Reduced Google matrix analysis of directed biological networks

José Lages
Équipe de Physique Théorique, Institut UTINAM, CNRS,
Université de Bourgogne Franche-Comté, Besançon

Dima L. Shepelyansky
Laboratoire de Physique Théorique, CNRS,
Université Paul Sabatier, Toulouse

Andrei Zinovyev
Institut Curie, Inserm, PSL Université, Paris

1ère Journée autour des mathématiques et de l'informatique en analyse de données et d'imagerie en oncologie
Institut de Cancérologie de Lorraine
June 12 2018

Projects

- **ApliGoogle** project (2016-2018) funded by MASTODONS CNRS Mission interdisciplinarité
Partners : LPT, CNRS, UPS, Toulouse / UTINAM, CNRS, UBFC, Besançon / I. Curie, Inserm, PSL, Paris / IRIT, CNRS, UPS, Toulouse

- **APEX** project (2017-2020) funded by Région Bourgogne Franche-Comté.
Researchers : J. Lages (PI), G. Rollin, C. Coquidé, D. Viennot, V. Pouthier, P. Joubert, S. Diakité, G. Jolicard

- **GNETWORKS** project (2018-2021) funded by ISITE-UBFC (PIA).
Researchers : J. Lages (PI), G. Rollin, C. Coquidé, D. Viennot, V. Pouthier, P. Joubert, S. Diakité, G. Jolicard

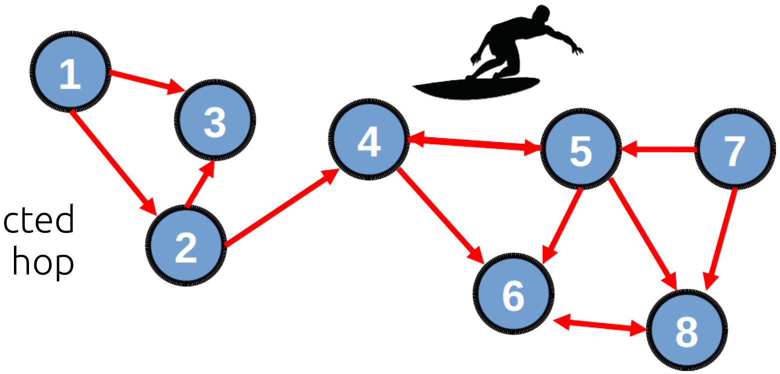
3 projects devoted to the physical analysis of complex networks and the application of Google matrix based analysis to complex systems

Examples of complex systems seen as directed networks

Data	Nodes	Directed links
WWW	Webpages	Hyperlinks
Wikipedia	Articles	Intrawiki citations
Social networks	Members	Acquaintances
World trade	Goods x countries	Importations / exportations
Omics	Proteins	Causal relations / interactions
Linux	Kernel commands	Command successions
DNA	Words of letters A,T,G,C	Word successions
Go game	Plaquettes / patterns	Pattern successions

Non exhaustive list ...

How Google search works



From Markov (1906) to Brin & Page (1998)

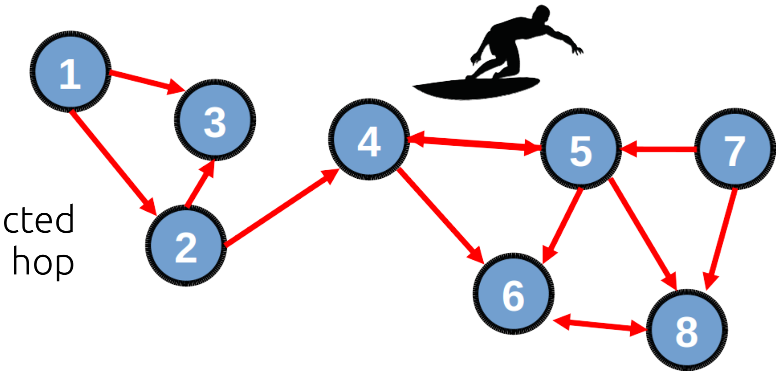
Markovian process : a random surfer probe the structure of a directed network. A each step, the surfer choose randomly an adjacent node to hop and continue its journey.

Adjacency matrix

$$A_{ij} = \begin{cases} 1 \text{ si } j \to i \\ 0 \text{ si } j \not\to 1 \end{cases}$$

$$\mathbf{A} = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 & 1 & 0 \end{pmatrix}$$

How Google search works



From Markov (1906) to Brin & Page (1998)

Markovian process : a random surfer probe the structure of a directed network. A each step, the surfer choose randomly an adjacent node to hop and continue its journey.

Adjacency matrix

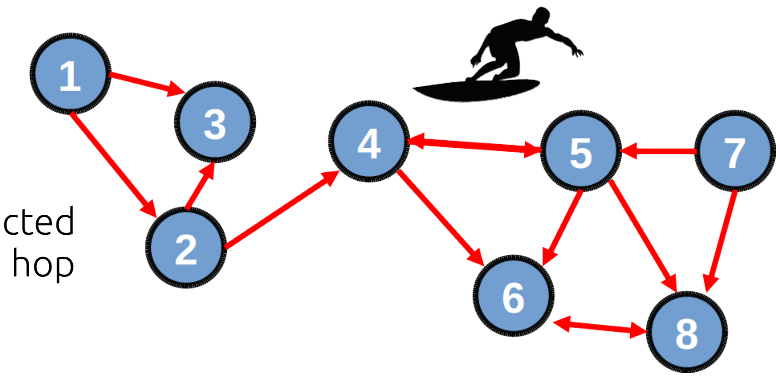
$$A_{ij} = \begin{cases} 1 \text{ si } j \to i \\ 0 \text{ si } j \nrightarrow 1 \end{cases}$$

Stochastic matrix

$$S_{ij} = \begin{cases} A_{ij} / \sum_{k=1}^{N} A_{kj} & \text{si} \sum_{k=1}^{N} A_{kj} \neq 0 \\ 1/N & \text{sinon} \end{cases}$$

$$\mathbf{S} = \begin{pmatrix} 0 & 0 & 1/8 & 0 & 0 & 0 & 0 & 0 \\ 1/2 & 0 & 1/8 & 0 & 0 & 0 & 0 & 0 \\ 1/2 & 1/2 & 1/8 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1/2 & 1/8 & 0 & 1/3 & 0 & 0 & 0 \\ 0 & 0 & 1/8 & 1/2 & 0 & 0 & 1/2 & 0 \\ 0 & 0 & 1/8 & 1/2 & 1/3 & 0 & 0 & 1 \\ 0 & 0 & 1/8 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1/8 & 0 & 1/3 & 1 & 1/2 & 0 \end{pmatrix}$$

How Google search works



From Markov (1906) to Brin & Page (1998)

Markovian process : a random surfer probe the structure of a directed network. A each step, the surfer choose randomly an adjacent node to hop and continue its journey.

Adjacency matrix

$$A_{ij} = \begin{cases} 1 \text{ si } j \to i \\ 0 \text{ si } j \nrightarrow 1 \end{cases}$$

Stochastic matrix

$$S_{ij} = \begin{cases} A_{ij} / \sum_{k=1}^{N} A_{kj} & \text{si } \sum_{k=1}^{N} A_{kj} \neq 0 \\ 1/N & \text{sinon} \end{cases}$$

Google matrix

$$G_{ij} = \alpha S_{ij} + (1-\alpha)/N$$
$$\text{avec } 0.5 < \alpha < 1$$

Perron-Frobenius operator

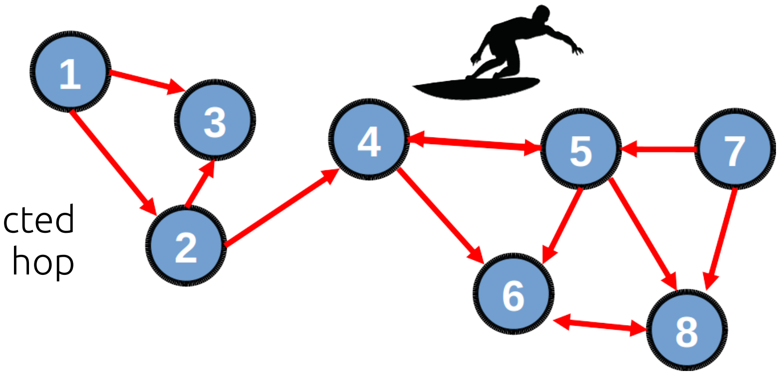
$$\mathbf{G} = \begin{pmatrix} 1/40 & 1/40 & 1/8 & 1/40 & 1/40 & 1/40 & 1/40 & 1/40 \\ 17/40 & 1/40 & 1/8 & 1/40 & 1/40 & 1/40 & 1/40 & 1/40 \\ 17/40 & 17/40 & 1/8 & 1/40 & 1/40 & 1/40 & 1/40 & 1/40 \\ 1/40 & 17/40 & 1/8 & 1/40 & 7/24 & 1/40 & 1/40 & 1/40 \\ 1/40 & 1/40 & 1/8 & 17/40 & 1/40 & 1/40 & 17/40 & 1/40 \\ 1/40 & 1/40 & 1/8 & 17/40 & 7/24 & 1/40 & 1/40 & 33/40 \\ 1/40 & 1/40 & 1/8 & 1/40 & 1/40 & 1/40 & 1/40 & 1/40 \\ 1/40 & 1/40 & 1/8 & 1/40 & 7/24 & 33/40 & 17/40 & 1/40 \end{pmatrix}$$

$\alpha = 0.8$

How Google search works

From Markov (1906) to Brin & Page (1998)

Markovian process : a random surfer probe the structure of a directed network. A each step, the surfer choose randomly an adjacent node to hop and continue its journey.



Adjacency matrix

$$A_{ij} = \begin{cases} 1 \text{ si } j \to i \\ 0 \text{ si } j \nrightarrow 1 \end{cases}$$

Stochastic matrix

$$S_{ij} = \begin{cases} A_{ij} / \sum_{k=1}^{N} A_{kj} & \text{si } \sum_{k=1}^{N} A_{kj} \neq 0 \\ 1/N & \text{sinon} \end{cases}$$

Google matrix

$$G_{ij} = \alpha S_{ij} + (1-\alpha)/N$$
$$\text{avec } 0.5 < \alpha < 1$$

Perron-Frobenius operator

PageRank vector

$$\mathbf{P} = \lim_{n\to\infty} \mathbf{P}^{(n)} = \lim_{n\to\infty} G^n \mathbf{P}^{(0)}$$

$P_i^{(n)}$ is the probability that random surfer arrives at node i at the nth step.

