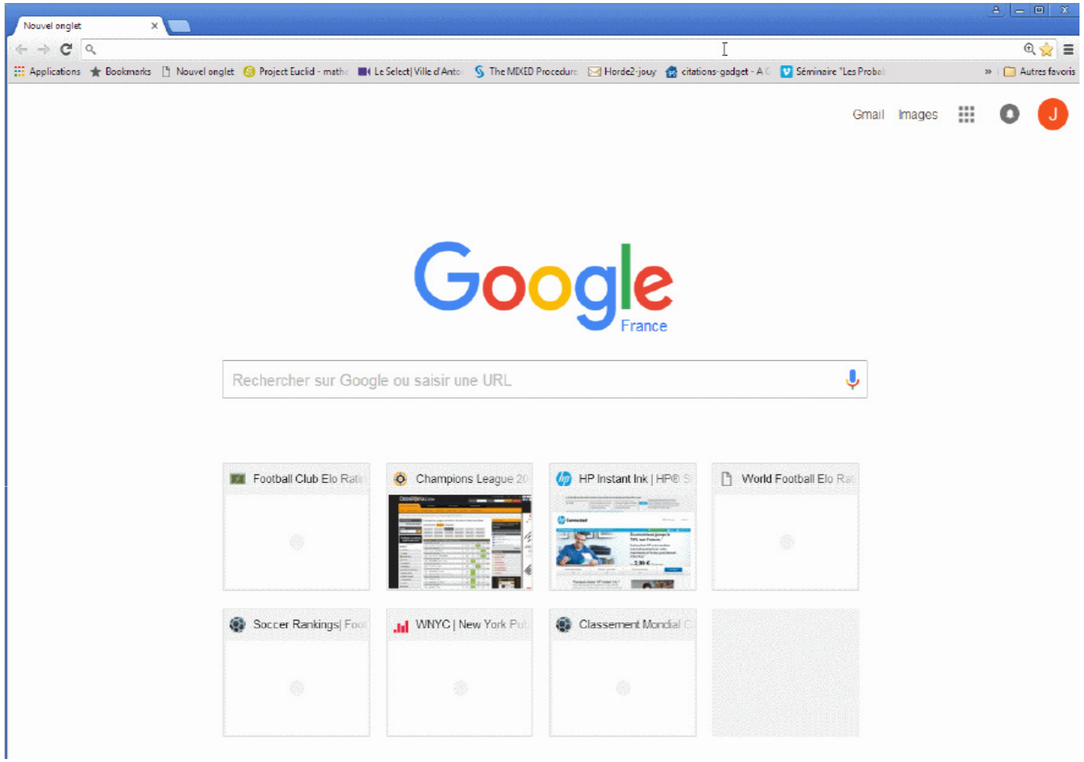
Technical issues with JIF

- **Field dependent criterion**
  - Medicine vs Probability & Statistics
- **Type of paper**
  - Review vs Research articles
- **Self Citations**
  - Included
- **Time window**
  - Standard 2 yrs (« accidental choice »), 5 yrs also published
  - Half-Life of cited papers quite longer (>10 yrs in stat) (Garfield & Sher, 1963)
- **Pareto Principle** (highly skewed distribution of citations)
  - 90 % IF of Nature due to 25% papers
- **Equal weight to all citations**
  - Does not depend of the importance of the citing journal

Technical issues with JIF

- **Asymmetry of the items valid for the numerator & denominator of JIF**
  - Numerator : any type of material
  - Denominator: only **citable items** with meeting abstracts, editorials, letters, corrections, book reviews, biographical items, theses, technical reports, patents **excluded**
  - « Non citable items are not uncitable » , they are cited, counted in the numerator but excluded from the denominator (Liu et al, 2016)
  - Lancet IF reduced by 40% when just counting citations from citable items (Moed et al, 1995)
  - Corrections needed for journals publishing a high proportion of non-citable items eg letters & editorials
  - Consistency in the source items of N & D in Scopus

GOOGLE



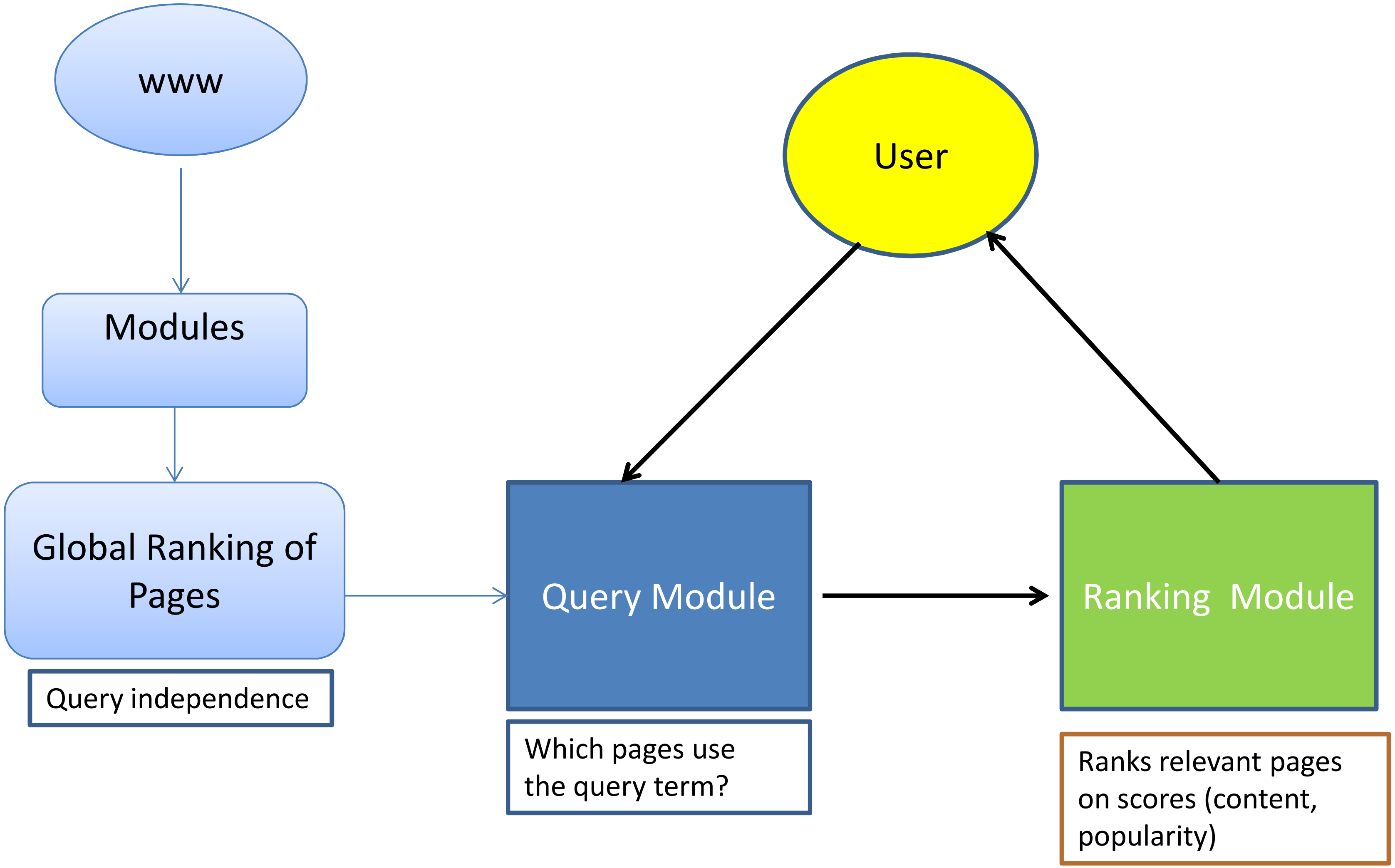
- Name Google: misspelling of Googol=10^100
- Headquarters in CA, Mountain View, Googleplex 10^Googol
- Cofounded by Larry Page & Sergey Brin in 1998
- Holding company « Alphabet » offering many products

- **Google Search** (Search Engine)
- **Google Chrome** (web browser
- Gmail (E mail)
- Google Drive (Cloud Storage)
- **You Tube** (video)
- Google translate
- **Google maps** (web mapping)
- Adsense (media advertisements)
- **Android** (mobile OS)
- Google Hangouts (video communication)
- Google Fiber (broadbrand internet and cable television access)
- Nexus (tablets)
- Google Pixel (smartphone)
- Google cars (self driving)
- **Waze** (GPS based geographical navigation)

Web Search Engines

• GOOGLE	71.1%
• BING	10.6%
• BAIDU	8.7%
• YAHOO	7.5%

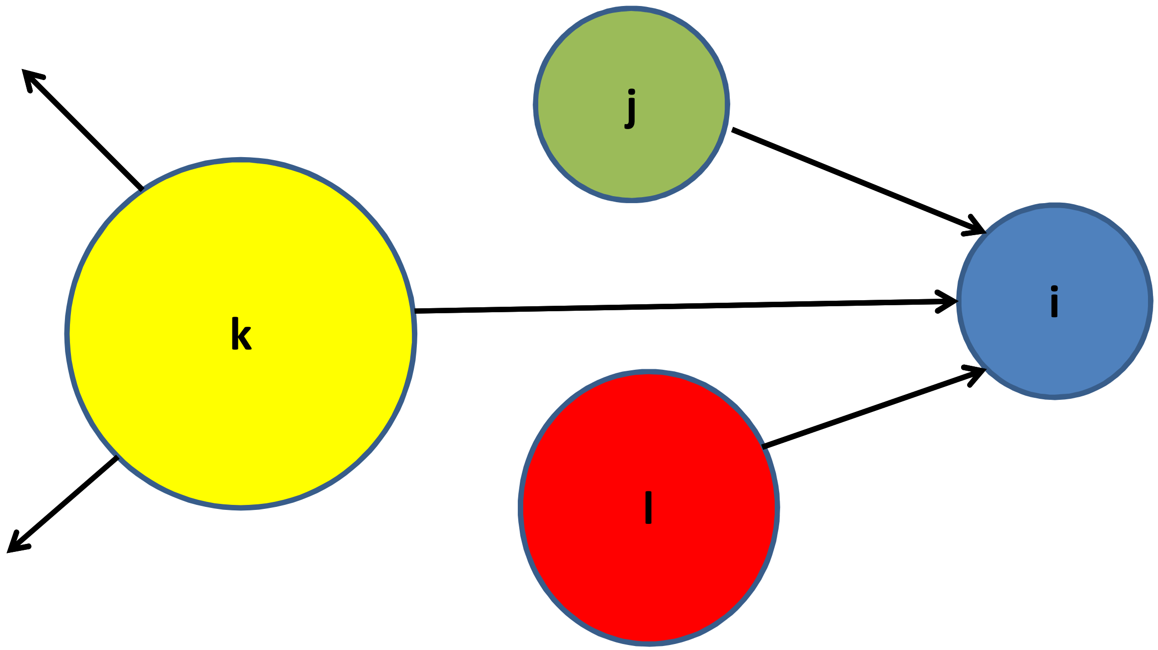

What is PageRank?
bloggingwithkids.com



Scheme of a Web Search Engine Machine (from Langville & Meyer, 2012)

