Physica A 542 (2020) 123199

Contents lists available at ScienceDirect

Physica A

journal homepage: www.elsevier.com/locate/physa

What is the central bank of Wikipedia?

Denis Demidov [a], Klaus M. Frahm [b], Dima L. Shepelyansky [b,*]

[a] Kazan Branch of Joint Supercomputer Center, Scientific Research Institute of System Analysis, Russian Academy of Sciences, 42011 Kazan, Russia

[b] Laboratoire de Physique Théorique, IRSAMC, Université de Toulouse, CNRS, UPS, 31062 Toulouse, France

ARTICLE INFO

Article history:
Received 2 June 2019
Received in revised form 30 September 2019
Available online 16 January 2020

ABSTRACT

We analyze the influence and interactions of 60 largest world banks for 195 world countries using the reduced Google matrix algorithm for the English Wikipedia network with 5 416 537 articles. While the top asset rank positions are taken by the banks of China, with China Industrial and Commercial Bank of China at the first place, we show that the network influence is dominated by USA banks with Goldman Sachs being the central bank. We determine the network structure of interactions of banks and countries and PageRank sensitivity of countries to selected banks. We also present GPU oriented code which significantly accelerates the numerical computations of reduced Google matrix.

1. Introduction

The world influence of the bank system was clearly demonstrated by the financial crisis of 2007–2008 whose impact appeared on financial, economic and political levels of many world countries (see e.g. [1,2]). The world wide contamination propagation of this crisis pushed forward the scientific analysis of bank networks which allows to detect interconnections of financial flows between banks [3–6]. For a bank network it is important to know what is the central bank since it may propagate its influence via the financial network to many other banks all over the world. On a first glance one could expect that a bank with the biggest total asset will be the most influential one. The list of the largest 60 world banks, ranked by their total assets, is available at [7] on the basis of the S&P Global Market Intelligence report 2018 [8]. The top position of rank index $K_a = 1$ is taken by Industrial and Commercial Bank (ICB) of China. The distribution of these 60 banks over world countriesis given in Table 1 and their names and AssetRank index K_a are given in Table 2. The first 4 positions are taken by banks of China and the first US bank appears only at 6th position taken by JPMorgan Chase & Co. The 5th position is taken by Mitsubishi UFJ Financial Group from Japan.

However, in the network of nodes the influence is determined by the links between nodes. Indeed, the PageRank algorithm was proposed by Brin and Page in 1998 [9] with the aim to rank the nodes of the World Wide Web (WWW). The method uses the construction of the Google matrix of Markov chain transitions between network nodes connected by directed links. This efficient algorithm became the foundation of the Google search engine [9,10]. The efficiency of this approach has been demonstrated for various types of directed networks including WWW, Wikipedia networks [11] and the world trade networks [12,13]. The PageRank probability vector being the eigenvector of the largest eigenvalue of the Google matrix determines the most important and influential nodes on the network [9,10].

* Corresponding author.
E-mail addresses: dennis.demidov@gmail.com (D. Demidov), frahm@irsamc.ups-tlse.fr (K.M. Frahm), dima@irsamc.ups-tlse.fr (D.L. Shepelyansky).

https://doi.org/10.1016/j.physa.2019.123199
0378-4371/

Table 1
List of 15 countries associated with the 60 richest banks from [7] with n_b being the number of banks for each country appearing in the bank list. The second column represents α_2 ISO 3166-1 country codes. The colors mark five groups with group index n_g: US and Canada ($n_g = 1$, red), Switzerland and UK ($n_g = 2$, olive), EU countries without UK ($n_g = 3$, green), India, Australia and Japan ($n_g = 4$, cyan) and China ($n_g = 5$, blue).

