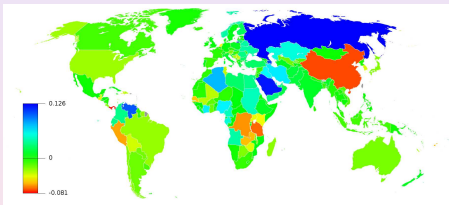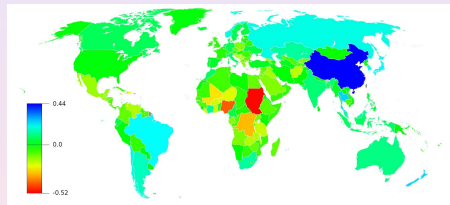# Reduced Google matrix: applications to directed networks

Dima Shepelyansky (CNRS, Toulouse)
www.quantware.ups-tlse.fr/dima

with L.Ermann (CNEA TANDAR), K.Frahm (LPT), K.Jaffres-Runser (N7 Tlse), J.Lages (U Besancon); thanks to UN COMTRADE
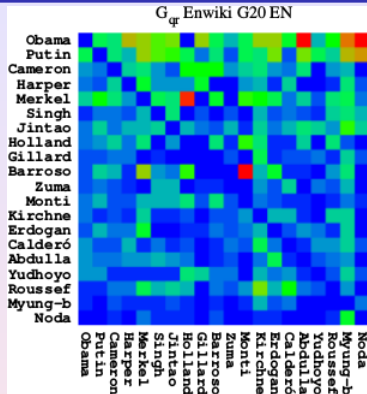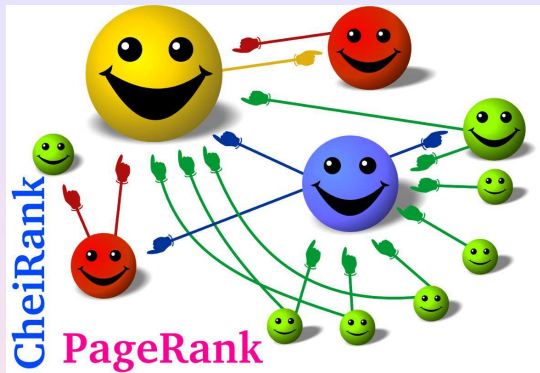

* Markov (1906) $\rightarrow$ Brin and Page (1998)
* reduced Google matrix of directed networks (brief introduction)
* Applications: multiproduct world trade network (UN COMTRADE), Wikipedia Ranking of World Universitis (WRWU), world terror networks ...

(1906) Markov vs Wigner (1955)

1945: Nuclear physics → Wigner (1955)→ Random Matrix Theory
1991: WWW, small world social networks → Markov (1906) → Google matrix

Despite the importance of large-scale search engines on the web, very little academic research has been done on them.

S.Brin and L.Page, Comp. Networks ISDN Systems **30**, 107 (1998)

Google matrix construction rules

Markov chains (1906) and Directed networks
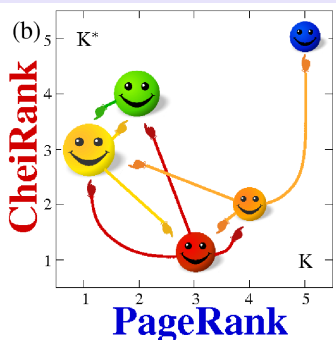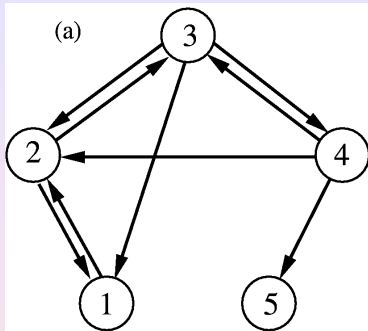
For a directed network with N nodes the adjacency matrix **A** is defined as $A_{ij} = 1$ if there is a link from node j to node i and $A_{ij} = 0$ otherwise. The weighted adjacency matrix is

$$S_{ij} = A_{ij} / \sum_k A_{kj}$$

In addition the elements of columns with only zeros elements are replaced by $1/N$.

Google matrix construction rules

Google Matrix and Computation of PageRank

$P = SP \Rightarrow P$= stationary vector of **S**; can be computed by iteration of **S**.