Basic Page Rank/Principles

- PageRank is based on link structure to capture « importance » of webpage
- Pages (j,k,l) pointing to i =inlinks to i & outlinks from j,k,l
- Pages having high PR (k) are given more weight than j,l
- Pages (k) linking to other pages (k) are given less weight than j,l



Brin & Page's 1998 article

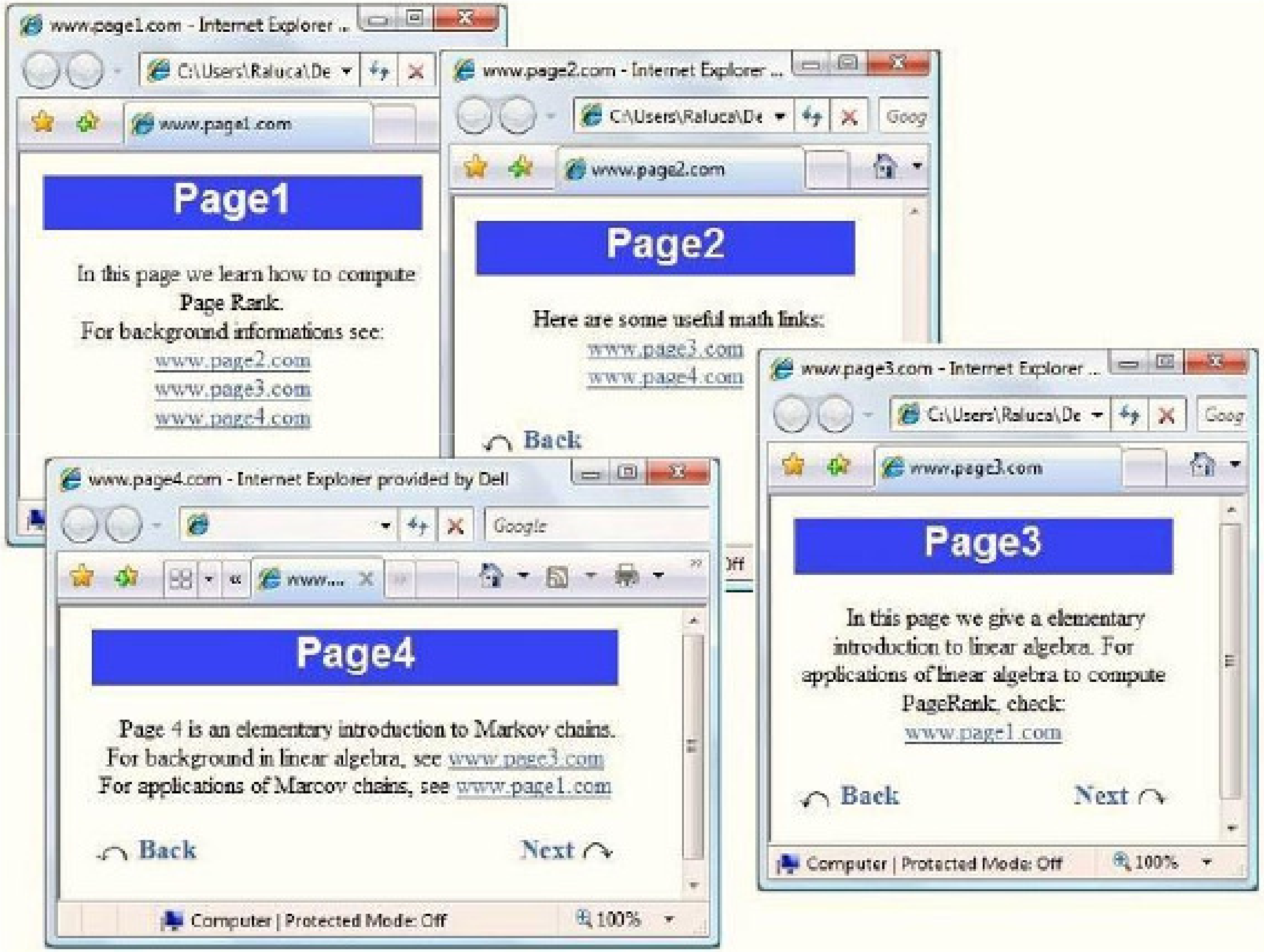
2.1.1 Description of PageRank Calculation

Academic citation literature has been applied to the web, largely by counting citations or backlinks to a given page. This gives some approximation of a page's importance or quality. PageRank extends this idea by not counting links from all pages equally, and by normalizing by the number of links on a page. PageRank is defined as follows:

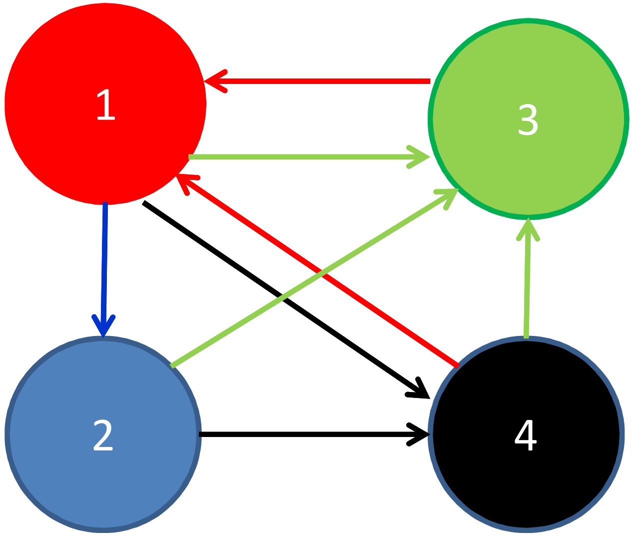
> *We assume page A has pages T1...Tn which point to it (i.e., are citations). The parameter d is a damping factor which can be set between 0 and 1. We usually set d to 0.85. There are more details about d in the next section. Also C(A) is defined as the number of links going out of page A. The PageRank of a page A is given as follows:*
>
> $$PR(A) = (1-d) + d\,(PR(T1)/C(T1) + \ldots + PR(Tn)/C(Tn))$$
>
> *Note that the PageRanks form a probability distribution over web pages, so the sum of all web pages' PageRanks will be one.*

PageRank or *PR(A)* can be calculated using a simple iterative algorithm, and corresponds to the principal eigenvector of the normalized link matrix of the web. Also, a PageRank for 26 million web pages can be computed in a few hours on a medium size workstation. There are many other details which are beyond the scope of this paper.

Basic Page Rank/Example



Basic Page Rank/Network & Algorithm



r(1)=r(3)/1+r(4)/2
r(2)=r(1)/3
r(3)= r(1)/3+ r(2)/2+ r(4)/2
r(4)= r(1)/3 +r(2)/2

$\{1\} \rightarrow \{2,3,4\}$
$\{2\} \rightarrow \{3,4\}$
$\{3\} \rightarrow \{1\}$
$\{4\} \rightarrow \{3,1\}$

B_i = subset of pages (nodes) pointing to i

d_j = #outlinks from page j

$$r_i = \sum_{j \in B_i} \frac{r_j}{d_j}$$

Basic Page Rank/Ex/Solutions

0	0.25	0.25	0.25	0.25
1	0.375	0.08333333333	0.3333333333	0.2083333333
2	0.4375	0.125	0.2708333333	0.1666666667
3	0.3541666667	0.1458333333	0.2916666667	0.2083333333
4	0.3958333333	0.1180555556	0.2951388889	0.1909722222
5	0.390625	0.1319444444	0.2864583333	0.1909722222
6	0.3819444444	0.1302083333	0.2916666667	0.1961805556
7	0.3897569444	0.1273148148	0.2905092593	0.1924189815
8	0.38671875	0.1299189815	0.2897858796	0.1935763889
9	0.3865740741	0.12890625	0.2906539352	0.1938657407
10	0.3875868056	0.1288580247	0.2902440201	0.1933111497
23	0.3870967818	0.129032167	0.2903226496	0.1935484017
34	0.3870967742	0.1290322579	0.2903225807	0.1935483871

r=(12, 4, 9, 6)/31

Basic Page Rank/Basic theory

$\{1\} \rightarrow \{2,3,4\}$

$\{2\} \rightarrow \{3,4\}$

$\{3\} \rightarrow \{1\}$

$\{4\} \rightarrow \{1,3\}$

$$\underbrace{\mathbf{A}}_{\substack{Ajacency\\ Matrix}} = \begin{bmatrix} 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 \\ 1 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 \end{bmatrix} \quad \mathbf{A1} = \underbrace{\mathbf{d}^{out}}_{\#\text{outdegrees}} = \{a_{i+}\} = \begin{bmatrix} 3 \\ 2 \\ 1 \\ 2 \end{bmatrix}$$

$$\mathbf{1}^T \mathbf{A} = \underbrace{\mathbf{d}^{in,T}}_{\#\text{in degrees}} = \begin{bmatrix} 2 & 1 & 3 & 2 \end{bmatrix}$$

$$\underbrace{\mathbf{P}}_{\text{Transition Matrix}} = \mathbf{D}^{-1}\mathbf{A} = \begin{bmatrix} 0 & 1/3 & 1/3 & 1/3 \\ 0 & 0 & 1/2 & 1/2 \\ 1 & 0 & 0 & 0 \\ 1/2 & 0 & 1/2 & 0 \end{bmatrix} \quad \mathbf{D} = Diag\left(\mathbf{d}^{out}\right)$$

$$p_{ij} = \Pr(s^{(n+1)} = j \mid s^{(n)} = i) = a_{ij} / a_{i+}$$

PageRank algorithm

$$\mathbf{r}^{(k+1)T} = \mathbf{r}^{(k)T}\mathbf{P} \quad r_j^{(k+1)} = \sum_{i=1}^{N} p_{ij} r_i^{(k)} \quad j = 1, 2, \ldots, N$$

Basic Page Rank/Markov Chain

$\mathbf{P} = \left\{ p_{ij} \right\} = \mathbf{D}^{-1}\mathbf{A}$ transition matrix

$\mathbf{r}^{(1)T} = \mathbf{r}^{(0)T}\mathbf{P}, \mathbf{r}^{(2)T} = \mathbf{r}^{(1)T}\mathbf{P} = \mathbf{r}^{(0)T}\mathbf{P}^2, \mathbf{r}^{(n)T} = \mathbf{r}^{(0)T}\mathbf{P}^n$