Country	α_2	n_b	n_g
United States	**US**	6	1
United Kingdom	**UK**	5	2
Germany	**DE**	3	3
Switzerland	**CH**	2	2
France	**FR**	5	3
Canada	**CA**	5	1
Netherlands	**NL**	2	3
Italy	**IT**	3	3
India	**IN**	1	4
Australia	**AU**	4	4
Spain	**ES**	2	3
China	**CN**	13	5
Finland	**FI**	1	3
Japan	**JP**	7	4
Denmark	**DK**	1	3

In [3] it was proposed to use the PageRank vector and to consider that the central bank is the one at the top of PageRank probability with PageRank index $K_b = 1$. The analysis was done for the Canadian Large Value Transfer System with 14 banks. While this study presents an interesting approach to the analysis of bank influence it is clear that the network of 14 Canadian banks is too small to make any conclusion for the world influence of banks. Also it is highly difficult to obtain information on monetary transfers between banks on the world scale due to the high secrecy of bank operations. At the same time it is clear that the monetary transfers do not provide the whole information about the real influence of banks since it includes also political, social, historical and other types of links and relations which are not directly expressed in money.

Due to these reasons we use another approach based on the Wikipedia network analysis. Indeed, Wikipedia accumulates a great amount of human knowledge and supersedes other encyclopedia such as Encyclopedia Britannica [14]. The academic analysis of information contained in Wikipedia finds more and more applications as reviewed in [15,16]. Scientific analysis shows that the quality of Wikipedia articles is growing [17].

Thus we take here the whole network of English Wikipedia dated by May 2017 [18] containing $N = 5\,416\,537$ articles (nodes) with $N_\ell = 122\,232\,932$ directed links between them generated by citations of articles in other articles. The global Google matrix of this directed network is created by the standard rules described in [9–11]. To analyze the interactions and influence of banks we select $N_b = 60$ largest banks from [7], given also in Table 2, and $N_c = 195$ world countries. Then we perform the reduced Google matrix (REGOMAX) analysis [19,20] of these selected $N_r = N_b + N_c = 255$ nodes. The REGOMAX approach allows to determine interactions between banks and countries by direct and indirect links between these selected nodes. The indirect links take into account all possible pathways from one node to another one via the global network of 5 millions of Wikipedia articles. The efficiency of the REGOMAX method has been demonstrated for various examples such as interactions between politicians [20], countries [21], world universities [22] and cancer networks [23].

The results of our analysis show that the central bank of Wikipedia is not at all ICB China with the largest asset $K_a = 1$ but Goldman Sachs which has the asset rank $K_a = 35$. We also determine the influence of 60 banks on 195 world countries.

We note that the power of matrix method for financial analysis has been demonstrated in [24,25]. The new mathematical element here is that we work with the directed networks and matrices belonging to the Perron–Frobenius operator class.

The complete data and additional figures obtained in this work are available at [26] including the list of 195 countries.

We briefly describe the Google matrix methods in Section 2, the rank plane of banks is analyzed in Section 3, the reduced Google matrix of banks and countries is described in Section 4, the network structure of friends and followers for banks and countries is shown and discussed in Section 5, the world country sensitivity to specific banks are obtained in Section 6 and the discussion of the results is given in Section 7.

2. Google matrix methods

The Google matrix G is constructed from the adjacency matrix A_{ij} with elements 1 if article (node) j points to article (node) i and zero otherwise. The matrix elements have the standard form $G_{ij} = \alpha S_{ij} + (1-\alpha)/N$ [9–11], where S is the matrix of Markov transitions with elements $S_{ij} = A_{ij}/k_{out}(j)$ and $k_{out}(j) = \sum_{i=1}^{N} A_{ij} \neq 0$ being the out-degree of node j (number of outgoing links); $S_{ij} = 1/N$ if j has no outgoing links (dangling node). The parameter $0 < \alpha < 1$ is the damping factor. We use the standard value $\alpha = 0.85$ [10] noting that for the range $0.5 \leq \alpha \leq 0.95$ the results are not sensitive to α [10,11]. For a random surfer, moving from one node to another, the probability to jump to any node is $(1-\alpha)$.

Table 2
Table of 60 most important banks. K_b is the relative rank number obtained from the Wikipedia PageRank of the Wikipedia article of each bank and K_a is the rank number due to the total value of assets of the Bank. The colors and α_2 are as in Table 1. The name "ICB China" is an abbreviation of "Industrial and Commercial Bank of China".