To remove convergence problems:

- Replace columns of 0 (dangling nodes) by $\frac{1}{N}$:

$$S = \begin{pmatrix} 0 & 1/2 & 1/3 & 0 & 1/5 \\ 1 & 0 & 1/3 & 1/3 & 1/5 \\ 0 & 1/2 & 0 & 1/3 & 1/5 \\ 0 & 0 & 1/3 & 0 & 1/5 \\ 0 & 0 & 0 & 1/3 & 1/5 \end{pmatrix} \qquad S^* = \begin{pmatrix} 0 & 1/3 & 0 & 0 & 0 \\ 1/2 & 0 & 1/2 & 0 & 0 \\ 1/2 & 1/3 & 0 & 1 & 0 \\ 0 & 1/3 & 1/2 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

- To remove degeneracies of $\lambda = 1$, replace **S** by **Google matrix**

  $\mathbf{G} = \alpha\mathbf{S} + (1 - \alpha)\frac{\mathbf{E}}{N}$; $\quad GP = \lambda P$ => Perron-Frobenius operator

- α models a random surfer with a random jump after approximately 6 clicks (usually $\alpha = 0.85$); PageRank vector => P at $\lambda = 1$ ($\sum_j P_j = 1$).

- CheiRank vector P^*: $G^* = \alpha\mathbf{S}^* + (1 - \alpha)\frac{\mathbf{E}}{N}$, $G^*P^* = P^*$
  (**S*** with inverted link directions)
  Chepelianskii arXiv:1003.5455 (2010) ...

Reduced Google matrix

A selected network of interest with $N_r < N$ nodes called reduced network. Block structure of G matrix:

$$G = \begin{pmatrix} G_{rr} & G_{rs} \\ G_{sr} & G_{ss} \end{pmatrix}$$

with s index for scattering network $N_s = N - N_r$.
Reduced G_R matrix

$$G_{\mathrm{R}} P_r = P_r \quad , \quad G_{\mathrm{R}} = G_{rr} + G_{rs}(\mathbf{1} - G_{ss})^{-1} G_{sr} = G_{pr} + G_{rr} + G_{qr}$$

Useful expansion

$$(\mathbf{1} - G_{ss})^{-1} = \mathcal{P}_c \frac{1}{1 - \lambda_c} + \mathcal{Q}_c \sum_{l=0}^{\infty} \bar{G}_{ss}^l$$

with projector $\mathcal{P}_c = \psi_R \psi_L^T$ on eigenstate of maximal eigenvalue λ_c of G_{ss}, the complementary projector $\mathcal{Q}_c = \mathbf{1} - \mathcal{P}_c$ and $\bar{G}_{ss} = \mathcal{Q}_c G_{ss} \mathcal{Q}_c$.
K.Frahm, DS arxiv:1602.02394 (2016)

Top historical figures of 24 Wikipedia editions

2DRanking of Wikipedia articles; top 100 historical figures;
comparison with historical studies of M.Hart (37 and 43 percent overlap)
35 centuries and all countries by birth place; 17 millions wiki-articles
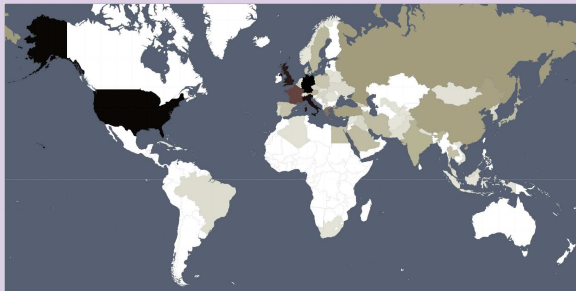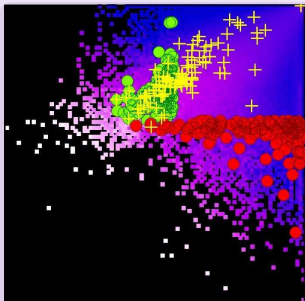
A.Zhirov, O.Zhirov, DLS EPJB (2010); Y.-H.Eom, K.M.Frahm, A.Benczur, DLS EPJB
(2013); Y.-H.Eom, DLS PLoS ONE (2013), Y.-H.Eom,P.Aragon, D.Laniado,
A.Kaltenbrunner, S.Vigna, DLS arXiv2014 - PLoS ONE (2015)