If $\mathbf{P}$ irreducible (graph strongly connected) and chain aperiodic

(not visited in a number of steps multiple of an integer>1)

Unique Stationary distribution $\tilde{\mathbf{r}}$ of the DTMC $\mathbf{r}^{(n)} \underset{n\to\infty}{\to} \tilde{\mathbf{r}}$ for any $\mathbf{r}^{(0)}$

$\tilde{\mathbf{r}}^T \to \left(p_{i1}^{(n)}, ..., p_{ij}^{(n)}, ..., p_{iN}^{(n)} \right)$ for any row i where $p_{ij}^{(n)} = \left[\mathbf{P}^n \right]_{ij}$

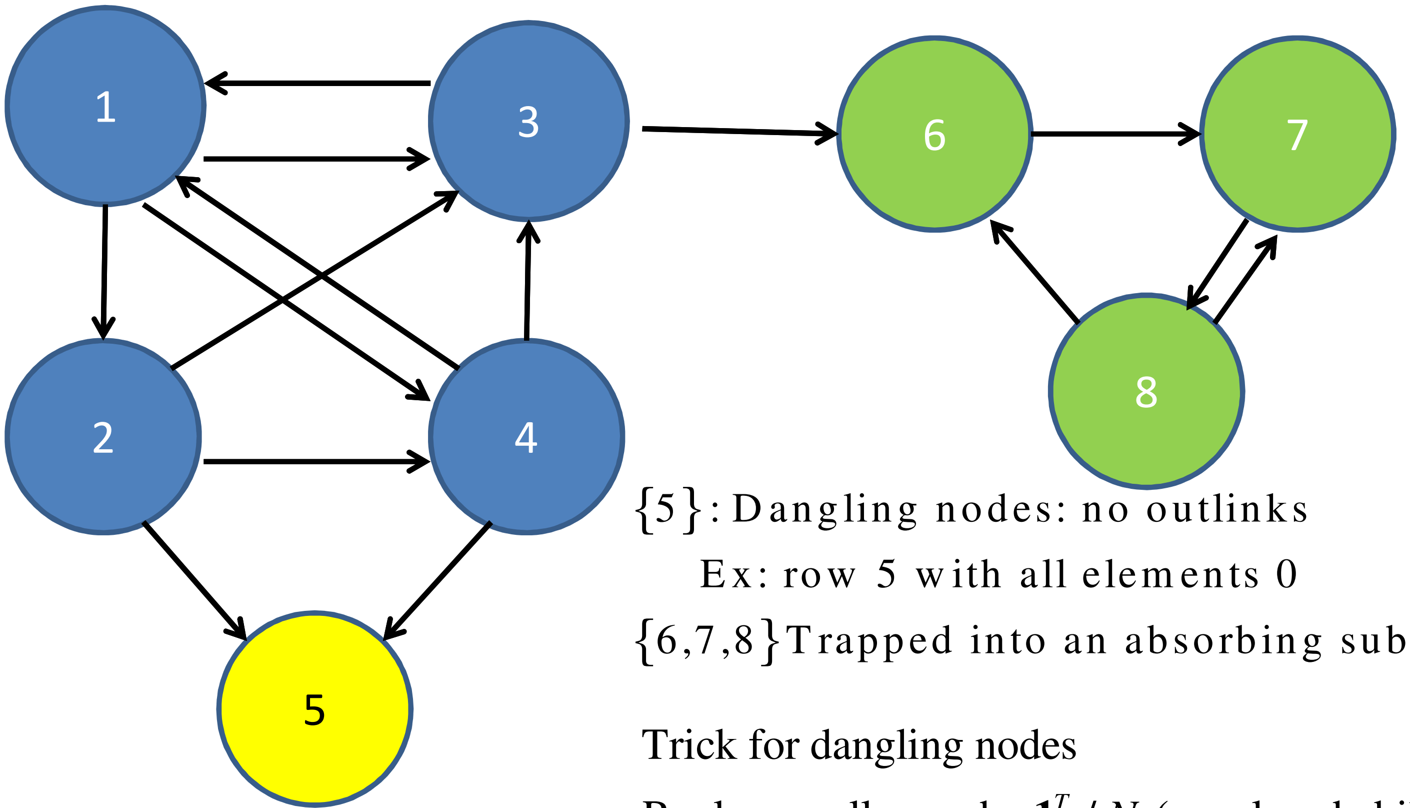
$\boxed{\tilde{\mathbf{r}}^T = \tilde{\mathbf{r}}^T\mathbf{P}} \quad \tilde{r}_j = \sum_{i=1}^{N} p_{ij}\tilde{r}_i \quad j = 1, 2, ..., N$

$\tilde{\mathbf{r}}$ =Left eigenvector pertaining to the dominant eigenvalue of $\mathbf{P}$

(here 1) such that $\left\| \tilde{\mathbf{r}} \right\|_1 = 1$

$\tilde{r}_j$ = long-term fraction of time the process spends at page (node) j

Basic Page Rank/Google Matrix



$\{5\}$: Dangling nodes: no outlinks

Ex: row 5 with all elements 0

$\{6,7,8\}$ Trapped into an absorbing subchain

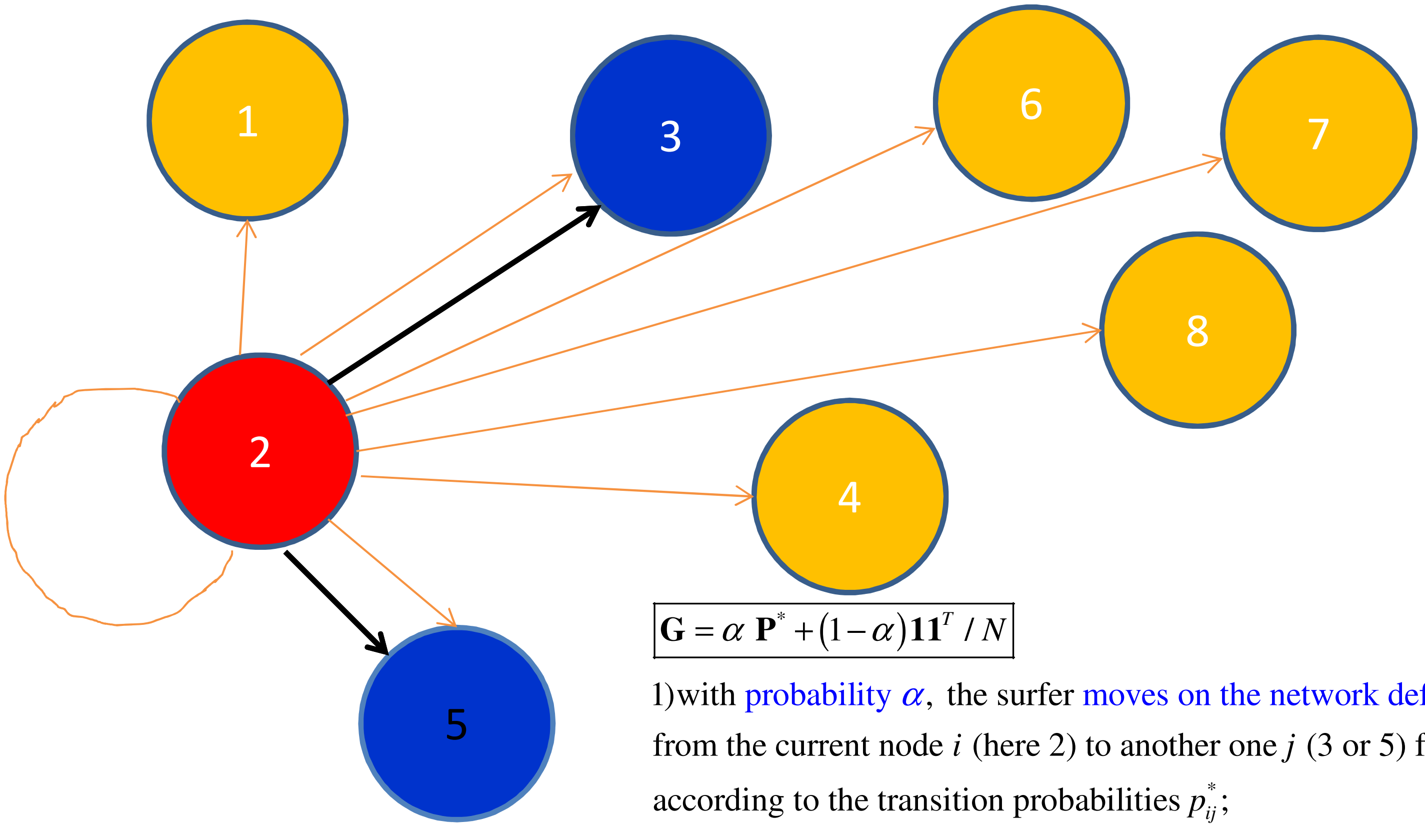
Trick for dangling nodes

Replace null rows by $\mathbf{1}^T / N$ (equal probabilities)

If $\mathbf{b} = \{b_i\}_{1 \le i \le N}$ with $b_i = 1$ if $i = DN$, 0 otherwise

Then $\mathbf{P}^* = \mathbf{P} + \mathbf{b}\mathbf{1}^T / N$

Basic Page Rank/Random Surfer



$$\boxed{\mathbf{G} = \alpha\, \mathbf{P}^{*} + (1-\alpha)\mathbf{1}\mathbf{1}^{T} / N}$$

1)with probability α, the surfer moves on the network defined by $\mathbf{P}^{*}$
from the current node i (here 2) to another one j (3 or 5) following an edge
according to the transition probabilities p_{ij}^{*};

2)with probability 1-α, the surfer "teleports" at random to any node j
More generally with transition probabilities $(u_1,\ldots,u_j,..,u_N)$

α = damping factor usually set empirically to 0.85

G called the Google matrix is stochastic, irreducible, aperiodic

PageRank is the the stationary distribution of the corresponding DTMC
defined as $\mathbf{r}^{T} = \mathbf{r}^{T}\mathbf{G}$ and calculated using the Power Method

Basic Page Rank/Calculation

$$\boxed{\mathbf{G} = \alpha\,\mathbf{P}^{*} + (1-\alpha)\mathbf{1}\mathbf{1}^{T}/N}$$

$$\boxed{\mathbf{r}^{(k+1)T} = \alpha\mathbf{r}^{(k)T}\mathbf{P}^{*} + (1-\alpha)\mathbf{1}^{T}/N}$$ Power M: Brin & Page

$$\boxed{\mathbf{r}^{(k+1)T} = \alpha\mathbf{r}^{(k)T}\mathbf{P} + \left[\alpha\mathbf{r}^{(k)T}\mathbf{b} + (1-\alpha)\right]\mathbf{1}^{T}/N}$$