K_b	Bank	α_2	K_a
1	Goldman Sachs	US	35
2	Citigroup	US	13
3	Bank of America	US	9
4	HSBC	UK	7
5	JPMorgan Chase	US	6
6	Barclays	UK	18
7	Deutsche Bank	DE	15
8	Morgan Stanley	US	38
9	UBS	CH	34
10	Wells Fargo	US	11
11	Credit Suisse	CH	40
12	BNP Paribas	FR	8
13	Royal Bank of Canada	CA	24
14	Société Générale	FR	19
15	Standard Chartered	UK	48
16	ING Group	NL	26
17	Bank of Montreal	CA	52
18	UniCredit	IT	28
19	Scotiabank	CA	42
20	State Bank of India	IN	55
21	Commonwealth Bank	AU	43
22	Banco Santander	ES	16
23	Crédit Agricole	FR	10
24	National Australia Bank	AU	50
25	Westpac	AU	47
26	Australia and New Zealand BG	AU	45
27	Lloyds Banking Group	UK	23
28	ICB China	CN	1
29	Nordea	FI	46
30	Mitsubishi UFJ Financial Group	JP	5
31	Banco Bilbao Vizcaya Argentaria	ES	41
32	Bank of China	CN	4
33	Toronto-Dominion Bank	CA	25
34	Canadian Imperial B. of Commerce	CA	59
35	Royal Bank of Scotland Group	UK	29
36	Mizuho Financial Group	JP	17
37	Rabobank	NL	44
38	Commerzbank	DE	54
39	Intesa Sanpaolo	IT	32
40	China Construction Bank	CN	2
41	Agricultural Bank of China	CN	3
42	Danske Bank	DK	53
43	Sumitomo Mitsui Financial Group	JP	14
44	Groupe BPCE	FR	20
45	Cassa Depositi e Prestiti	IT	57
46	DZ Bank	DE	51
47	Bank of Communications	CN	21
48	Resona Holdings	JP	56
49	Sumitomo Mitsui Trust Holdings	JP	60
50	Crédit Mutuel	FR	39
51	China Merchants Bank	CN	31
52	China Minsheng Bank	CN	36
53	Japan Post Bank	JP	12
54	Norinchukin Bank	JP	27
55	Ping An Bank	CN	58
56	China CITIC Bank	CN	37
57	Shanghai Pudong Development B.	CN	33
58	Industrial Bank (China)	CN	30
59	Postal Savings Bank of China	CN	22
60	China Everbright Bank	CN	49

The right PageRank eigenvector of G is the solution of the equation $GP = \lambda P$ for the unit eigenvalue $\lambda = 1$. The PageRank $P(j)$ values give positive probabilities to find a random surfer on a node j ($\sum_j P(j) = 1$). We order all nodes by decreasing probability P numbered by PageRank index $K = 1, 2, \ldots, N$ with a maximal probability at $K = 1$ and minimal at $K = N$. The numerical computation of $P(j)$ is done efficiently with the PageRank algorithm described in [9,10].

We also consider the original network with inverted direction of links. After inversion the Google matrix G^* is constructed via the same procedure with $G^*P^* = P^*$. The matrix G^* has its own PageRank vector $P^*(j)$ called CheiRank [27] (see also [11]). Its probability values can be again ordered in a decreasing order with CheiRank index K^* with highest P^* at $K^* = 1$ and smallest at $K^* = N$. On average, the high values of P (P^*) correspond to nodes with many ingoing (outgoing) links [10,11].