Top historical figures of 24 Wikipedia editions

Top global PageRank historical figures: Carl Linnaeus, Jesus, Aristotle ...
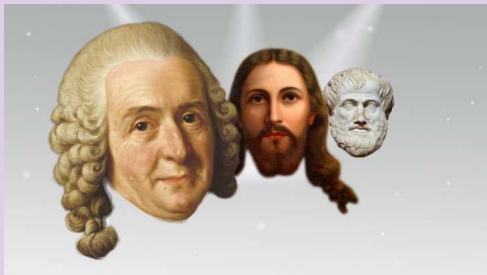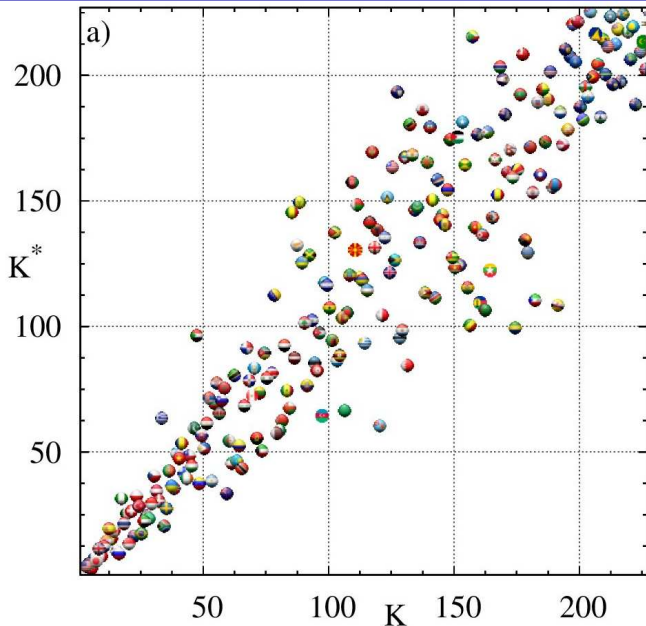

Media highlights: The Guardian, The Independent, The Washington Post, France24, EC CORDIS
==>Uppsala Universitet: "Carl Linnaeus ranked most influential person of all time" ... (about 20 countries)
Competitors: MIT Pantheon project http://pantheon.media.mit.edu (2014); Stony-Brook NY http://www.whoisbigger.com/ (2014)
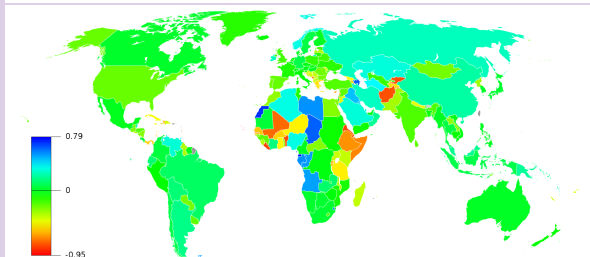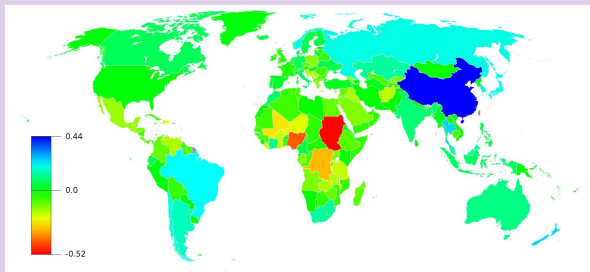
Ranking of World Trade (2008 all commodities)

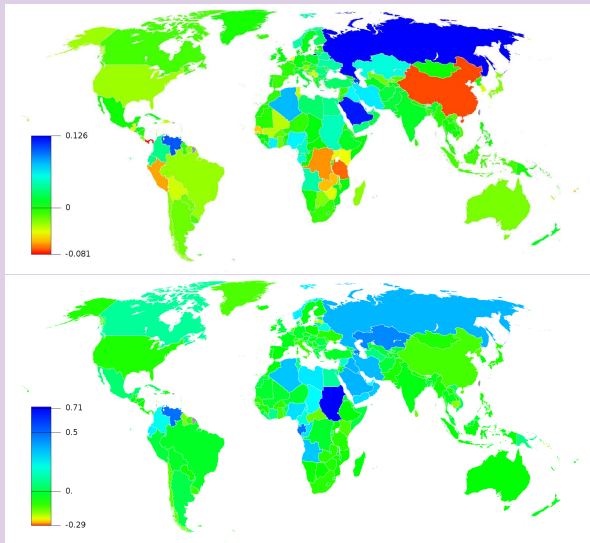a)

CheiRank-PageRank balance (2008)