$\mathbf{P}$ is the $\mathbf{G}$ matrix eigenvector associated with eigenvalue 1

$$\mathbf{P} = \mathbf{G}\mathbf{P}$$

$$\mathbf{P} = \begin{pmatrix} 0.031094525687305597 \\ 0.04353233614756617 \\ 0.06094527086606558 \\ 0.06729412361797826 \\ 0.07044998599586171 \\ 0.3518167935609489 \\ 0.03109452568730597 \\ 0.34377243843697143 \end{pmatrix}$$

Distribution $P(K)$

where K is the rank index:

$P(1) = 0.3518167935609489$ 6

$P(2) = 0.34377243843697143$ 8

$P(3) = 0.07044998599586171$ 5

$P(4) = 0.06729412361797826$ 4

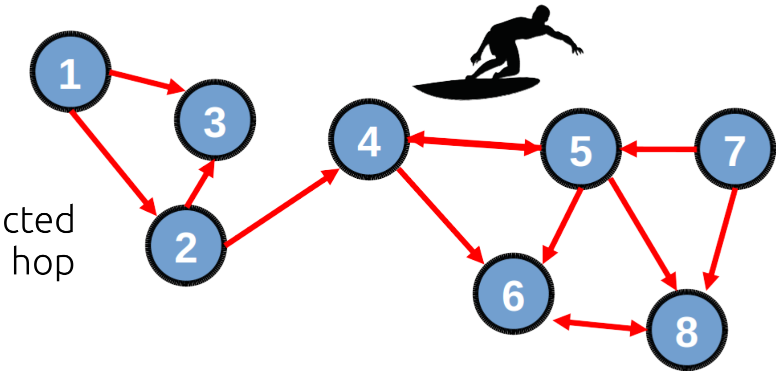
$P(5) = 0.06094527086606558$ 3

$P(6) = 0.04353233614756617$ 2

$P(7) = P(8) = 0.03109452568730597$ 1 7

How Google search works

From Markov (1906) to Brin & Page (1998)



Markovian process : a random surfer probe the structure of a directed network. A each step, the surfer choose randomly an adjacent node to hop and continue its journey.

Adjacency matrix

$$A_{ij} = \begin{cases} 1 \text{ si } j \rightarrow i \\ 0 \text{ si } j \nrightarrow 1 \end{cases}$$

Stochastic matrix

$$S_{ij} = \begin{cases} A_{ij} / \sum_{k=1}^{N} A_{kj} & \text{si } \sum_{k=1}^{N} A_{kj} \neq 0 \\ 1/N & \text{sinon} \end{cases}$$

Google matrix

$$G_{ij} = \alpha S_{ij} + (1-\alpha)/N$$
$$\text{avec } 0.5 < \alpha < 1$$

Perron-Frobenius operator

PageRank vector

$$\mathbf{P} = \lim_{n\to\infty} \mathbf{P}^{(n)} = \lim_{n\to\infty} G^n \mathbf{P}^{(0)}$$

$P_i^{(n)}$ is the probability that random surfer arrives at node i at the nth step.

$\mathbf{P}$ is the $\mathbf{G}$ matrix eigenvector associated with eigenvalue 1

$$\mathbf{P} = \mathbf{GP}$$

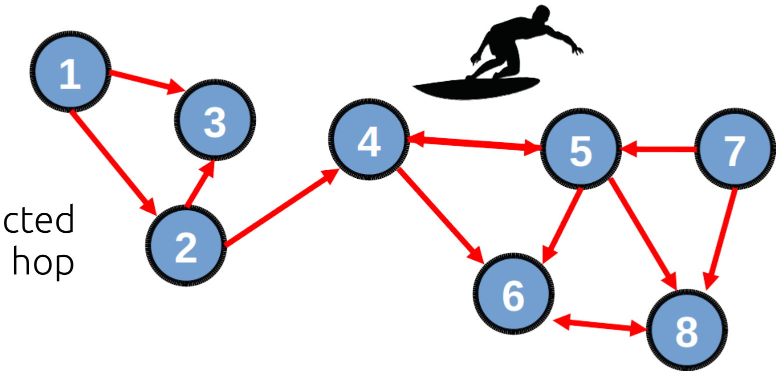
The most important node is the one with the highest probability.
"**Recursive definition**": the more a node is pointed by important nodes, the more it is important.

PageRank measures the influence of a node.

PageRank is at the heart of Google search engine (Brin, Page '98).

How Google search works

From Markov (1906) to Brin & Page (1998)



Markovian process : a random surfer probe the structure of a directed network. A each step, the surfer choose randomly an adjacent node to hop and continue its journey.

Adjacency matrix

$$A_{ij} = \begin{cases} 1 \text{ si } j \to i \\ 0 \text{ si } j \nrightarrow 1 \end{cases}$$

Stochastic matrix

$$S_{ij} = \begin{cases} A_{ij} / \sum_{k=1}^{N} A_{kj} & \text{si } \sum_{k=1}^{N} A_{kj} \neq 0 \\ 1/N & \text{sinon} \end{cases}$$

Google matrix

$$G_{ij} = \alpha S_{ij} + (1-\alpha)/N$$
$$\text{avec } 0.5 < \alpha < 1$$

Perron-Frobenius operator

PageRank vector

$$\mathbf{P} = \lim_{n\to\infty} \mathbf{P}^{(n)} = \lim_{n\to\infty} G^n \mathbf{P}^{(0)}$$

$P_i^{(n)}$ is the probability that random surfer arrives at node i at the nth step.

$\mathbf{P}$ is the $\mathbf{G}$ matrix eigenvector associated with eigenvalue 1

$$\mathbf{P} = \mathbf{G}\mathbf{P}$$

The most important node is the one with the highest probability.
"**Recursive definition**": the more a node is pointed by important nodes, the more it is important.

PageRank measures the influence of a node.

PageRank is at the heart of Google search engine (Brin, Page '98).

CheiRank vector $\mathbf{P}^* = \mathbf{G}^*\mathbf{P}^*$

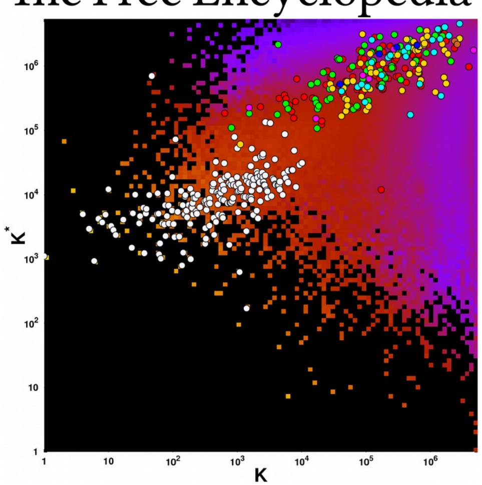
Similar to PageRank of the inverted network. With inverted adjacency matrix elements $A^*_{ij} = A_{ji}$, it is possible to define the stochastic matrix elements $S^*_{ij} \neq S_{ji}$, and the Google matrix elements $G^*_{ij} \neq G_{ji}$ associated to the inverted network (Fogaras '03, Chepelianskii '10).

"**Recursive definition**": the more a node pointed toward important nodes, the more it is important.

CheiRank measures the diffusion/the communication of a node.

Ranking of infectious diseases and countries in 2017 English Wikipedia according to PageRank algorithm. The color code distinguishes type of infectious diseases: **bacterial**, **viral**, **parasitic**, **fongic**, **prionic**, **multiple origin**, and **other origin**.