The REGOMAX method is described in detail in [19–21,23]. It allows to compute efficiently a "reduced Google matrix" of size $N_r \times N_r$ that captures the full contributions of direct and indirect pathways happening in the full Google matrix between N_r nodes of interest. For the selected N_r nodes their PageRank probabilities are the same as for the global network with N nodes, up to a constant multiplicative factor taking into account that the sum of PageRank probabilities over N_r nodes is unity. The computation of G_R provides a decomposition of G_R into matrix components that clearly distinguish direct from indirect interactions: $G_R = G_{rr} + G_{pr} + G_{qr}$ [20]. Here G_{rr} is given by the direct links between selected N_r nodes in the global G matrix with N nodes. In fact, G_{pr} is rather close to the matrix in which each column is given by the PageRank vector P_r, ensuring that PageRank probabilities of G_R are the same as for G (up to a constant multiplier). Hence G_{pr} does not provide much information about direct and indirect links between selected nodes.

The most interesting role is played by G_{qr}, which takes into account all indirect links between selected nodes appearing due to multiple pathways via the global network nodes N (see [19,20]). The matrix $G_{qr} = G_{qrd} + G_{qrnd}$ has diagonal (G_{qrd}) and non-diagonal (G_{qrnd}) parts with G_{qrnd} describing indirect interactions between nodes. The exact formulas for all three components of G_R are given in [19,20]. The numerical computation methods of all three components of G_R and numerical REGOMAX code developed in [19] are given in [28].

In this work we also ported the original REGOMAX code to NVIDIA CUDA technology using VexCL library [29,30], with two notable modifications. First, we reordered the global Google matrix using Cuthill–McKee method [31] in order to reduce the bandwidth of the matrix, and thus improve cache locality of the code. This resulted in about 30% reduction of the computational time. Second, the computation of the reduced matrix was performed in batches of 8–20 columns, which allowed us to further improve the performance of the method by another 50%. The computations were run at the OLYMPE CALMIP cluster [32] using NVIDIA Tesla V100 GPUs. Overall, we managed to achieve an approximately 8x speedup with respect to the original OpenMP implementation which was performed on a dual socket Intel(R) Xeon(R) CPU E5-2620 v2 @ 2.10 GHz system (2 × 6 cores). The REGOMAX-GPU code is available at [33].

With the matrix G_R and its components we can analyze the PageRank sensitivity with respect to specific links between $N_r = 255$ nodes. To measure the sensitivity of a country c to a bank b we change the matrix element $G_R(b \to c)$ by a factor $(1 + \delta)$ with $\delta \ll 1$ and renormalize to unity the sum of the column elements associated with bank b. Then we compute the logarithmic derivative of PageRank probability $P(c)$ associated to country c: $D(b \to c, c) = d \ln P(c)/d\delta$ (diagonal sensitivity). This approach already demonstrated its efficiency as reported in [22,34].

We point out that the REGOMAX algorithm does not change the PageRank probabilities of N_r selected nodes compared to their relative values in the global network of N nodes: it simply multiplies them by a constant factor related to probability normalization on N_r nodes instead of N nodes [20]. We also point out that we use G_R of size $N_r = 255$ including 60 banks and 195 countries. This is convenient for investigation of influence of banks on countries. However, we also checked that the interactions between banks only (with REGOMAX of size $N_r = 60$) are very close to their values inside the bank sector obtained with the REGOMAX algorithm for banks and countries at $N_r = 60 + 195$.

3. Rank plane of banks

We determine the PageRank index of countries and banks for $N_r = 255$. As discussed in [11,21] the top PageRank positions of the global matrix G are taking by countries. The PageRank local index K_b marking only banks is given in Table 2. The distribution of banks on the plane of PageRank K_b and AssetRank K_a is shown in Fig. 1 (top panel). While the top AssetRank $K_a = 1$ is taken by ICB China its PageRank index is only $K_b = 28$ while the top PageRank position $K_b = 1$ belongs to Goldman Sachs even if its AssetRank is $K_a = 35$. *Thus the central bank of Wikipedia is Goldman Sachs bank*. While the top 4 positions of AssetRank are taken by banks of China the top 4 PageRank positions are taken by USA banks (plus HSBC of UK). Since Wikipedia includes a huge amount of knowledge accumulated by humanity we can say the most strong political and social influence belongs to USA banks.