$\mathbf{P}$: Extremely sparse matrix, never formed nor stored

Alternative by solving a Linear System

$$\mathbf{x}^{T}\left[\mathbf{I} - \alpha\mathbf{P}^{*}\right] = (1-\alpha)\mathbf{1}^{T}/N \text{ with } \mathbf{r}^{T} = \mathbf{x}^{T}/|\mathbf{x}|_{1}$$

Application to Rating & Ranking Scientific Journals/Cross Citation Matrix

« Fundamental entity which contains the informatrion describing the flow of influence among units » Pinski & Narin, 1976

Cason & Lubotsky, 1936, Psychological B Daniel & Louttit, 1953, Psychology

Kessler, 1964, Physics Price, 1981, Scientometrics

Todorov & Glänezel, 1988, Review, J of Information Science

Stigler, Stigler & Friedland, 1992, Economics; Stigler, 1994 Varin et al, 2016

Cij=No of references journal i gives in a given period (T1) to papers published previously by journal j within a given target window (T2) ex T2=[t-1,t-2]

In Short Cij: No of cites i gives to j

T1: usually 1yr ex 2015

T2: Target Window: 2 yrs eg 2014,2013 or 5 yrs eg 2014, 2013,2012,2011,2010

Rows: i citing (issuing references)

Columns: j cited (receiving citations)

R &R Journals/Cross Citation Matrix/Ex

Cross Citation Matrix for a subset of 12 journals from Varin et al (2016)

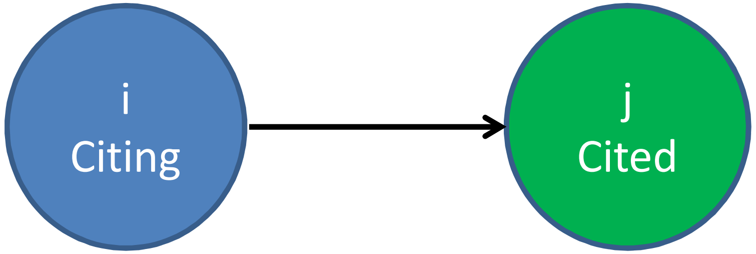
	AOS	BCS	BKA	CSSC	ENVR	ISR	JABES	JASA	JRSS-B	JSS	STCMP	STMOD
AOS	291	15	38	0	0	3	0	121	76	4	5	0
BCS	29	191	60	2	0	0	7	129	45	0	11	3
BKA	74	31	75	0	1	0	0	78	41	2	4	0
CSSC	19	11	11	29	1	3	1	29	12	4	2	4
ENVR	6	8	8	0	31	2	8	20	14	2	1	3
ISR	0	2	5	0	0	20	0	6	4	0	1	0
JABES	2	30	5	0	3	0	16	20	5	2	0	0
JASA	126	69	93	3	11	5	8	232	101	0	13	2
JRSS-B	38	5	21	0	0	1	0	40	55	2	8	0
JSS	7	13	7	0	2	1	0	13	12	91	1	4
STCMP	34	13	17	1	4	0	1	45	33	0	24	4
SMOD	2	5	9	00	0	1	0	13	8	3	0	7

Citing in rows; Cited in columns

Cij: References given in 2010 by journals i to papers published in journals j during 2001-2010

AOS: Annals of Statistics, BCS: Biometrics, BKA: Biometrika, CSSC: Communication in Statistics-Simulation & Computation, ENVR: Evironmentrics, ISR: International Statistical Review, JABES: Journal of Agricultural, Biological and Environmental Statistics, JASA: Journal of the American Statistical Science, JRSS-B: Journal of the Royal Statistical Society, Series B, JSS: Journal of Statistical Software, STCMP: Statistics & Computing, STMOD: Statistical Modelling

R&R Journals/EigenFactor



Transition Matrix $\mathbf{P}$: From i citing to j cited

$$p_{ij} = c_{ij} / c_{i+}$$

c_{ij} = No of references issued by i during 1 yr to papers of j published 5yrs before

c_{i+} = No of references issued by i during 1 yr to all papers published 5yrs before

Row i= $\left[p_{i1}, p_{i2}, ..., _{ij}, ..., p_{iN} \right]$

$$c_{i+} = \sum_{j=1}^{N} c_{ij}, \ \sum_{j=1}^{N} p_{ij} = 1$$

Note : $\mathbf{P}$ called $\mathbf{H}^T$ in Bergstrom (2007) Eigenfactor: measuring the value and prestige of scholarly journals. College & Research Libraries News, 68, 314-316

R& R Journals/EF & AI

Dangling Nodes: null rows replaced by $\tilde{\mathbf{a}}^T = (\tilde{a}_i)^T_{1 \le i \le N}$ with $\tilde{a}_i = a_i / \sum_{k=1}^{N} a_k$

$b = (b_i)_{1 \le i \le N}$ $b_i = 1$ for $i \in$ DN, 0 otherwise

a_i = No of articles published by i during the 5yr time window $\mathbf{P}^* = \mathbf{P} + \mathbf{b}\tilde{\mathbf{a}}^T$

Google Matrix: $\boxed{\mathbf{G} = \alpha \, \mathbf{P}^* + (1-\alpha)\mathbf{1}\tilde{\mathbf{a}}^T}$

Let $\mathbf{r}$ "Weighted Page Rank" be the leading Eigenvector vector $\mathbf{r}^T\mathbf{G} = \mathbf{r}^T$

$$r_j = \alpha \sum_{i=1}^{N} p^*_{ij} r_i + (1-\alpha)\tilde{a}_j \underbrace{\sum_{i=1}^{N} r_i}_{1}$$

Then EigenFactor defined as $\boxed{\mathrm{EF}_i = \frac{\left[\mathbf{P}^T\mathbf{r}\right]_i}{\sum_{k=1}^{N}\left[\mathbf{P}^T\mathbf{r}\right]_k}}$ Total influence of journal i

EF defined not as $\mathbf{r}$ but using $\mathbf{P}$ without self citations, dangling nodes, teleportation

The Article Influence (AI) score $\boxed{\mathrm{AI}_i = [\mathbf{EF}]_i / \tilde{a}_i}$

AI=Per article influence as JIF; Size free score

Journal Abbreviations (from Varin et al, 2016)

Journal name	Abbreviation
American Statistician	AmS
Annals of the Institute of Statistical Mathematics	AISM
Annals of Statistics	AoS
Australian and New Zealand Journal of Statistics	ANZS
Bernoulli	Bern
Biometrical Journal	BioJ
Biometrics	Bcs
Biometrika	Bka
Biostatistics	Biost
Canadian Journal of Statistics	CJS
Communications in Statistics—Simulation and Computation	CSSC
Communications in Statistics—Theory and Methods	CSTM
Computational Statistics	CmpSt
Computational Statistics and Data Analysis	CSDA
Environmental and Ecological Statistics	EES
Environmetrics	Envr
International Statistical Review	ISR
Journal of Agricultural, Biological and Environmental Statistics	JABES
Journal of the American Statistical Association	JASA
Journal of Applied Statistics	JAS
Journal of Biopharmaceutical Statistics	JBS
Journal of Computational and Graphical Statistics	JCGS
Journal of Multivariate Analysis	JMA

Journal name	Abbreviation
Journal of Nonparametric Statistics	JNS
Journal of the Royal Statistical Society, Series A	JRSS-A
Journal of the Royal Statistical Society, Series B	JRSS-B
Journal of the Royal Statistical Society, Series C	JRSS-C
Journal of Statistical Computation and Simulation	JSCS
Journal of Statistical Planning and Inference	JSPI
Journal of Statistical Software	JSS
Journal of Time Series Analysis	JTSA
Lifetime Data Analysis	LDA
Metrika	Mtka
Scandinavian Journal of Statistics	SJS
Stata Journal	StataJ
Statistical Methods in Medical Research	SMMR
Statistical Modelling	StMod
Statistica Neerlandica	StNee
Statistical Papers	StPap
Statistical Science	StSci
Statistica Sinica	StSin
Statistics	Stats
Statistics and Computing	StCmp
Statistics in Medicine	StMed
Statistics and Probability Letters	SPL
Technometrics	Tech
Test	Test

R& R Journals/Example

Rating & Ranking of 47 Statistical Journals in 2015 on Impact Factor (IF), EigenFactor (EF) and Article Influence Scores (AI) from citations given in 2015 to items published 2 yrs (IF) or 5 yrs before (EF,AI)

(Source Journal Citation Reports)

No	Journal	IF	No	Journal	EF	No	Journal	AI
1	**SMMR**	4.634	1	AOS	0.04095	1	JRSS-B	7.822
2	JRSS-B	4.222	2	STMED	0.03354	2	JSS	5.283
3	AOS	2.780	3	JASA	0.03342	3	AOS	5.099
4	JSS	2.379	4	JSS	0.03124	4	STSCI	3.757
5	STSCI	2.213	5	CSDA	0.02068	5	JASA	3.641
6	*BIOST*	2.109	6	JRSS-B	0.01916	6	SMMR	2.877
7	STCMP	1.786	7	BKA	0.01630	7	BKA	2.785
8	JCGS	1.755	8	BCS	0.01548	8	JCGS	2.299
9	JASA	1.725	9	JSPI	0.01460	9	*BIOST*	2.017
10	JRSS-A	1.702	10	JMA	0.01440	10	STCMP	2.011
11	STMED	1.533	11	STSIN	0.01200	11	STSIN	1.978
12	TECH	1.435	12	SPL	0.01142	12	ISR	1.890
13	BERN	1.372	13	STSCI	0.01000	13	BERN	1.692
14	BCS	1.360	14	BERN	0.00955	14	STATAJ	1.639
15	JRSS-C	1.354	15	STCMP	0.00934	15	TECH	1.604
16	STATAJ	1.292	16	JCGS	0.00915	16	TEST	1.582
17	ISR	1.240	17	*BIOST*	0.00842	17	BCS	1.548
18	AMS	1.215	18	SMMR	0.00740	18	JRSS-A	1.459
19	TEST	1.207	19	CSTM	0.00554	19	STMED	1.394
20	CSDA	1.179	20	STATAJ	0.00550	20	JRSS-C	1.331
21	ENVR	1.160	21	JSCS	0.00526	21	SJS	1.304
22	BKA	1.130	22	SJS	0.00479	22	JTSA	1.153
23	JTSA	1.000	23	JRSS-A	0.00457	23	JMA	1.072
24	STMOD	0.932	24	JAS	0.00454	24	ENVR	1.009

R & R Journals/Example

Rating & Ranking of 47 Statistical Journals in 2015 on Impact Factor (IF), EigenFactor (EF) and Article Influence Scores (AI) from citations given in 2015 to items published 2 yrs (IF) or 5 yrs before (EF,AI)