WIKIPEDIA
The Free Encyclopedia

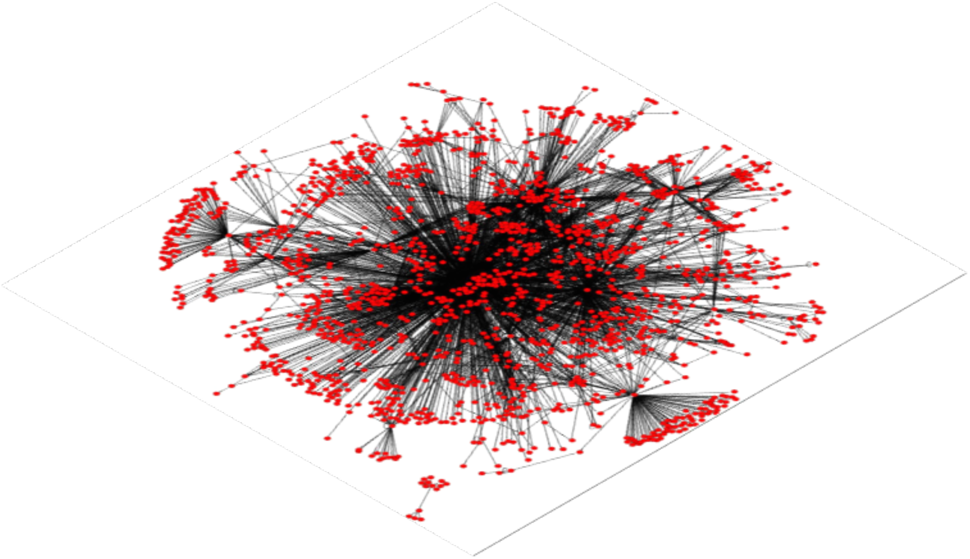


Rank	Disease or country	Rank	Disease or country	Rank	Disease or country	Rank	Disease or country
1	United States	228	Haemophilus influenzae	294	Cysticercosis	360	Bolivian hemorrhagic fever
	…	229	Tetanus	295	Babesiosis	361	Blastomycosis
105	Sudan	230	H. papillomavirus inf.	296	Bacteroides	362	Cutaneous larva migrans
106	Tuberculosis	231	West Nile fever	297	Pneumocystis pneumonia	363	H. metapneumovirus
107	Uganda	232	Schistosomiasis	298	Viral pneumonia	364	Zygomycosis
	…	233	Herpes simplex	299	Cryptococcosis	365	Trichuriasis
114	Somalia	234	Gonorrhea	300	Hepatitis E	366	Granuloma inguinale
115	HIV/AIDS	235	Pertussis	301	Acinetobacter	367	Hymenolepiasis
116	Ivory Coast	236	African trypanosomiasis	302	Chlamydophila pneumoniae	368	Clonorchiasis
	…	237	Rubella	303	Q fever	369	HFRS[g]
128	Fiji	238	Hepatitis A	304	Pinworm inf.	370	Buruli ulcer
129	Malaria	239	Cytomegalovirus	305	Shigellosis	371	LCM[h]
130	Mali	240	Botulism	306	Gas gangrene	372	Yersinia pseudotuberculosis
	…	241	Mumps	307	Bacillus cereus	373	Pediculosis pubis
140	Oman	242	Creutzfeldt–Jakob disease	308	Kuru (disease)	374	BK virus
141	Pneumonia	243	Toxoplasmosis	309	SSPE[b]	375	GSS[i]
142	Smallpox	244	Candidiasis	310	Group A streptococcal inf.	376	Kingella kingae
143	Suriname	245	Chlamydia inf.	311	Roseola	377	H. granulocytic anaplasmosis
	…	246	Rickettsia	312	Meningococcal disease	378	Microsporidiosis
162	Malawi	247	Infectious mononucleosis	313	H. parainfluenza viruses	379	Pneumococcal inf.
163	Cholera	248	Onchocerciasis	314	Burkholderia	380	Opisthorchiasis
164	Togo	249	Scabies	315	Onychomycosis	381	Nocardiosis
	…	250	Brucellosis	316	Aspergillosis	382	Taeniasis
185	San Marino	251	Chagas disease	317	CCHF[c]	383	Bartonellosis
186	Influenza	252	Shingles	318	Relapsing fever	384	Anaplasmosis
187	Saint Lucia	253	Filariasis	319	Ascariasis	385	Tinea capitis
188	Measles	254	Hookworm inf.	320	Strongyloidiasis	386	Colorado tick fever
189	Palau	255	Leishmaniasis	321	Glanders	387	Baylisascaris
190	Typhoid fever	256	Leptospirosis	322	Psittacosis	388	Fasciolopsiasis
191	Marshall Islands	257	Pelvic inflammatory disease	323	Listeriosis	389	Group B streptococcal inf.
192	Equatorial Guinea	258	Norovirus	324	Caliciviridae	390	Pasteurellosis
193	Dominica	259	Cellulitis	325	PML[d]	391	Head lice infestation
194	Guinea-Bissau	260	Trichinosis	326	Rickettsialpox	392	Angiostrongyliasis
195	Syphilis	261	Rotavirus	327	Tinea versicolor	393	Isosporiasis
196	Comoros	262	Hantavirus	328	Campylobacteriosis	394	Argentine hemorrhagic fever
197	Djibouti	263	Legionnaires' disease	329	Naegleriasis	395	Diphyllobothriasis
198	Yellow fever	264	Histoplasmosis	330	Murine typhus	396	Heartland virus
199	Bubonic plague	265	Clostridium difficile inf.	331	Tinea cruris	397	Cyclosporiasis
200	Fed. States of Micronesia	266	Rocky Mountain spotted fever	332	Fusobacterium	398	Carrion's disease
201	Poliomyelitis	267	Enterococcus	333	Rift Valley fever	399	Balantidiasis
202	Tuvalu	268	Bacterial vaginosis	334	Lassa fever	400	Tinea manuum
203	Leprosy	269	Giardiasis	335	Chancroid	401	Sporotrichosis
204	Sepsis	270	Bacterial pneumonia	336	Cat-scratch disease	402	Venezuelan hemorrhagic fever
205	Nauru	271	Amoebiasis	337	Neonatal conjunctivitis	403	Blastocystosis
206	St. Vincent & Grenadines	272	H. respiratory syncytial virus	338	Toxocariasis	404	Tinea barbae
207	Meningitis	273	Athlete's foot	339	Astrovirus	405	Yersiniosis
208	Kiribati	274	Trichomoniasis	340	Fifth disease	406	Tinea nigra
209	Plague (disease)	275	Epidemic typhus	341	Staphylococcal inf.	407	Chromoblastomycosis
210	Saint Kitts and Nevis	276	Hemolytic-uremic syndrome	342	Vibrio parahaemolyticus	408	Dientamoebiasis
211	Typhus	277	Marburg virus	343	Prevotella	409	Brazilian hemorrhagic fever
212	Antigua and Barbuda	278	Trachoma	344	Fatal familial insomnia	410	Gnathostomiasis
213	São Tomé & Príncipe	279	Rhinovirus	345	Anisakis	411	Mycoplasma pneumonia
214	Diphtheria	280	Salmonellosis	346	Ehrlichiosis	412	Capillariasis
215	SARS[a]	281	Coccidioidomycosis	347	VEE[e]	413	White piedra
216	Anthrax	282	Cryptosporidiosis	348	Molluscum contagiosum	414	HME[j]
217	Hepatitis C	283	Myiasis	349	MERS[f]	415	Metagonimiasis
218	Foodborne illness	284	Enterovirus	350	Monkeypox	416	Pediculosis corporis
219	Hepatitis B	285	Chytridiomycosis	351	Fasciolosis	417	Black piedra
220	Ebola virus disease	286	Tularemia	352	Paracoccidioidomycosis	418	H. bocavirus
221	Common cold	287	Kawasaki disease	353	Hand, foot, and mouth disease	419	Ehrlichiosis ewingii inf.
222	Rabies	288	Chikungunya	354	Vibrio vulnificus	420	Desmodesmus
223	Dengue fever	289	Hepatitis D	355	Actinomycosis	421	Rhinosporidiosis
224	Helicobacter pylori	290	Dracunculiasis	356	Ureaplasma urealyticum	422	Free-living Amoebozoa inf.
225	Lyme disease	291	Keratitis	357	Tinea corporis	423	Geotrichosis
226	Chickenpox	292	Lymphatic filariasis	358	Melioidosis	424	A. haemolyticum[k]
227	Staphylococcus	293	Echinococcosis	359	Paragonimiasis	425	Mycetoma

Abbreviations H. and inf. stand for Human and infection. [a]SARS: Severe acute respiratory syndrome. [b]SSPE: Subacute sclerosing panencephalitis. [c]CCHF: Crimean–Congo hemorrhagic fever. [d]PML: Progressive multifocal leukoencephalopathy. [e]VEE: Venezuelan equine encephalitis virus. [f]MERS: Middle East respiratory syndrome. [g]HFRS: Hantavirus hemorrhagic fever with renal syndrome. [h]LCM: Lymphocytic choriomeningitis. [i]GSS: Gerstmann–Sträussler–Scheinker syndrome. [j]HME: Human monocytotropic ehrlichiosis. [k]A. haemolyticum: Arcanobacterium haemolyticum.

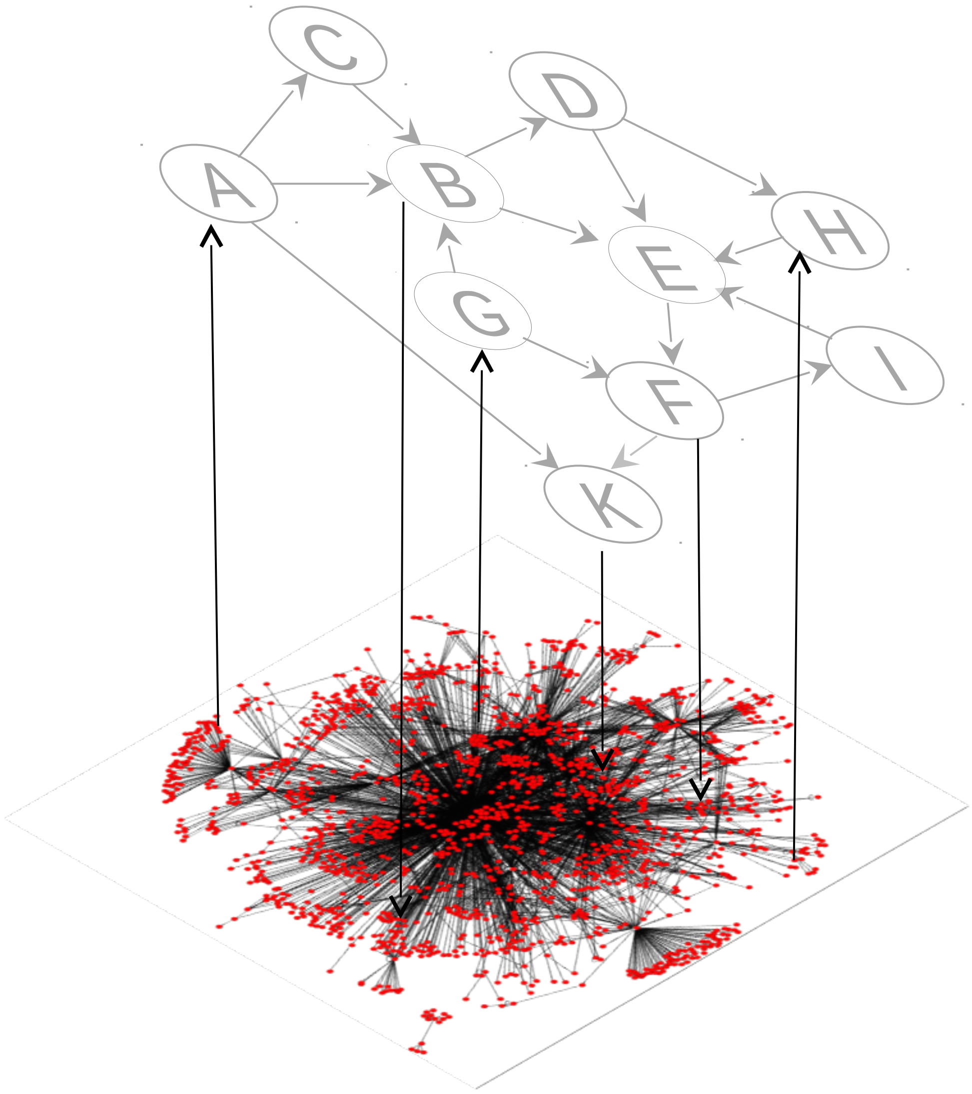
Reduced Google matrix

Consider a network with $N \gg 1$ nodes.