In contrast to the PageRank vector, which is reflecting the influence, the CheiRank vector reflects the communicative features of nodes. Thus the banks on the top positions of CheiRank index of banks K^*_b have the most communicative articles (nodes) among banks. The distribution of banks of the PageRank-CheiRank plane (K_b, K^*_b) is shown in the bottom panel of Fig. 1. The top CheiRank positions $K^*_b = 1, 2, 3$ are taken by Goldman Sachs, UBS, Bank of America, while ICB China has $K^*_b = 39$ showing very low communicativity of its Wikipedia article.

4. Reduced google matrix

Using the REGOMAX algorithm we determine the matrix G_R and its 3 components for $N_r = 255$ banks and countries. These matrices are available at [26]. We determine the weight of each component as the sum of all matrix elements divided by N_r (the weight of G_R is by definition $W_R = 1$). The weights of the components G_{pr}, G_{rr}, G_{qr} and G_{qrnd} are respectively $W_{pr} = 0.8912$, $W_{rr} = 0.0402$, $W_{qr} = 0.06859$ and $W_{qrnd} = 0.04417$. The weights of components G_{rr}, G_{qr} are

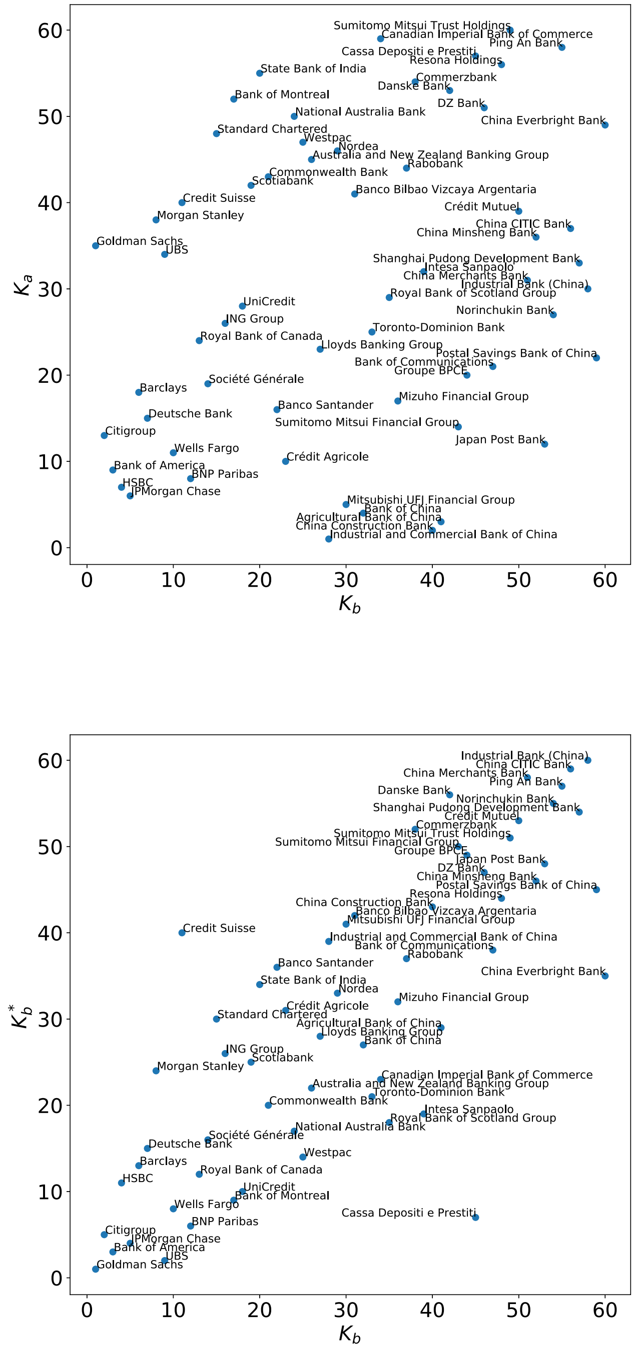

Fig. 1. Top panel: PageRank-AssetRank plane of 60 banks with the world largest assets from [7]; bottom panel: PageRank-CheiRank plane of the same banks.

relatively small but as it was discussed above and in [20–22] they provide the most interesting information on interactions of nodes.