$B_c = (P_c{}^* - P_c)/(P_c{}^* + P_c)$ (top - CheiRank-PageRank; bottom -Export-Import volume; multiproduct world trade $N_c = 227\,countries, N_p = 61\,products, N = 13847 \Longrightarrow \text{UN COMTRADE})$
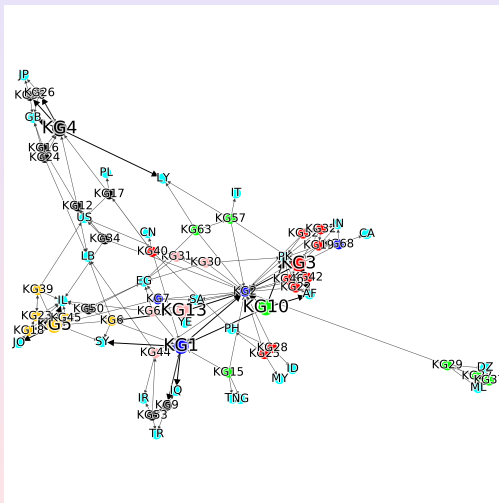
Sensitivity to petrolem price (2008)

$B_c = (P_c^* - P_c)/(P_c^* + P_c)$, color => $dB_c/d\delta_{petroleum})$

# Analysis of world terror networks (ENWiki2017)

reduced Google matrix for 64 world countries and 95 terrorist groups (from ENWiki2017 with 5.4 millions articles)

Friendship network structure from $G_{qr} + G_{rr}$ with the top terrorist groups (marked by their respective colors) and countries (marked by cyan color); it is shown with 2 friends for top terrorist groups of each category and top friend 2 countries for each group. El Zant, Frahm, Jaffres-Runser, DS (EPJB 2018)

# Analysis of world terror networks (ENWiki2017)

reduced Google matrix for 64 world countries and 95 terrorist groups (from ENWiki2017 with 5.4 millions articles)
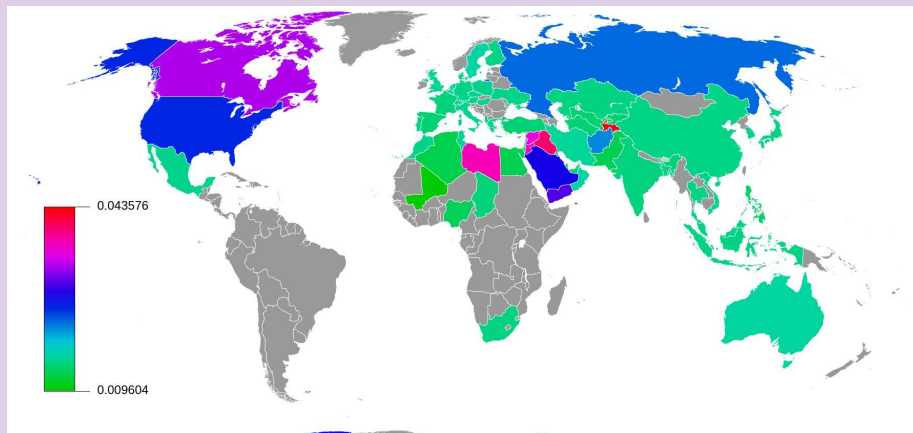
PageRang sensitivity S to weight link variation ISIS - country
El Zant, Frahm, Jaffres-Runser, DS (EPJB 2018)

Shaghai vs. Wikipedia World University Ranking

Top 100 (Shanghai ARWU2017 - left, wiki WRWU2017 - right), about 20 millions wiki-articles of 24 language editions
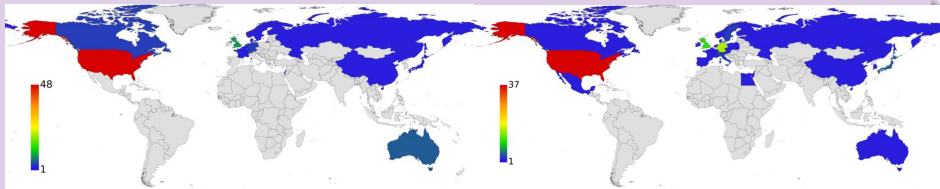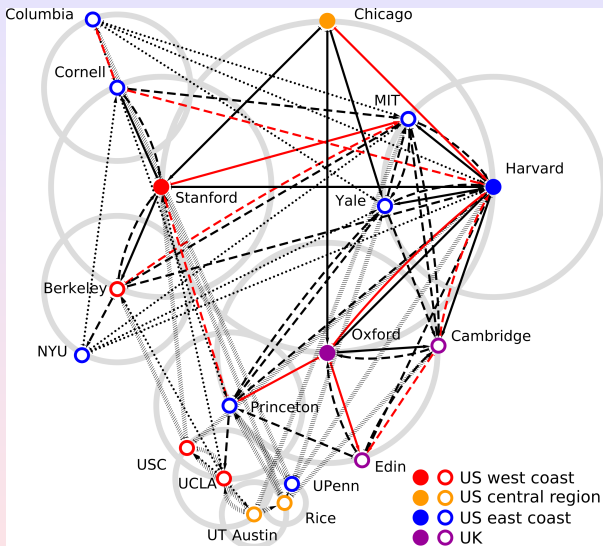