Reduced Google matrix

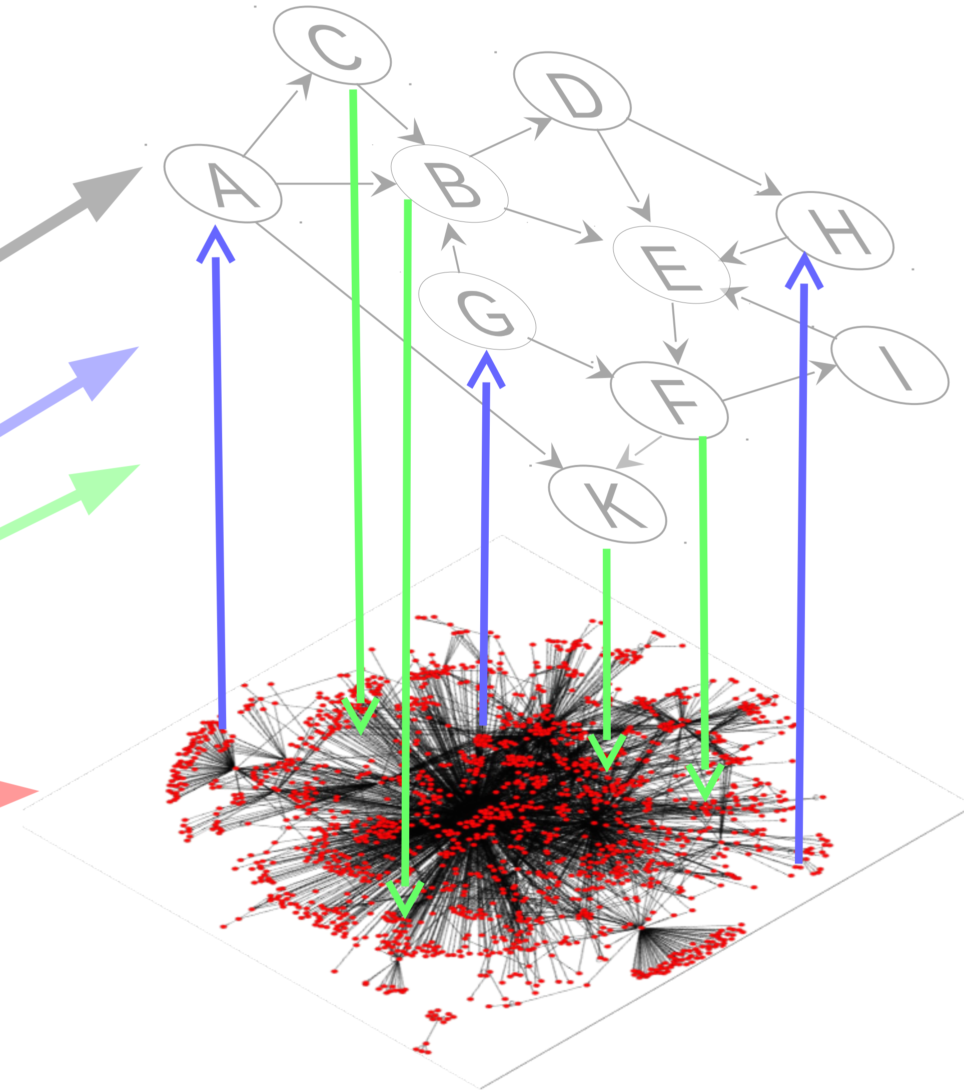
Consider a network with $N \gg 1$ nodes.
Consider a sub-network (a community) of $N_r \ll N$ nodes.



Reduced Google matrix

Consider a network with $N \gg 1$ nodes.
Consider a sub-network (a community) of $N_r \ll N$ nodes.
The Google matrix of the size N network and the associated PageRank vector can be written as

$$\mathbf{G} = \begin{pmatrix} \mathbf{G}_{rr} & \mathbf{G}_{rs} \\ \mathbf{G}_{sr} & \mathbf{G}_{ss} \end{pmatrix},$$



$$\mathbf{P} = \begin{pmatrix} \mathbf{P}_r \\ \mathbf{P}_s \end{pmatrix} \qquad \mathbf{G}\mathbf{P} = \mathbf{P}$$

Reduced Google matrix

Consider a network with $N \gg 1$ nodes.
Consider a sub-network (a community) of $N_r \ll N$ nodes.
The Google matrix of the size N network and the associated PageRank vector can be written as

$$\mathbf{G} = \begin{pmatrix} \mathbf{G}_{rr} & \mathbf{G}_{rs} \\ \mathbf{G}_{sr} & \mathbf{G}_{ss} \end{pmatrix}, \qquad \mathbf{P} = \begin{pmatrix} \mathbf{P}_r \\ \mathbf{P}_s \end{pmatrix}$$

$$\mathbf{G}\mathbf{P} = \mathbf{P}$$

We define the reduced Google matrix $\mathbf{G}_R$ associated to the community of size N_r such as

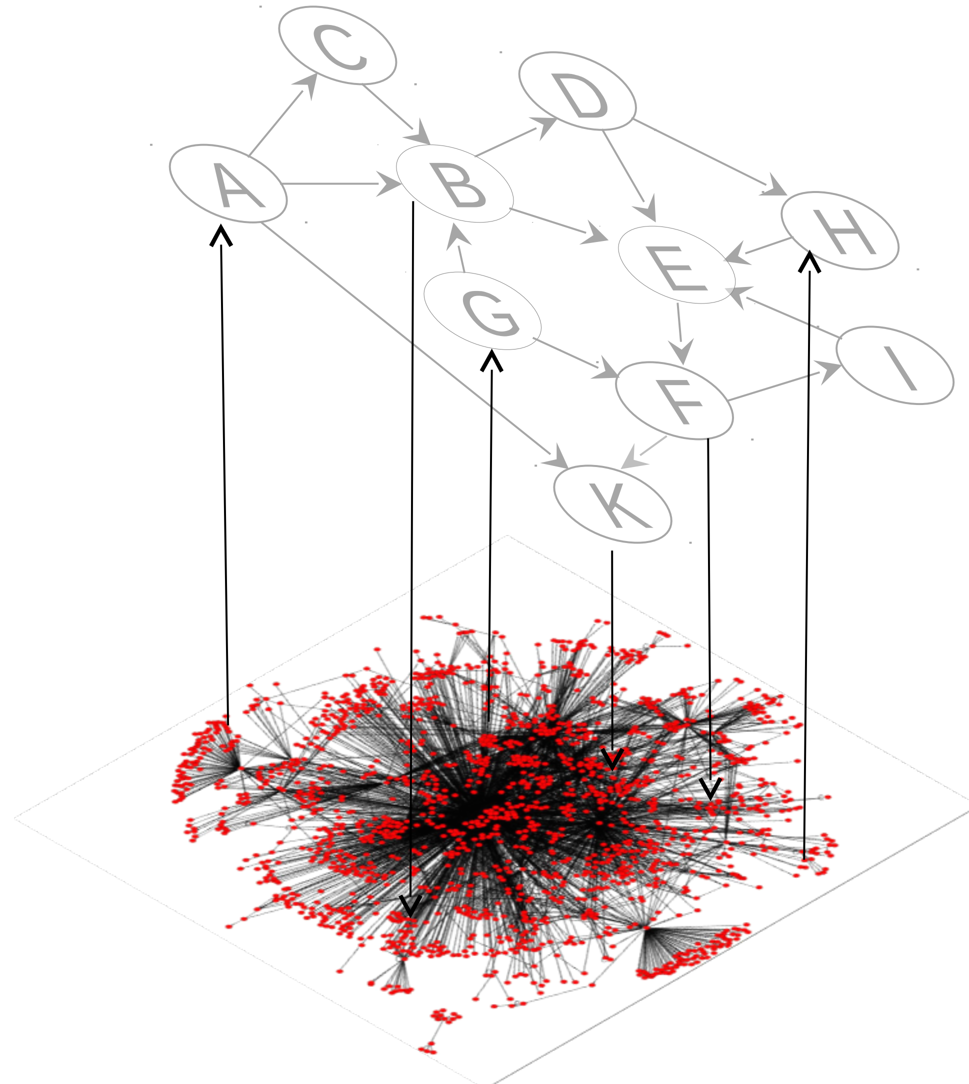
$$\mathbf{G}_R \mathbf{P}_r = \mathbf{P}_r$$

The reduced Google matrix can be written

$$\mathbf{G}_R = \mathbf{G}_{rr} + \mathbf{G}_{rs}\left(\mathbf{1} - \mathbf{G}_{ss}\right)^{-1}\mathbf{G}_{sr}$$

Contribution from direct links

Contribution from indirect links (scattering term)



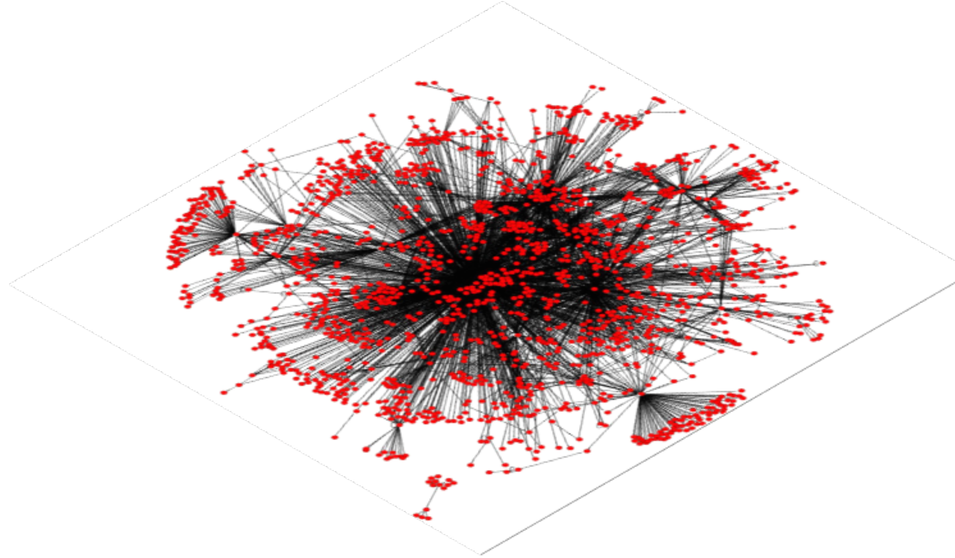
J. Lages, D. Shepelyansky, A. Zinovyev, PLoS ONE 13(1): e0190812 (2018)
K. M. Frahm, and D. L. Shepelyansky, arXiv:1602.02394 [physics.soc-ph]

Very slow convergence since the eigenvalue λ_c of $\mathbf{G}_{ss} \sim \mathbf{G}$ is very close to 1.

$$\left(\mathbf{1} - \mathbf{G}_{ss}\right)^{-1} = \sum_{l=0}^{\infty} \mathbf{G}_{ss}^{l}$$

Reduced Google matrix

Consider a network with $N \gg 1$ nodes.
Consider a sub-network (a community) of $N_r \ll N$ nodes.
The Google matrix of the size N network and the associated PageRank vector can be written as



$$\mathbf{G} = \begin{pmatrix} \mathbf{G}_{rr} & \mathbf{G}_{rs} \\ \mathbf{G}_{sr} & \mathbf{G}_{ss} \end{pmatrix}, \qquad \mathbf{P} = \begin{pmatrix} \mathbf{P}_r \\ \mathbf{P}_s \end{pmatrix}$$

$$\mathbf{G}\mathbf{P} = \mathbf{P}$$

We define the reduced Google matrix $\mathbf{G}_R$ associated to the community of size N_r such as

$$\mathbf{G}_R \mathbf{P}_r = \mathbf{P}_r$$

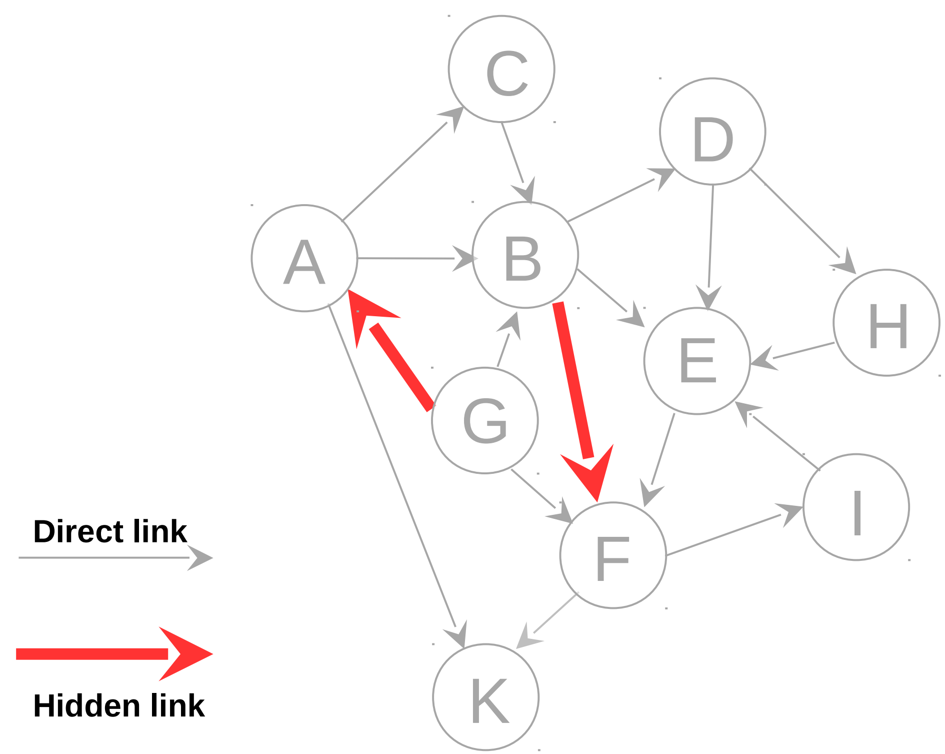
The reduced Google matrix can be written

$$\mathbf{G}_R = \mathbf{G}_{rr} + \mathbf{G}_{pr} + \mathbf{G}_{qr}$$

Contribution from direct links

Contribution from hidden links

Contribution from « PageRank »