Source: Journal Citation Reports

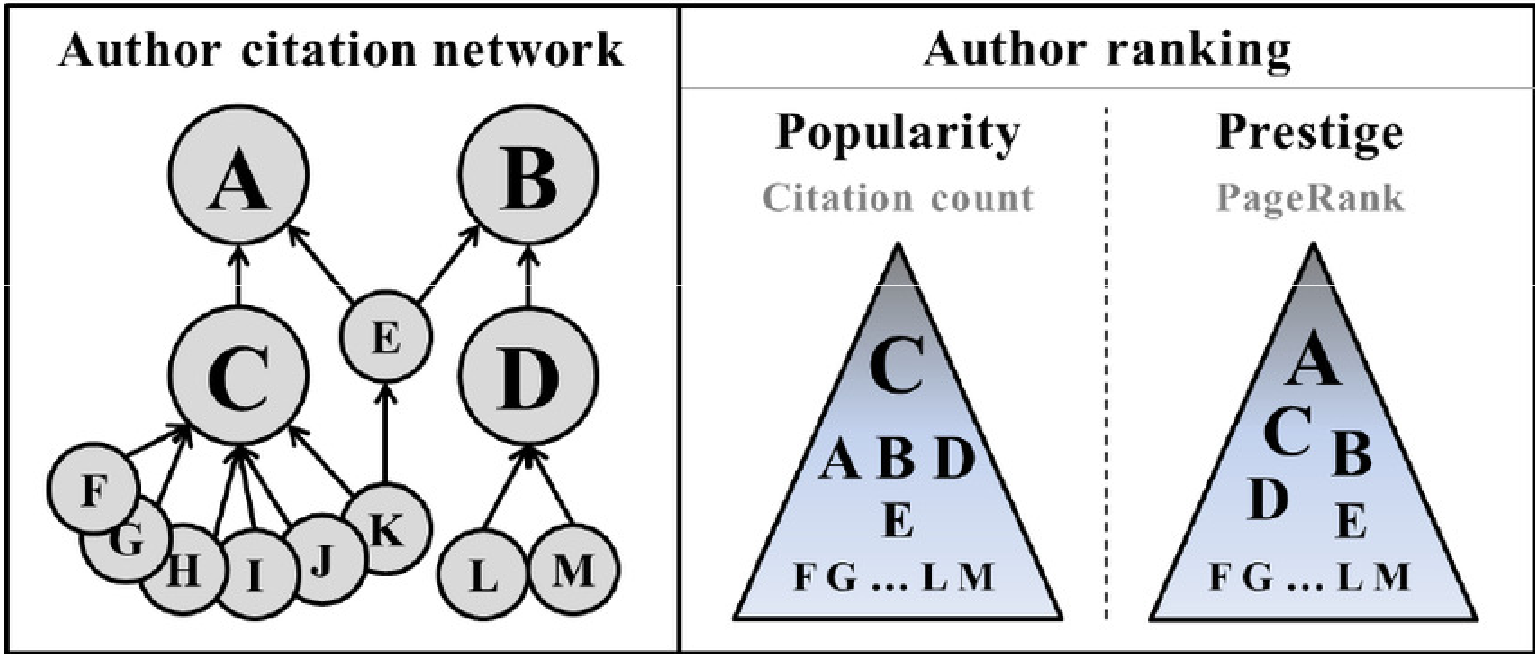
No	Journal	IF	No	Journal	EF	No	Journal	AI
25	SJS	0.908	25	ENVR	0.00443	25	CSDA	0.939
26	JBS	0.882	26	TECH	0.00439	26	AMS	0.915
27	JMA	0.857	27	JTSA	0.00415	27	AISM	0.836
28	STSIN	0.838	28	JRSS-C	0.00413	28	STMOD	0.830
29	LDA	0.810	29	BIOJ	0.00346	29	LDA	0.812
30	JABES	0.790	30	TEST	0.00335	30	BIOJ	0.791
31	STPAP	0.781	31	AISM	0.00315	31	JSPI	0.770
32	EES	0.769	32	JBS	0.00313	32	CJS	0.753
33	AISM	0.768	33	JNS	0.00302	33	STNEE	0.738
34	JSCS	0.749	34	ISR	0.00292	34	JNS	0.726
35	JSPI	0.727	35	CSSC	0.00260	35	JABES	0.690
36	BIOJ	0.683	36	AMS	0.00248	36	JBS	0.522
37	MTKA	0.595	37	STPAP	0.00245	37	JSCS	0.506
38	STATS	0.530	38	CJS	0.00237	38	STPAP	0.504
39	CMPST	0.520	39	CMPST	0.00221	39	MTKA	0.502
40	SPL	0.506	40	MTKA	0.00205	40	ANZS	0.495
41	JNS	0.446	41	STATS	0.00180	41	SPL	0.493
42	STNEE	0.432	42	LDA	0.00178	42	EES	0.442
43	ANZS	0.431	43	JABES	0.00172	43	CMPST	0.417
44	JAS	0.419	44	STMOD	0.00148	44	STATS	0.358
45	CJS	0.413	45	STNEE	0.00138	45	JAS	0.343
46	CSSC	0.397	46	EES	0.00118	46	CSSC	0.251
47	CSTM	0.300	47	ANZS	0.00096	47	CSTM	0.239

R& R Journals/IF & AI

Despite high correlation between IF & AI, there are major changes in ranking

Journal	IF	AI	Quartile	Items published
BIOMETRICS	1.360	1.548	Q2	661
CSDA	1.179	0.939	Q2	1457
BIOMETRIKA	1.130	2.785	Q1	387

Rating & Ranking Journals/Prestige vs Popularity



R& R/Previous works/Pinski-Narin

Pinski G & Narin F (1976): Influence Weight

$$W_j^{(k+1)} = \sum_{i=1}^{N} \frac{c_{ij}}{s_j} W_i^{(k)} = 1, 2, \ldots, N \quad ; s_i = \sum_{k=1}^{N} c_{ik} = c_{i+}$$

normalized such that $\sum_{j=1}^{N} W_j^{(k+1)} s_j / \sum_{j=1}^{N} s_j = 1$

$W_j^{(0)} = 1$, $\boxed{W_j^{(1)} = c_{+j} / c_{j+}}$ = cites received/cites given by j

(Arbel & Robert, 2015)

$$\underbrace{W_j s_j / s_+}_{r_j} = \sum_{i=1}^{N} \frac{c_{ij}}{s_i} \underbrace{W_i s_i / s_+}_{r_i} = \sum_{i=1}^{N} p_{ij} \underbrace{W_i s_i / s_+}_{r_i}, \text{, with } \sum_{i=1}^{N} r_i = 1$$

Equivalent to $\boxed{\mathbf{r}^T \mathbf{P} = \mathbf{r}^T}$ with $\boxed{W_i = \frac{r_i}{s_i / s_+}}$

s_i / s_+ = Proportion of references journal i gives to other journals

R& R/Previous works/Pinski-Narin

Influence Weight **Pinski G & Narin F** (1976)

Equivalent to $\boxed{\mathbf{r}^T \mathbf{P} = \mathbf{r}^T}$ and $\boxed{W_i = \frac{r_i}{s_i / s_+}}$

$s_i / s_+ =$ Proportion of references journal i gives to other journals

$\text{AI}_i = \frac{r_i}{\tilde{a}_i} \quad \tilde{a}_i = a_i / a_+ =$ Prop. of total papers from journal i published in $t-1,...,t-T2$

Total influence: $\text{TI}_i = W_i s_i = s_+ r_i \propto "EF"_i$ without teleportation (α=1)

Influence per citation "Weight influence": W_i

Influence per publication: $\text{IPP}_i = W_i s_i / \tilde{a}_i$ $\boxed{\text{AI without teleportation } (\alpha=1)}$

R& R/Related approaches

Invariant Method : Palacios-Huerta I & Volij (2004)

(Axiomatic approach)

$$v_j = \sum_{i=1}^{N} \frac{c_{ij}}{a_j} \frac{a_i}{c_{i+}} v_i \quad j = 1, 2, \ldots, N$$

c_{i+} / a_i=Reference intensity $\quad v^T = v^T D_a D_c^{-1} C D_a^{-1}$

$$v^T D_a = v^T D_a D_c^{-1} C \Leftrightarrow \underline{v}^T = \underline{v}^T D_c^{-1} C$$

$v \approx$ 'Article Influence' without teleportation

R& R/Related approaches

Audience Factor: Zitt & Small (2008), Zitt (2010)

Citations weighted by the reciprocal of the average no of references per article m_i from the citing journal i in $t-1,\ldots,t-5$ yrs (to correct for field dependency of IF)

$$AF_j = \sum_{i=1}^{N} \frac{c_{ij}}{a_j} w_i \quad w_i = m_+ / m_i$$

If $v_i = cst$ (ie α=0 in EF) $\Rightarrow IM_i \propto AF_i$

R& R /Previous works/SCImago

Borja González-Pereira, Vicente P Guerrero-Bote, **Félix Moya - Anegón**

(2010) *The SJR indicator. Journal of informetrics* 4, 379–391

$$PSJR_j^{(k+1)} = \underbrace{(1-d-e)/N}_{(1)} + \underbrace{e\tilde{a}_j}_{(2)} + \underbrace{d\left(\sum_{i=1}^{N} p_{ij} PSJR_i^{(k)} C_F + \tilde{a}_j \sum_{i \in DN} PSJR_i^{(k)}\right)}_{(3)}$$

(1),(2)"teleportation" (3) Influence of journals citing j

$d = 0.9; e = 0.0999$ $\tilde{a}_i = a_i / a_+$; $C_F = (1\text{-}\mathbf{r}^T\mathbf{b}) / \left(\mathbf{r}^T\mathbf{P1}\right)$ normalizing constant

$SJR_i = PSJR_i / \tilde{a}_j$: Prestige Size Free Index (by article normalisation as in AI)

Main differences with EF, AI

- –PSJR based completely on Weighted PageRank
- –Damping factor: 0.9 in SJR vs 0.85 in EF
- –3yrs Time Window vs 5 yrs in EF,AI
- –Self citations included (max 33%)
- –Using Scopus data base (Elsevier) vs Thomson-Reuters

SCImago. (2007). *SJR — SCImago Journal & Country Rank.*

Retrieved July 21, 2015, *from http://www.scimagojr.com*

R&R/Bayesian reinterpretation of the G matrix

i^{th} (citing) row of **G** is $\boxed{\mathbf{G}_i^T = \alpha \mathbf{p}_i^T + (1-\alpha)\tilde{\mathbf{a}}^T}$

with element *j*: $g_{ij} = \alpha p_{ij} + (1-\alpha)\tilde{a}_j$

i^{th} row of **C** is $\mathbf{C}_i^T = (c_{i1}, c_{i1}, ..., c_{ij}, ..., c_{iN})$