Here, in Fig. 2 we show the parts of G_R belonging to the 60×60 corner corresponding its bank–bank sector. The whole matrix 255 is available at [26]. The weight of matrix elements of this 60×60 sector is relatively small with $W_R = 0.0800$ and $W_{pr} = 0.00598$, $W_{rr} = 0.03567$, $W_{qr} = 0.03843$ and $W_{qrnd} = 0.02267$. Thus we see that this sector gives only a relatively small contribution for the full matrix of banks and countries of size 255×255. However, this 60×60 sector describes important interactions between banks. Here we see that there is for example a strong indirect link in G_{qr} from the Postal Savings Bank of China ($K_b = 59$) to Goldman Sachs ($K_b = 1$) with matrix element 3.74×10^{-3} while in G_{rr} the

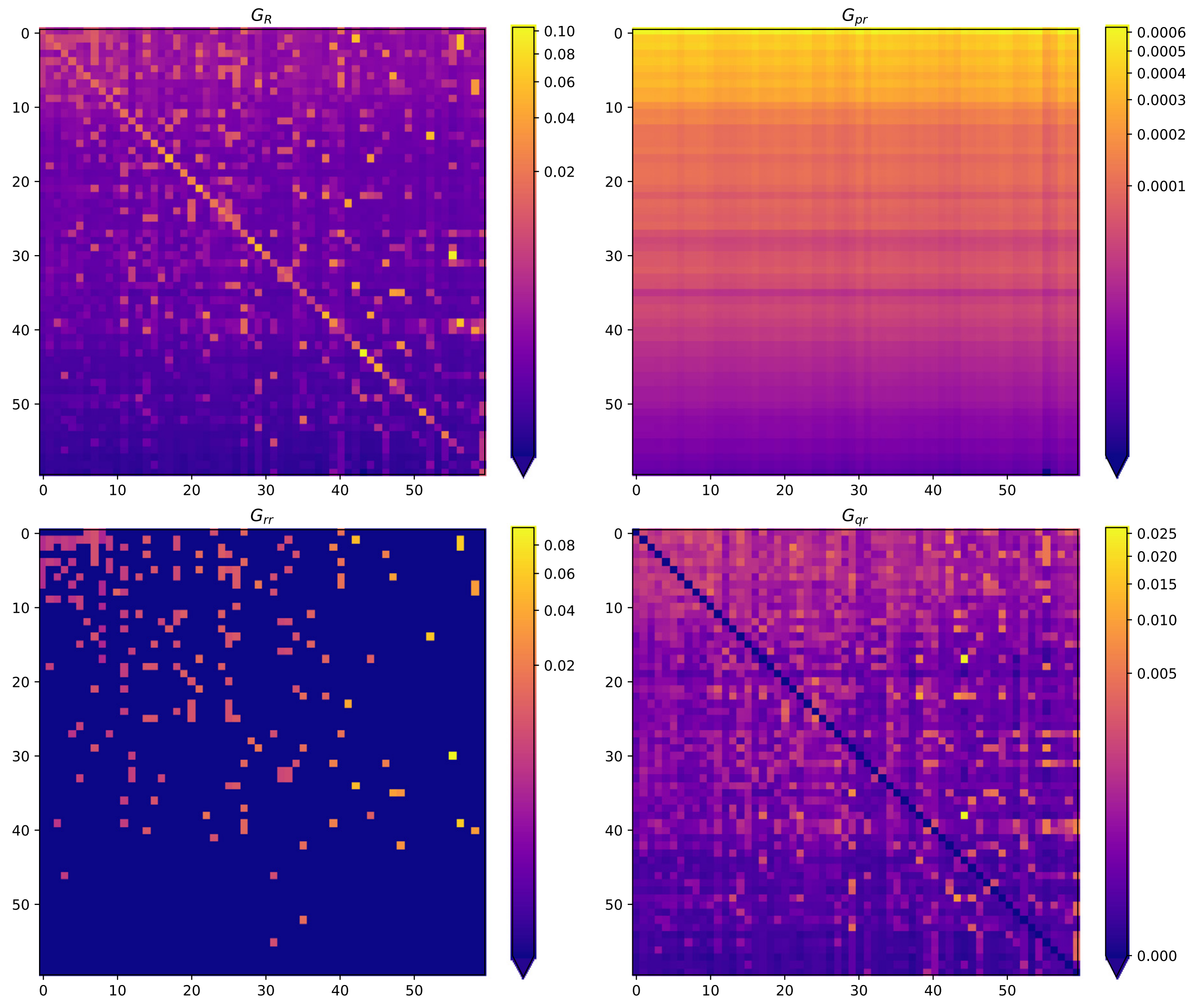

Fig. 2. Matrix elements of reduced Google matrix G_R and its 3 components G_{pr}, G_{rr}, G_{qr} (without diagonal) are shown for the bank sector 60×60 of the whole matrix 255×255; to improve the visibility in all 4 panels the color bars mark the values $G^{1/4}$ of matrix elements in power 1/4 (some values of G are marked on the color bars). Bank index on axis is given by the PageRank index K_b of Table 2.

same link is nearly absent with matrix element $2.77 \times 10^{-8} = (1 - \alpha)/N$, only non-zero due to the minimal damping factor contribution.

5. Networks of banks and countries