Reduced Google matrix

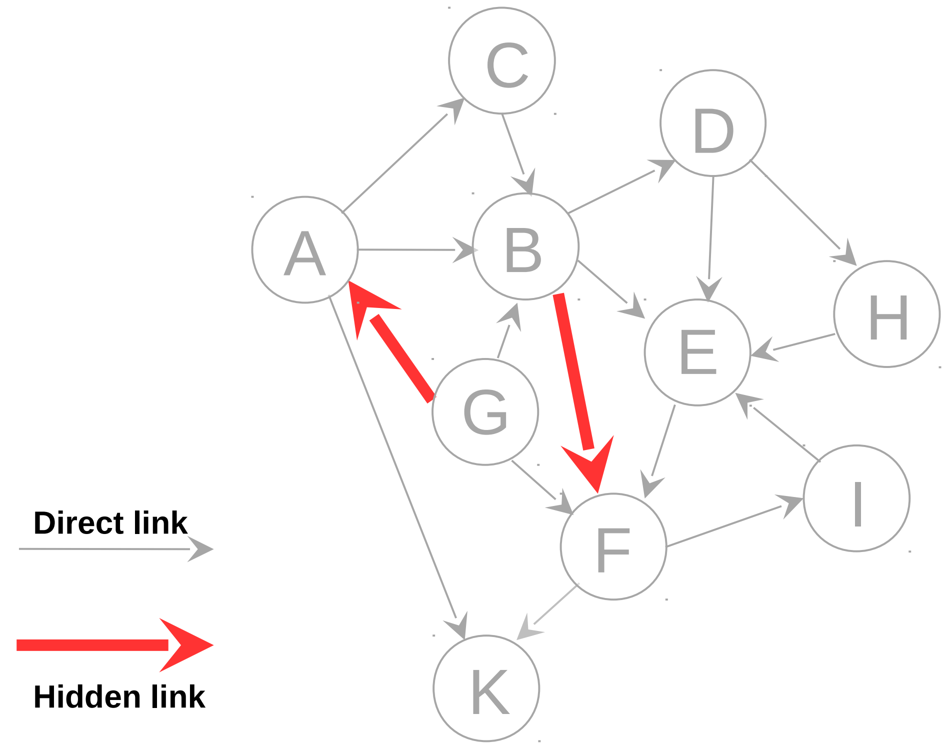
Consider a network with $N \gg 1$ nodes.
Consider a sub-network (a community) of $N_r \ll N$ nodes.
The reduced Google matrix can be written

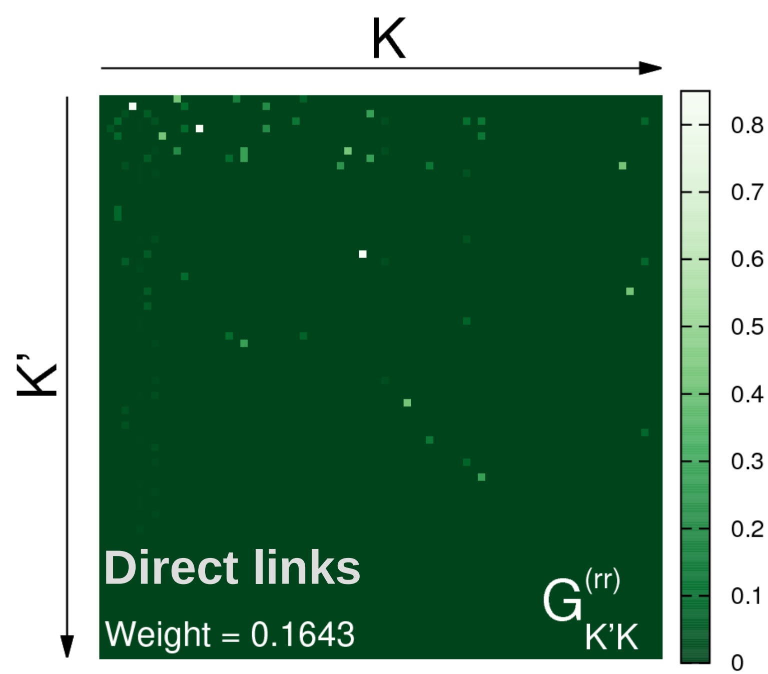
$$\mathbf{G}_R = \mathbf{G}_{rr} + \mathbf{G}_{pr} + \mathbf{G}_{qr}$$

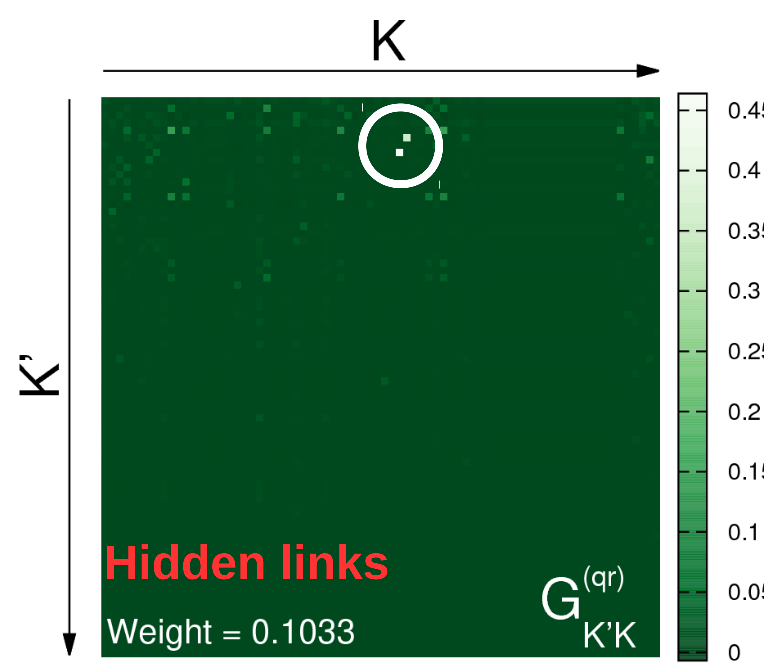
Contribution from direct links ($\mathbf{G}_{rr}$)

Contribution from « PageRank » ($\mathbf{G}_{pr}$)

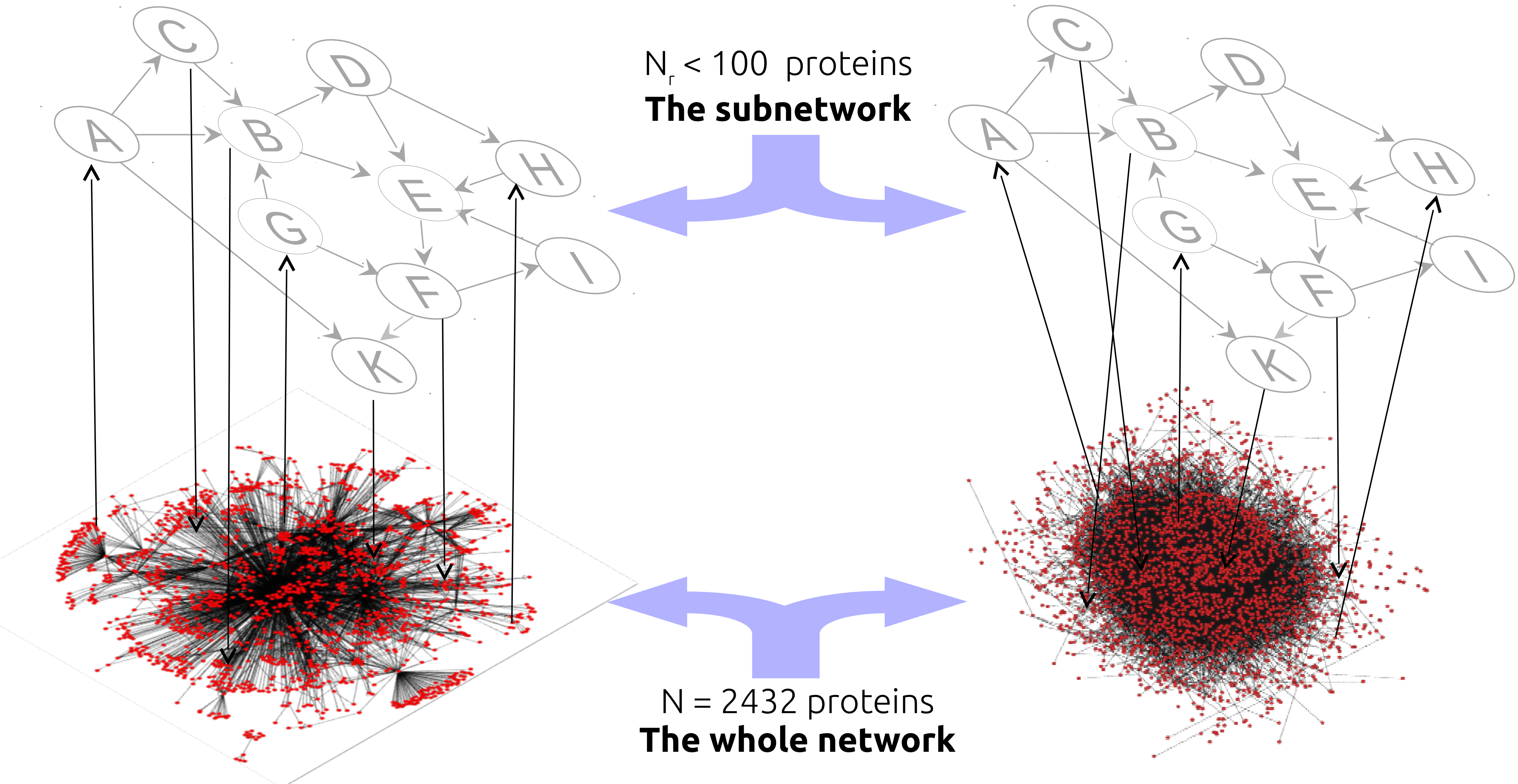
Contribution from hidden links ($\mathbf{G}_{qr}$)