ARWU2017: 1.Harvard, 2.Stanford, 3.Cambridge, 4.MIT, 5.Berkeley (...74.Rice; USA: 37)
WRWU2017: 1.Oxford, 2.Cambridge, 3.Harvard, 4.Columbia, 5.Yale (...357.Rice; USA: 56)
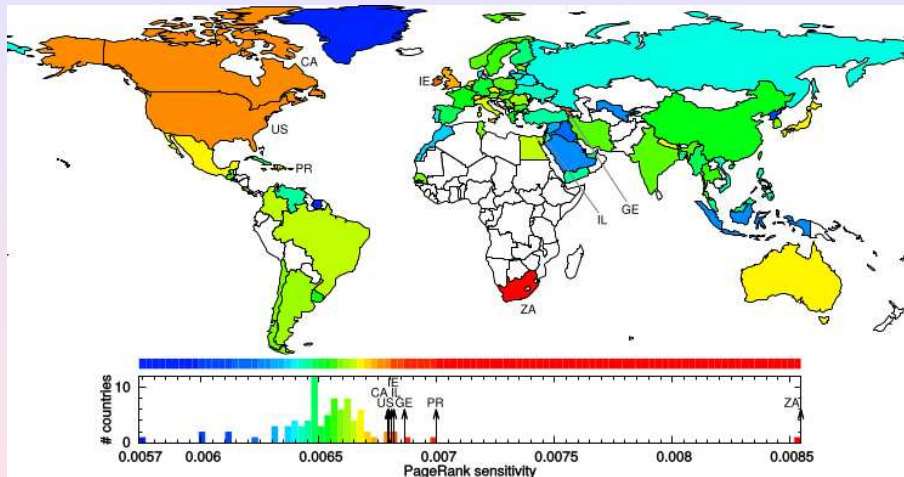60percent overlap for top 100
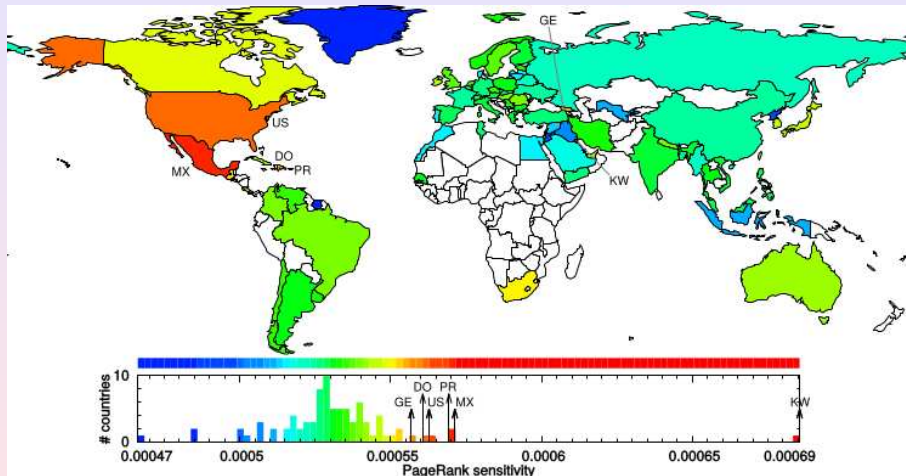Coquide, Lages, DS (in progress 2018) + EPJB **89**, 69 (2016)

G-reduced network of top 20 Univ of EnWiki2017 plus Rice Univ; color marks regional leaders, red links are purely indirect links, 4 level friendship links are shown: black solid, dashed, dotted, slash lines. Coquide, Lages, DS (in progress 2018)

Coquide, Lages, DS (in progress 2018)

Coquide, Lages, DS (in progress 2018)

Further applications of Markov chains and Google matrix ? → Artificial intelligence at NEXT

Google matrix: fundamentals, applications and beyond,
IHES workshop 15-18 Oct 2018 (see www.ihes.fr)