1) $\boxed{\mathbf{C}_i^T \sim \mathcal{M}(s_i, \boldsymbol{\theta}_i^T)}$ $\boldsymbol{\theta}_i = \{\theta_{ij}; j = 1, ..., N\}$ where $s_i = \sum_{j=1}^{N} c_{ij}$

s_i = No of citations given by i to all journals (outlinks from i)

2) $\boxed{\boldsymbol{\theta}_i \mid \boldsymbol{\gamma} \sim \mathcal{D}(\boldsymbol{\gamma}^T)}$ $\boldsymbol{\gamma} = \{\gamma_j; j = 1, ..., N\}$ (Dirichlet distribution)

3) $\boxed{\boldsymbol{\theta}_i \mid \boldsymbol{\gamma}, \mathbf{C}} \sim \mathcal{D}(\mathbf{C}_i^T + \boldsymbol{\gamma}^T)$ $\boldsymbol{\gamma} = \{\gamma_j; j = 1, ..., N\}$

$$E_P(\boldsymbol{\theta}_{ij}) = \hat{\boldsymbol{\theta}}_{ij} = \frac{c_{ij} + \gamma_j}{\underbrace{\sum_{j=1}^{J} c_{ij}}_{n_i} + \underbrace{\sum_{j=1}^{J} \gamma_j}_{\lambda}} = \underbrace{\boxed{\frac{s_i}{s_i + \lambda}}}_{\alpha_i} p_{ij} + \underbrace{\boxed{\frac{\lambda}{s_i + \lambda}}}_{1-\alpha_i} \underbrace{\frac{\gamma_j}{\lambda}}_{\pi_j}$$

i^{th} row of new **G** $\boxed{\mathbf{G}_i^{*T} = \alpha_i \mathbf{p}_i^T + (1-\alpha_i)\boldsymbol{\pi}^T}$

R&R/Bayesian reinterpretation of the G matrix

i^{th} row of new $\mathbf{G}^*$ $\boxed{\mathbf{G}_i^{*T} = \alpha_i \mathbf{p}_i^T + (1-\alpha_i)\boldsymbol{\pi}^T}$

1)Pretermined values assigned to

λ : tantamount to no of citations given by a journal

$\lambda = k\tilde{s}$ where $\tilde{s}$ = harmonic mean of s_i; $k = 0.15 - 0.25$

similar approach as for historical data

$$\boldsymbol{\pi} = (\pi_j)_{1 \le j \le N}$$

$\pi_j = c_{+j} / c_{++}$ = Prop of citations to j by all journals

R&R/Bayesian reinterpretation of the G matrix

i^{th} row of new $\mathbf{G}^*$ $\boxed{\mathbf{G}_i^{*T} = \alpha_i \mathbf{p}_i^T + (1-\alpha_i)\boldsymbol{\pi}^T}$

2) Empirical– Bayes via eg BLUP

$\lambda = \rho^{-1} - 1$ where ρ intra-class correlation

ANOVA estimators in Landis & Koch,1977

$\boxed{\hat{\pi}_j = \left(\sum_{i=1}^N \alpha_i p_{ij}\right) / \left(\sum_{i=1}^N \alpha_i\right)}$ weighted mean of row transition matrices

$\rho \to 0 \Rightarrow \alpha_i \to 0 \; g_{ij} \to \pi_j = c_{+j} / c_{++}$ for any $i \Rightarrow \pi_j / a_j \propto IF_j$

$\rho \to 1$ and/or s_i large $\Rightarrow \alpha_i = 1$, $\mathbf{G}_i^{*T} = \mathbf{p}_i^T$ (no teleportation) Pinski-Narin

More generally

$$\mathbf{G}_i^{*T} = \mathbf{p}_i^T \mathbf{A}_i^T + \boldsymbol{\pi}^T (\mathbf{I} - \mathbf{A}_i)^T \quad \text{with } \mathbf{A}_i = s_i \left(s_i I_{N-1} + \mathbf{W}_0 \mathbf{B}_0^{-1}\right)^{-1}$$

$\mathbf{B}_0, \mathbf{W}_0$ Between,Within components of var-covar among citing items

Quaas & VanVleck, 1980; Foulley et al, 1987

3) Hierarchical Bayes

$\boldsymbol{\pi}^T \sim \mathcal{D}(c^T) \; \lambda \sim \mathcal{G}(a,b)$

R&R/P matrix estimated via g-logit model

4) P estimated by alternative models: g-logit model with random effects

i^{th} row of $\mathbf{C}$ is $\mathbf{C}_i^T = \left(c_{i1}, c_{i1}, \ldots, c_{ij}, \ldots, c_{iN}\right)$

1) $\boxed{\mathbf{C}_i^T \sim \mathcal{M}\left(s_i, \boldsymbol{\theta}_i^T\right)}$ $\boldsymbol{\theta}_i = \left\{\theta_{ij}; j = 1, \ldots, N\right\}$ where $s_i = \sum_{j=1}^{N} c_{ij}$

s_i = No of citations given by i to all journals (outlinks from i)

2) $\theta_{ij} = \dfrac{\exp(\beta_j + u_{ij})}{1 + \sum_{j=1}^{N-1} \exp(\beta_j + u_{ij})}$ $j = 1, \ldots, N-1$

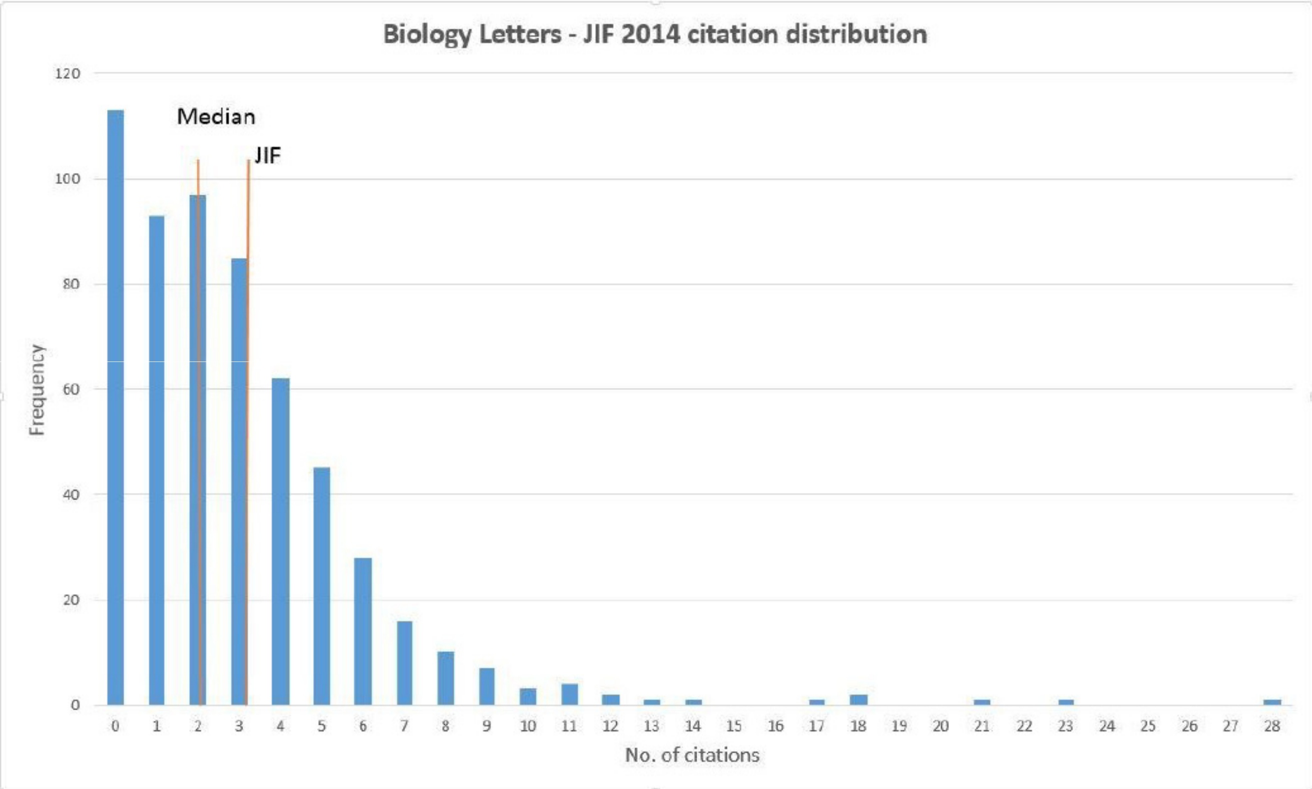
β_j : fixed effect

$\mathbf{u}_i = \left(u_{i1}, \ldots, u_{ij}, \ldots, u_{iJ-1}\right)$ $\mathbf{u}_i \sim \mathcal{N}\left(0, \mathbf{D}\right)$

$\mathbf{D}$ saturated, diagonal (homogeneous or not)

3) $\mathbf{D} \sim \mathcal{IW}(,)$ or $\left[\mathbf{D}\right]_{jj} \sim \mathcal{IG}(,)$