Googlomics : Inferring hidden causal relations between proteins



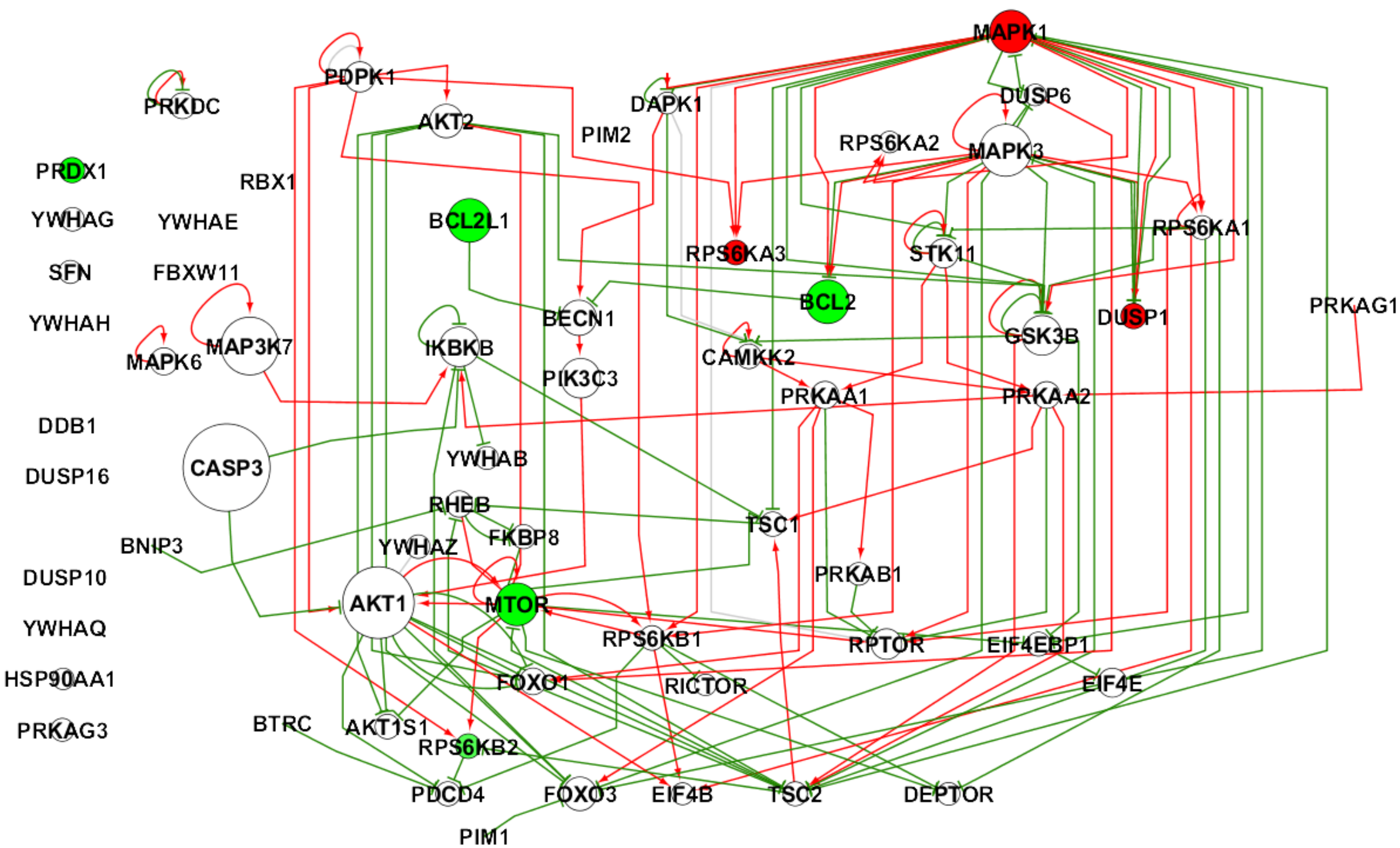
TRN1 (“**normal**”)
Normal B-Lymphocytes

TRN2 (“**cancer**”)
Leukemia cell line

J. Lages, D. Shepelyansky, A. Zinovyev, PLoS ONE 13(1): e0190812 (2018)

Googlomics : Inferring hidden causal relations between proteins

Inferring indirect (hidden) causal connections between **AKT-mTOR pathway members**



direct activation

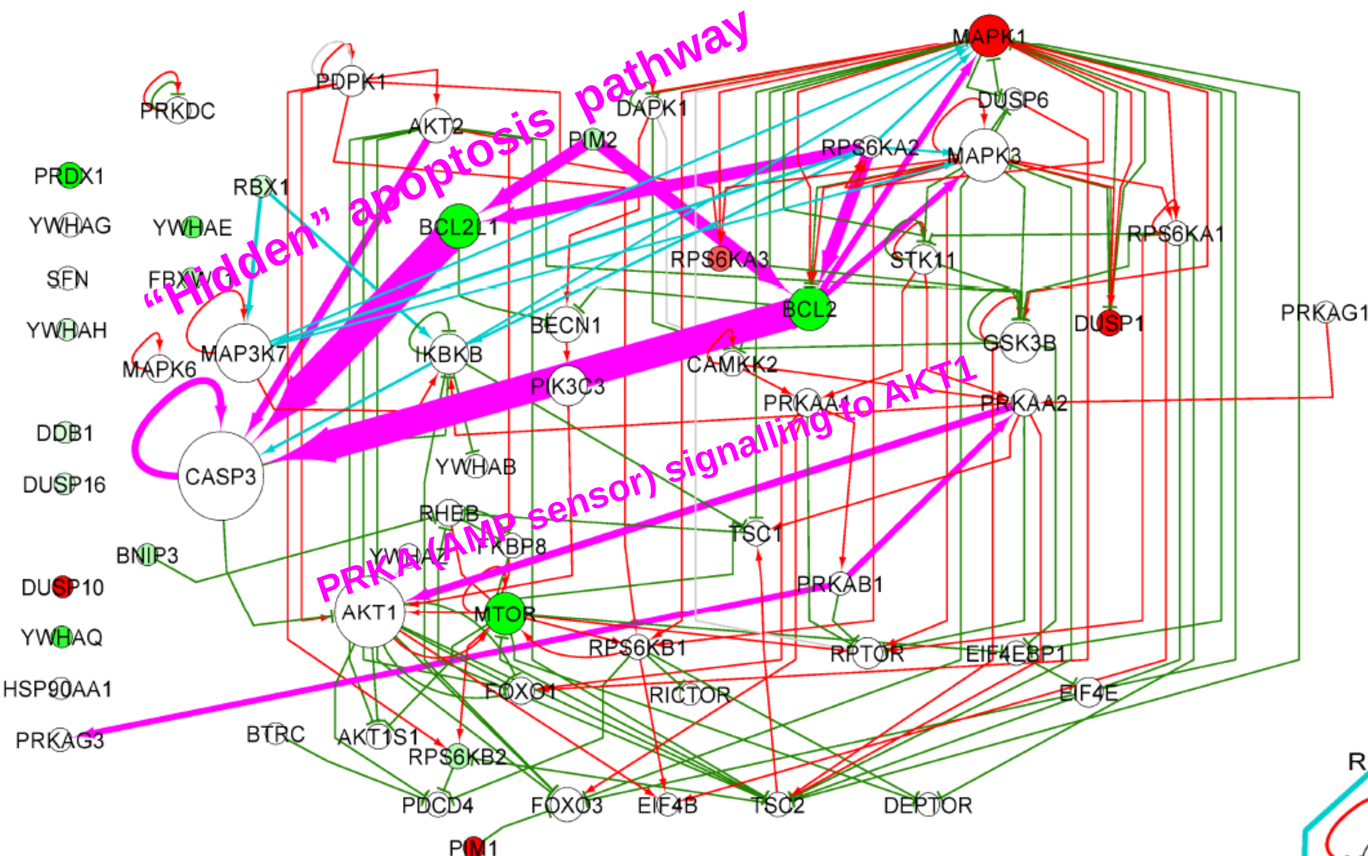
direct inhibition

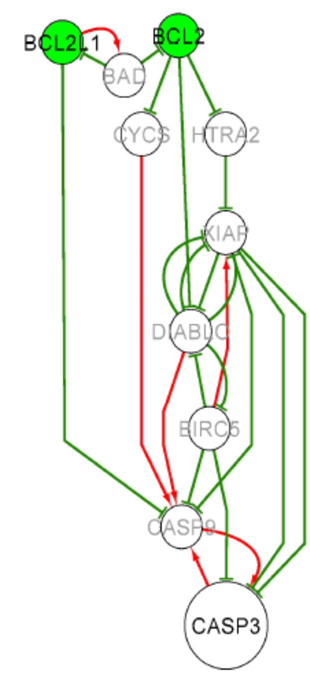
Subnetwork of 63 proteins

J. Lages, D. Shepelyansky, A. Zinovyev, PLoS ONE 13(1): e0190812 (2018)

Googlomics : Inferring hidden causal relations between proteins

Inferring indirect (hidden) causal connections between **AKT-mTOR pathway members**