At present the network presentation of link interactions between nodes gained a significant popularity due to its compact presentation of interactions (see e.g. [35]). In directed networks we will say that a node j has a friend i if j points to i and i has a follower j. Such a description has been used for interactions between politicians [20], countries [21] and interactions between infectious diseases and countries [34].

We construct similar types of networks from our reduced Google matrix G_R of size 255×255 for all 60 banks and 195 countries. For a better visibility we attribute all 60 banks to 5 groups according to geographical and historical links between countries and banks as it shown in Tables 1, 2. The top PageRank banks of each group represent the central group nodes being Goldman Sachs (red for North America), HSBC (olive for Europe non-EU), Deutsche Bank (green for Europe EU), State Bank of India (cyan for India, Japan and Australia), and ICB China (blue for China). These 5 top nodes define the set of level 0 nodes. Assuming that the network has already been constructed up to level j nodes (with integer $j \geq 0$) we determine for each level j bank node 4 bank friends (followers) with 4 largest bank matrix elements of G_R in the same column (row). If such a friend (follower) is not yet present in the network of levels up to j we add it to the network as level $(j + 1)$ node. Furthermore, we also add in the same way (up to) 2 country friends (followers) with 2 largest country matrix elements of G_R in the same column (row) for each level j bank node (we do not select followers or friends for country nodes once they appear in the network).