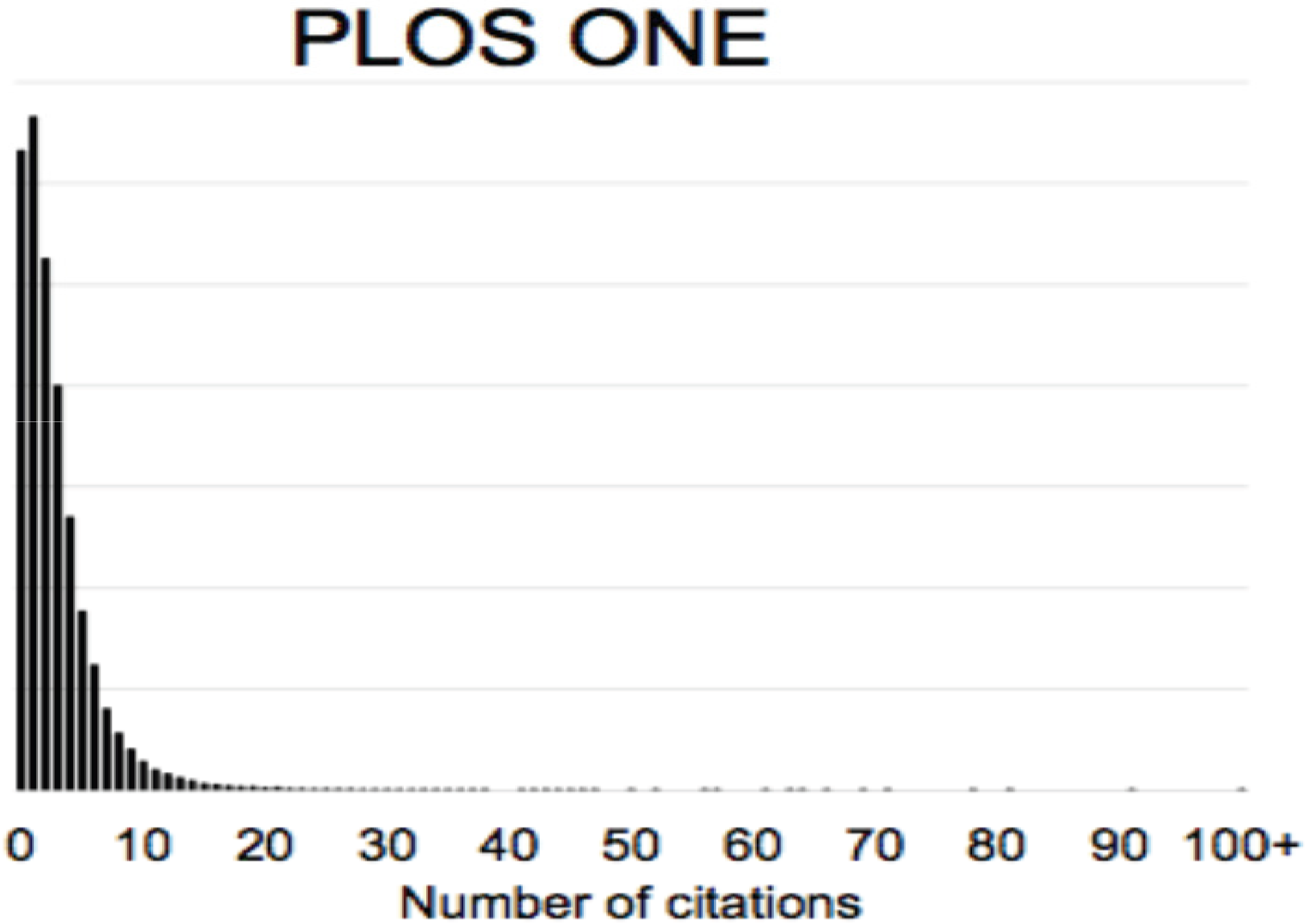
Citation distribution



0	113	19.72
1	93	35.95
2	97	52.88
3	85	67.71
4	62	78.53
5	45	86.38
6	28	91.27
7	16	94.07
8	10	95.81
9	7	97.03
10	3	97.56
11	4	98.25
12	2	98.6
13	1	98.78
14	1	98.95
15	0	98.95
16	0	
17	1	
18	2	
19	0	
20	0	
21	1	
22	0	
23	1	
24	0	
25	0	
26	0	
27	0	
28	1	

Larivière et al (2016) Citations (573) received in 2014 by Biology Letters of citable papers published in 2002-2013

Citation distribution



R&R/Poisson QS loglinear, Bradley-Terry & Stigler models

A) Poisson QS Loglinear model : Caussinus (1965) Goodman (2002)

1) $\boxed{C_{ij} \mid \mu_{ij} \sim_{id} \mathcal{P}\left(\mu_{ij}\right)}$ i citing in rows & j cited in columns

2) $\ln\left(\mu_{ij}\right) = \mu + \beta_i^{(1)} + \beta_j^{(2)} + \beta_{ij}^{(12)}$

3) $\beta_{ij}^{(12)} = \beta_{ji}^{(12)}$ for any $i \neq j = 1, 2, ..., N$

B)Bradley-Terry model for pairwise comparisons

$C_{ij} \mid N_{ij} \sim B\left(N_{ij}, \pi_{ij}\right)$ where $N_{ij} = C_{ij} + C_{ji}$

$\pi_{ij} = \mu_{ij} / \left(\mu_{ij} + \mu_{ji}\right)$

$$\ln \frac{\pi_{ij}}{\pi_{ji}} = \ln \frac{\mu_{ij}}{\mu_{ji}} = \left(\beta_j^{(2)} - \beta_j^{(1)}\right) - \left(\beta_i^{(2)} - \beta_i^{(1)}\right)$$

$\text{logit}\left(\pi_{ij}\right) = \theta_j - \theta_i;\ \theta_i = \beta_i^{(2)} - \beta_i^{(1)}$ Cited -Citing Effect /Reference

Stigler (1994) "Export Score" Varin et al (2015)

$\mathrm{N}_P = \left(N-1\right)\left(N+2\right)/2$; $\mathrm{DL}_R = \left(N-1\right)\left(N-2\right)/2$

R&R/Relationship between BT and AHP-ANR

Analytic Hierarchy and Network Process based on the Pairwise Comparison Matrix (PCM)

$$a_{ij} = \frac{1}{a_{ji}} = \frac{c_{ij}}{c_{ji}} \quad a_{ii}=1; \text{ AHP, } a_{ij} \approx w_j / w_i$$

Several methods to get w_i (Saaty, 1977)

Transition matrix: $[\mathbf{P}]_{ij} = a_{ij} / a_{i+}$

Solve for $\boxed{\mathbf{r}^T\mathbf{P} = \mathbf{r}^T}$ (Choo & Wedley, 2009)

Remember in BT: $\ln \mu_{ij} / \mu_{ji} = \theta_j - \theta_i$

If a_{ij} replaced by μ_{ij} / μ_{ji}, then

$$\boxed{r_i = \exp(\theta_i) / \left[\sum_{k=1}^{N} \exp(\theta_k)\right]}$$

Analysis of the 47x47 Cross Citation matrix

used in Varin et al without self citations

Based on ANOVA Landis & Koch (1977)

$\hat{\rho} = 0.0267 \qquad \hat{\lambda} = 36.381$

α_i varying from 0.51 to 0.97

$\bar{\alpha} = 0.841 \quad \alpha^{(med)} = 0.866$

$s^{(\mathrm{H})} = 173.82 \quad k = 0.209$

R&R/Scores of the 47 journals

NO	JOURNAL	EXS	IW	AI	AI-BLUP	IPP	SJR
1	AMS	1.12	1.118	0.725	0.758	0.682	0.678
2	AISM	1.028	0.977	0.664	0.671	0.641	0.509
3	AOS	2.967	3.012	3.498	3.5	3.663	3.826
4	ANZS	0.793	0.762	0.469	0.452	0.458	0.422
5	BERN	1.497	1.426	1.178	1.215	1.225	0.895
6	BIOJ	0.501	0.471	0.531	0.484	0.474	0.762
7	BCS	1.749	1.614	1.879	1.835	1.778	2.519
8	BKA	2.716	2.674	3.185	3.14	3.276	3.035
9	BIOST	1.444	1.36	0.472	0.452	0.465	0.594
10	CJS	1.016	1.003	1.415	1.375	1.397	1.249
11	CSSC	0.212	0.19	0.212	0.222	0.189	0.229
12	CSTM	0.356	0.291	0.349	0.371	0.299	0.304
13	CMPST	0.395	0.533	0.382	0.37	0.39	0.288
14	CSDA	0.443	0.453	0.583	0.596	0.543	1.054
15	EES	0.463	0.557	0.043	0.046	0.044	0.042
16	ENVR	0.672	0.727	0.445	0.438	0.426	0.401
17	ISR	1.039	1.178	0.668	0.812	0.668	0.501
18	JABES	0.635	0.615	0.189	0.191	0.179	0.183
19	JASA	2.645	2.605	3.792	3.804	3.887	4.113
20	JAS	0.182	0.173	0.173	0.174	0.155	0.177
21	JBS	0.325	0.273	0.092	0.084	0.083	0.165
22	JCGS	1.427	1.481	1.826	1.75	1.835	1.645
23	JMA	0.476	0.52	0.907	0.938	0.911	0.93
24	JNS	0.442	0.454	0.173	0.184	0.176	0.141
25	JRSS-A	1.502	1.67	0.96	0.926	0.951	1.292
26	JRSS-B	6.062	6.092	6.107	6.086	6.295	5.675
27	JRSS-C	0.817	0.808	0.463	0.477	0.461	0.421
28	JSCS	0.299	0.35	0.321	0.329	0.302	0.282
29	JSPI	0.539	0.56	0.546	0.57	0.533	0.607
30	JSS	0.335	0.403	0.053	0.059	0.053	0.053
31	JTSA	1.085	0.925	0.436	0.49	0.403	0.39
32	LTA	0.827	0.75	1.105	1.099	1.06	1.127
33	MTKA	0.623	0.535	0.275	0.282	0.253	0.252
34	SJS	1.45	1.505	1.811	1.827	1.829	1.426
35	STATAJ	0.767	0.759	0.228	0.24	0.213	0.464
36	STCMP	0.776	0.858	0.596	0.596	0.636	0.497
37	STATS	0.389	0.481	0.432	0.402	0.413	0.389
38	STMED	0.799	0.749	0.743	0.771	0.648	1.562
39	SMMR	0.527	0.432	0.078	0.08	0.069	0.109
40	STMOD	0.601	0.546	0.778	0.708	0.718	0.767
41	STNEE	0.674	0.69	0.414	0.427	0.413	0.32
42	STPAP	0.194	0.18	0.172	0.174	0.147	0.163
43	SPL	0.681	0.667	0.296	0.315	0.286	0.291
44	STSCI	0.839	0.961	4.415	4.458	4.5	3.567
45	STSIN	0.999	0.917	0.396	0.4	0.408	0.347
46	TECH	1.275	1.145	1.054	1.001	0.994	1.066
47	TEST	0.371	0.524	1.449	1.396	1.546	1.247

R&R/Applications to statistical journals

Rankings of 47 Statistical Journals according to several criteria

NO	EXS	AI	AI	AI-BLUP	IPP	SJR
1	JRSS-B	JRSS-B	JRSS-B	JRSS-B	JRSS-B	JRSS-B
2	AOS	AOS	STSCI	STSCI	STSCI	JASA
3	BKA	BKA	JASA	JASA	JASA	AOS
4	JASA	JASA	AOS	AOS	AOS	STSCI
5	BCS	JRSS-A	BKA	BKA	BKA	BKA
6	JRSS-A	BCS	BCS	BCS	JCGS	BCS
7	BERN	SJS	JCGS	SJS	SJS	JCGS
8	SJS	JCGS	SJS	JCGS	BCS	STMED
9	BIOST	BERN	TEST	TEST	TEST	SJS
10	JCGS	BIOST	CJS	CJS	CJS	JRSS-A
11	TECH	ISR	BERN	BERN	BERN	CJS
12	AMS	TECH	LTA	LTA	LTA	TEST
13	JTSA	AMS	TECH	TECH	TECH	LTA
14	ISR	CJS	JRSS-A	JMA	JRSS-A	TECH
15	AISM	AISM	JMA	JRSS-A	JMA	CSDA
16	CJS	STSCI	STMOD	ISR	STMOD	JMA
17	STSIN	JTSA	STMED	STMED	AMS	BERN
18	STSCI	STSIN	AMS	AMS	ISR	STMOD
19	LTA	STCMP	ISR	STMOD	STMED	BIOJ
20	JRSS-C	JRSS-C	AISM	AISM	AISM	AMS
21	STMED	ANZS	STCMP	STCMP	STCMP	JSPI
22	ANZS	STATAJ	CSDA	CSDA	CSDA	BIOST
23	STCMP	LTA	JSPI	JSPI	JSPI	AISM
24	STATAJ	STMED	BIOJ	JTSA	BIOJ	ISR