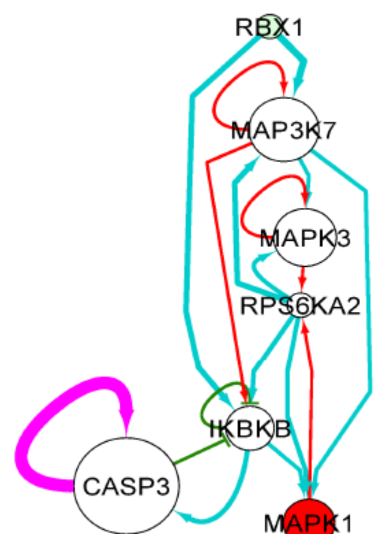
direct activation

direct inhibition

hidden (indirect)

hidden emergent oncogenic

Subnetwork of 63 proteins

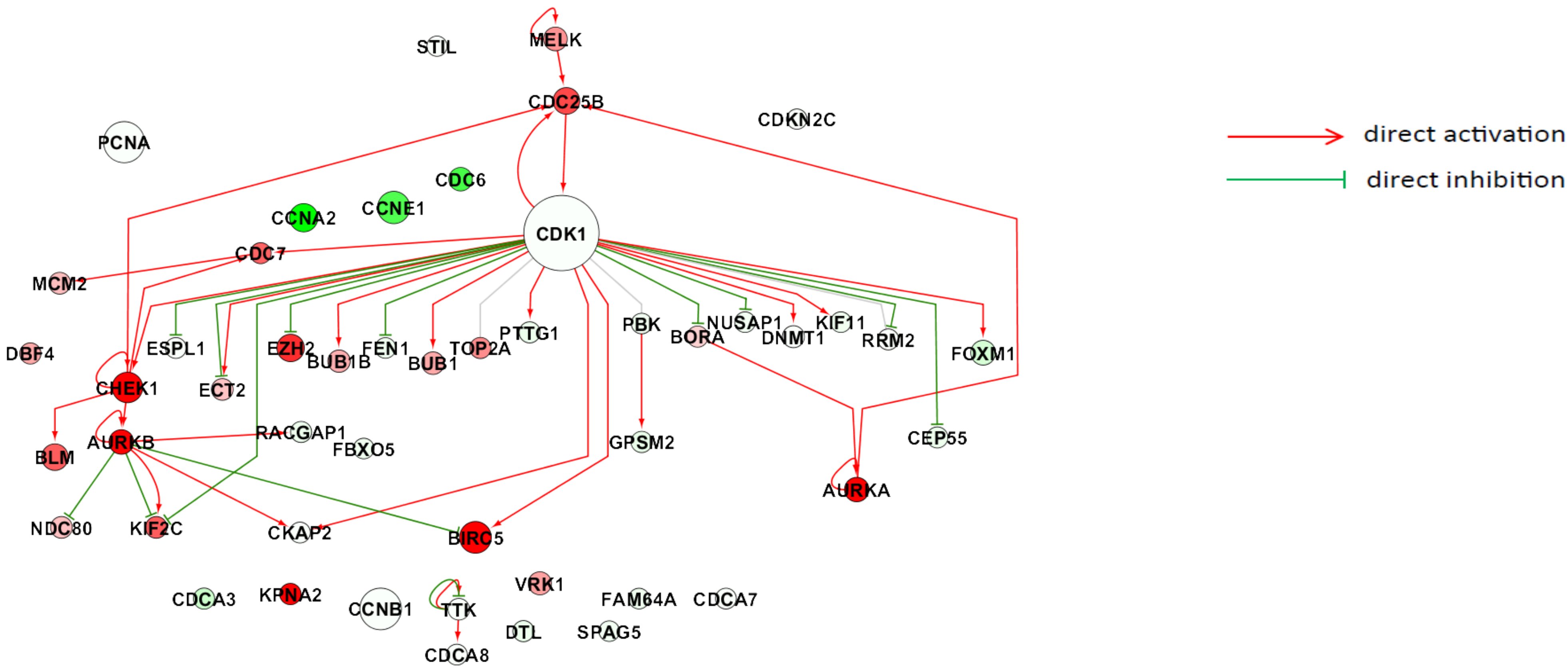


Emergent oncogenic signaling between RBX1 (cell cycle protein degradation proteasome) and MAPK1

J. Lages, D. Shepelyansky, A. Zinovyev, PLoS ONE 13(1): e0190812 (2018)

Googlomics : Inferring hidden causal relations between proteins

Genes of a proliferative signature resulted from pancancer transcriptomic analysis

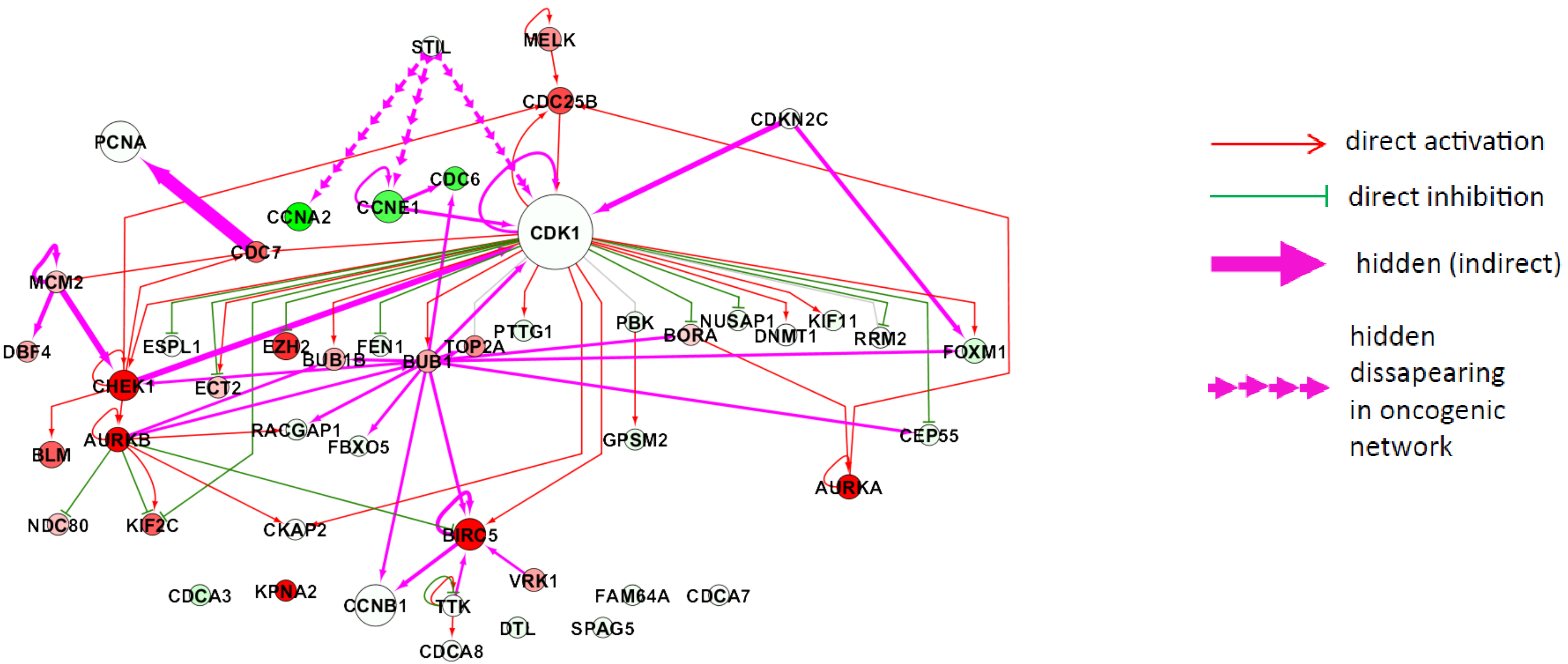


Subnetwork of 49 proteins

J. Lages, D. Shepelyansky, A. Zinovyev, PLoS ONE 13(1): e0190812 (2018)

Googlomics : Inferring hidden causal relations between proteins

Genes of a proliferative signature resulted from pancancer transcriptomic analysis



Subnetwork of 49 proteins

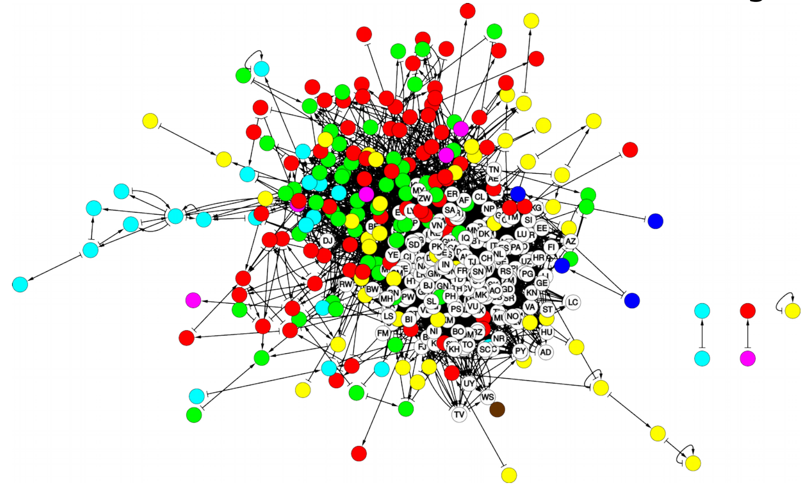
More genes are connected into the network
Emergence of a new "hidden" hub BUB1
Connection to PCNA (DNA replication and DNA repair)
Many cell cycle proteins improves in PageRank (AURK)
Connection between STIL (mitotic spindle checkpoint regulator) and CCNA2, CCNE1

J. Lages, D. Shepelyansky, A. Zinovyev, PLoS ONE 13(1): e0190812 (2018)

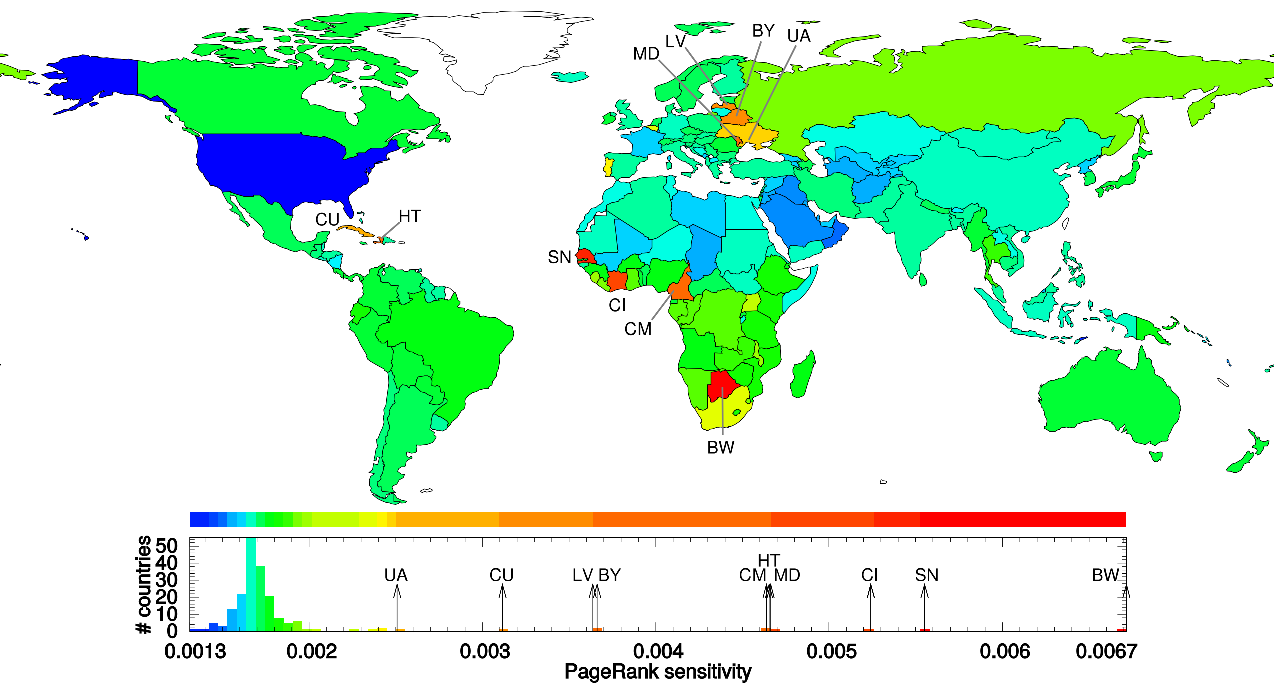
Influence of Infectious Diseases from Wikipedia Network Analysis

(preliminary results…)

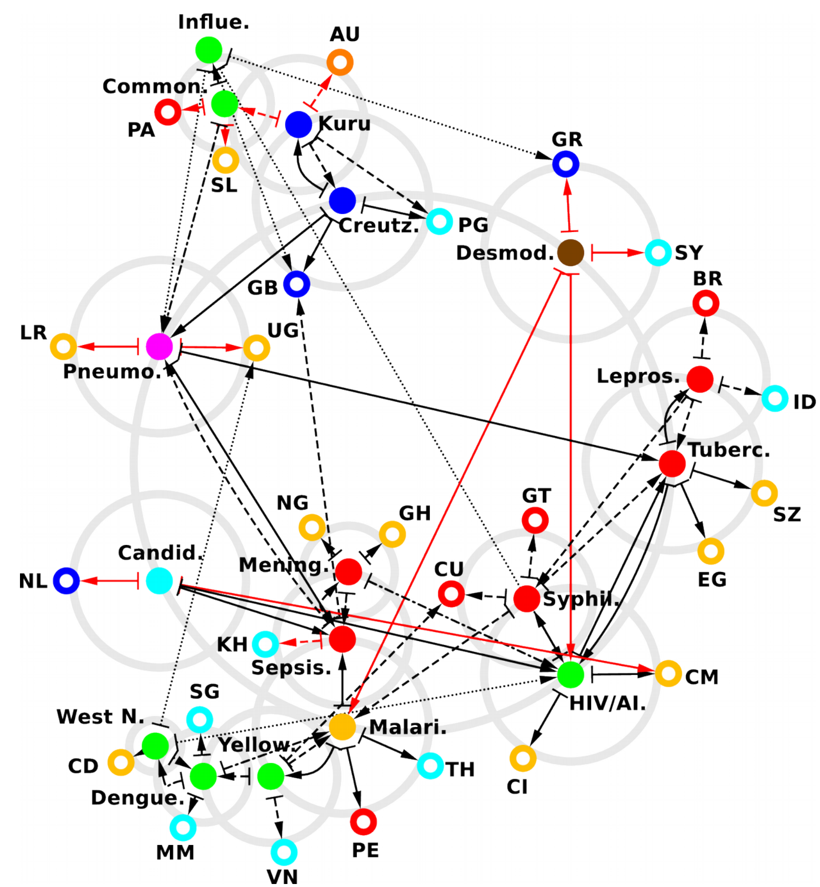
Subset of infectious diseases and countries in 2017 English Wikipedia



Sensitivity of countries to HIV



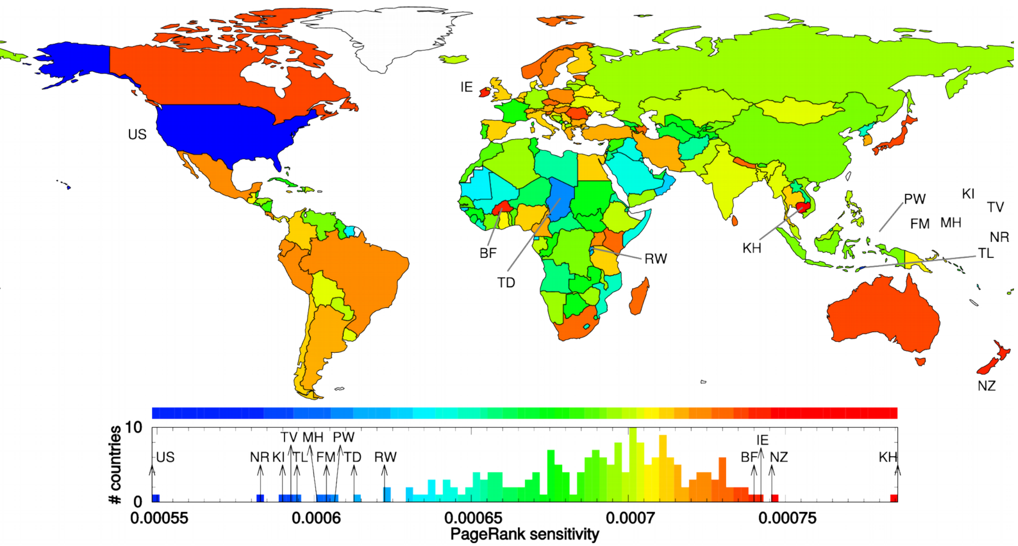
Reduced network of diseases and countries



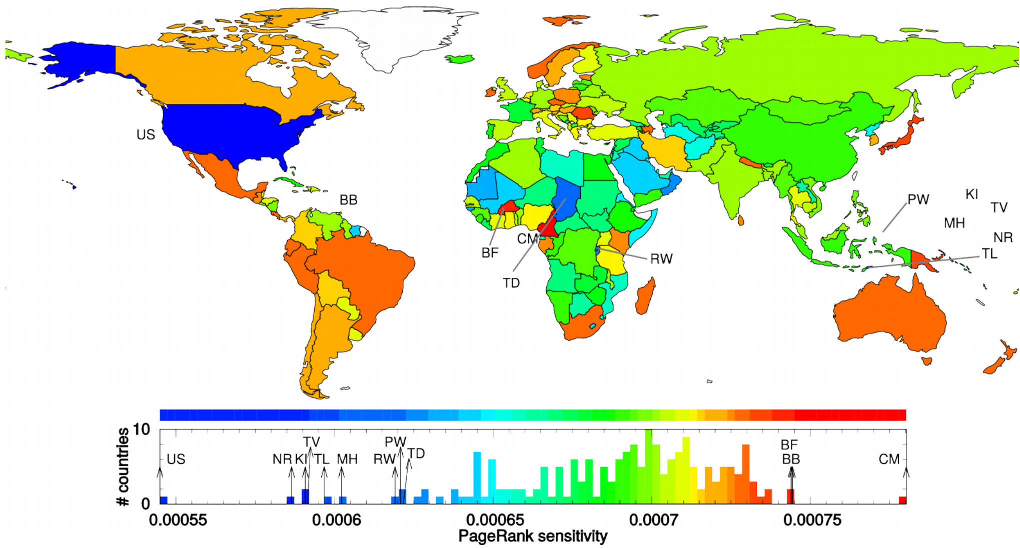
Influence of Cancers from Wikipedia Network Analysis

(preliminary results...)

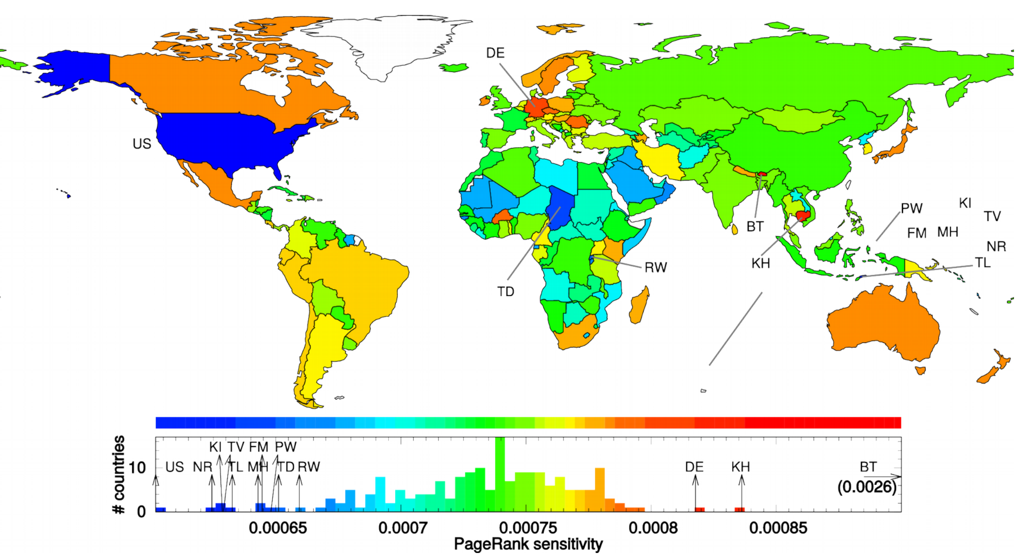
Sensitivity of countries to breast cancer



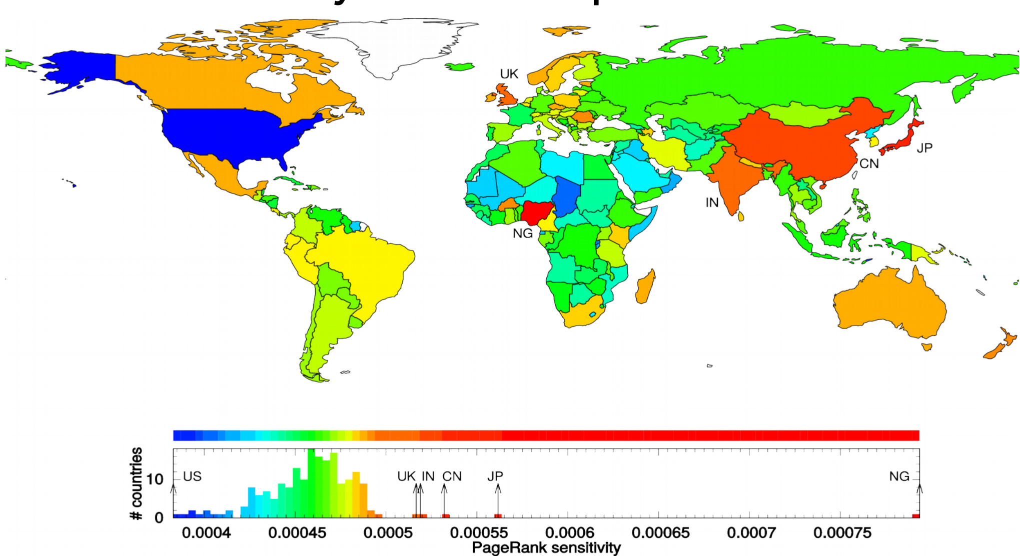
Sensitivity of countries to leukemia



Sensitivity of countries to lung cancer



Sensitivity of countries to prostate cancer



Thank You !