EXS : Export Score; IW: Influence Weight Pinski & Narin
AI : Article Influence, AI-BLUP : same but based on BLUP matrix
IPP : Influence per publication : Pinski & Narin (1976)
SJR: SCImago Journal Rank (per article normalization)
Citations received in2010 JCR from papers published in 2001-2010
(Varin et al, 2016)

R&R/Applications to statistical journals

Rankings of 47 Statistical Journals (continued)

NO	EXS	AI	AI	AI-BLUP	IPP	SJR
25	SPL	ENUR	BIOST	BIOJ	BIOST	STCMP
26	STNEE	STNEE	ANZS	JRSS-C	JRSS-C	STATAJ
27	ENUR	SPL	JRSS-C	BIOST	ANZS	ANZS
28	JABES	JABES	ENUR	ANZS	ENUR	JRSS-C
29	MTKA	JSPI	JTSA	ENUR	STNEE	ENUR
30	STMOD	EES	STATS	STNEE	STATS	JTSA
31	JSPI	STMOD	STNEE	STATS	STSIN	STATS
32	SMMR	MTKA	STSIN	STSIN	JTSA	STSIN
33	BIOJ	CMPST	CMPST	CSTM	CMPST	STNEE
34	JMA	TEST	CSTM	CMPST	JSCS	CSTM
35	EES	JMA	JSCS	JSCS	CSTM	SPL
36	CSDA	STATS	SPL	SPL	SPL	CMPST
37	JNS	BIOJ	MTKA	MTKA	MTKA	JSCS
38	CMPST	JNS	STATAJ	STATAJ	STATAJ	MTKA
39	STATS	CSDA	CSSC	CSSC	CSSC	CSSC
40	TEST	SMMR	JABES	JABES	JABES	JABES
41	CSTM	JSS	JNS	JNS	JNS	JAS
42	JSS	JSCS	JAS	STPAP	JAS	JBS
43	JBS	CSTM	STPAP	JAS	STPAP	STPAP
44	JSCS	JBS	JBS	JBS	JBS	JNS
45	CSSC	CSSC	SMMR	SMMR	SMMR	SMMR
46	STPAP	STPAP	JSS	JSS	JSS	JSS
47	JAS	JAS	EES	EES	EES	EES

R&R/Applications to statistical journals

Correlations among 6 different criteria of journal ratings

	EXS	IW	AI	AI-BLUP	IPP	SJR
EXS	1	0.996	0.845	0.844	0.847	0.857
IW	0.996	1	0.857	0.856	0.859	0.863
AI	0.845	0.857	1	0.999	0.999	0.978
AI-BLUP	0.844	0.856	0.999	1	0.999	0.977
IPP	0.847	0.859	0.999	0.999	1	0.975
SJR	0.857	0.863	0.978	0.977	0.975	1

EXS & IW : influence scores per citation

AI, AI-BLUP, IPP, SJR: influence scores per article

R&R/Discussion-Conclusion

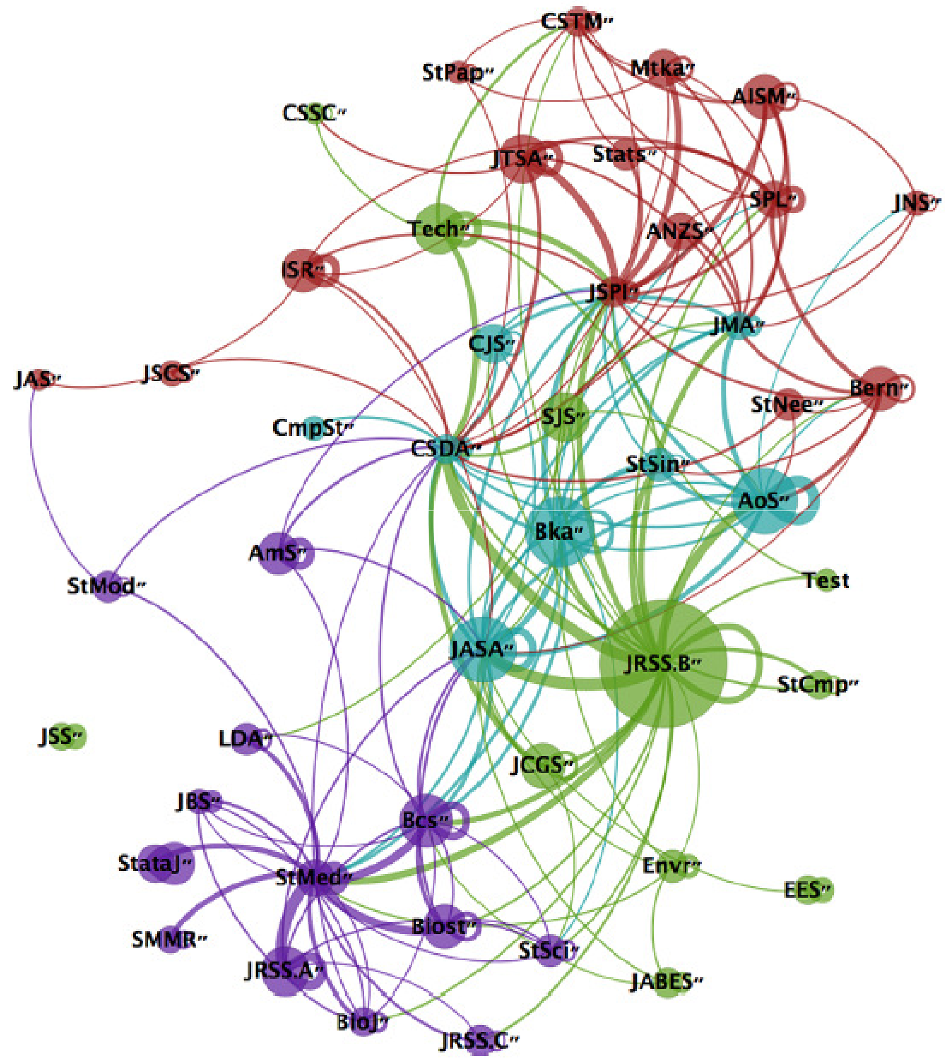
Model-based approaches better suited to Journal R&R than raw indicators

- -Hypotheses made explicit
- -Uncertainty available

Network approaches are attractive due to simplicity & computability

Several extensions made or in progress

- -Grouping Lasso (Varin et al, 2015)
- -Including Dynamics
- -Goodman's Row-Column models (Grah, 2016)
- -Stochastic Block Models (Wyse & White, 2016 in Varin et al)
- -Clustering with Modularity Classes (Arbel & Robert in Varin et al)



(Arbel & Robert, 2015)

R&R/Discussion-Conclusion

Practical side

- -Multidimensional rating (Prestige vs Popularity)
- -Solutions to avoid manipulation (Archambault & Larivière)
  - -Algorithm kept secret
  - -Monitoring journal behaviour (including sanctions)
- -Avoid misuse of indicators
  - -Indicators specific to different levels of evaluation
  - -Field dependency to take into account
  - -Poor correlation between JIF and citation rate of individuals

References

Agresti, A. (2013) Categorical Data Analysis, 3rd edn. New York: Wiley

Archambault, E. and Lariviere, V. (2009) History of the journal impact factor: contingencies and consequences. ` Scientometrics, 79, 635–649.

Bergstrom, C. (2007) Eigenfactor: measuring the value and the prestige of scholarly journals. Coll. Res. Lib. News

Brin, S. and Page, L. (1998) The anatomy of a large-scale hypertextual web search engine. Comput. Netwrks ISDN Syst., 30, 107–117.

Caussinus H (1966) Contribution à l'analyse statistique des tableaux de correlation. *Annales de la Faculté des Sciences de Toulouse*, 29, 77-182.

Christensen R (1990) Log-Linear Models. Springer-Verlag, 408 pages

Firth, D. (2009) Bradley–Terry models in R. J. Statist. Softwr., 12, 1–12.

Foulley J.L., Gianola D., Im S., 1987. Genetic evaluation for discrete polygenic traits in animal breeding. In Gianola D., Hammond K. (eds.). Statistical Methods for Genetics Improvement of Livestock, 1990, 361-409. Springer Verlag, Heidelberg.

Franceschet, M. (2012) The large-scale structure of journal citation networks. J. Am. Soc. Inform. Sci. Technol., 63, 837–842.

Garfield, E. (1955) Citation indices for Science. Science, 122, 108–111.

Grah S (2016) Ranking and Rating of Scientific Journals. M1 Master Report, Paris Sud Orsay, 54p

Garfield, E. (1972) Citation analysis as a tool in journal evaluation. Science, 178, 471–479.

Goodman L A (2002) Contributions to the statistical analysis of contingency tables: Notes on quasi-symmetry, quasi-independence, log-linear models, log bilinear models, and correspondence analysis models. *Annales de la Faculté des Sciences de Toulouse*, 6ème série, 11,(4), 525-540.

Palacios-Huerta, I. and Volij, O (2004) The measurement of intellectual influence. Econometrica, 72, 963–977

Quaas RL, Van Vleck (1980) Categorical trait sire evaluation by BLUP of future category frequencies. Biometrics, 36, 117-122

Seglen, P. O. (1997) Why the impact factor of journals should not be used for evaluating research. Br. Med. J., 314, 498–502.

Stigler SM (1994) Citations Patterns in the Journals of Statistics and Probability. *Statistical Science*, 9, 94-108

Varin C, Cattelan M, Firth D (2016) Statistical modelling of citation exchange between statistics journals. JR Statistical Society A,179,1-63

Waltman L, van Eck NJ (2010) The relation between Eigenfactor, audience factor and influence weight. J of the American Society for Information Science & Technolgy, 61, 1476-1486.

Zitt, M. and Small, H. (2008) Modifying the journal impact factor by fractional citation weighting: the audience factor. J. Am. Soc. Inform. Sci. Technol., 59, 1856–1860.