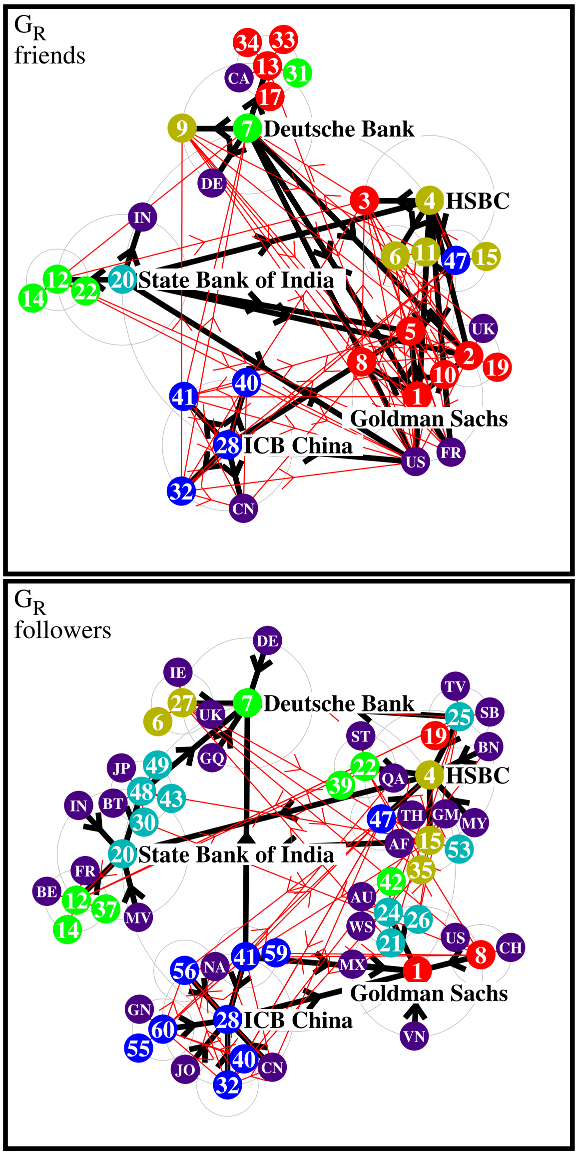

Fig. 3. Network of friends (top) and followers (bottom) from G_{R} matrix of 60 banks and 195 countries; 2 depth levels are shown; links starting from (for friends) or arriving to (for followers) level 0 nodes are shown by thick black arrows and similarly links starting from/arriving to level 1 nodes are shown by thin red arrows; bank nodes are marked by their PageRank index K_b, given in Table 2, and their group color defined in Table 1; countries are marked by their α_2 ISO 3166-1 country code and the color indigo.

The resulting friend (follower) network is shown in top (bottom) panel of Fig. 3 for G_R and maximal depth at level 2. The links to friends (from followers) are drawn as arrows (thick black line if $j = 0$ or thin red line if $j = 1$ in the above scheme). The nodes are drawn on gray circles of three different sizes for the levels 0, 1 and 2. A new (bank) node is attributed to the circle of the bank of the same group/color if there is a link. Otherwise it is attributed to the circle of the first (in K_b order) bank of the previous level to which a link exists.

We see that in the friend network of G_R there is a large cluster of USA banks located around Goldman Sachs with the two friend countries US, FR. The friends of ICB China are three other Chinese banks, one US bank and the two countries CN and US. The bank friends of the State Bank of India are two European and two US banks together with IN and US as country friends. The bank HSBC has two US bank friends, one other UK bank friend and a CN bank friend together with US and FR as country friends. The bank friends of the Deutsche Bank are from Switzerland (1 bank) and US (3 banks) and its country friends are DE and US.

In the follower network of G_R, we observe a lot of countries, quite low in the PageRank order for countries, which are level 1 or level 2 followers and do not have a bank of their own in the list of top 60 banks (i.e. countries which do not

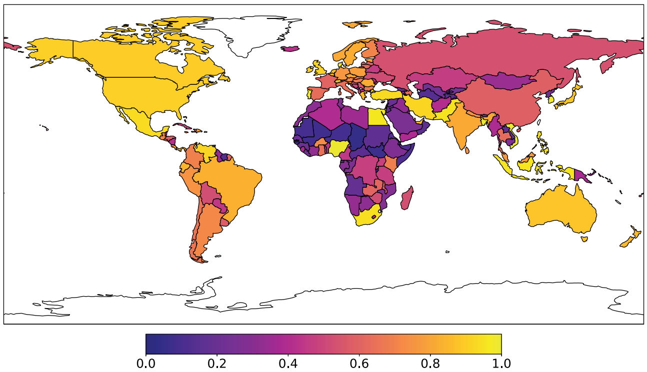

Fig. 4. Sensitivity D of world countries to Goldman Sachs shown by color changing from zero level (blue color for $D = 1.03 \times 10^{-3}$) to maximal unity level (yellow color for $D = 5.30 \times 10^{-4}$).

appear in Tables 1 or 2). The followers of ICB China are other CN Banks (exclusively on level 1 and mostly on level 2) and its country followers are CN and JO. For the other four top banks the main bank followers are quite often from banks of other groups/colors. For the Deutsche Bank and the State Bank of India, the corresponding country (DE or IN) is also a country follower but the other country follower is a small country (GQ or MV) which for GQ is even from a very different geographical region. For Goldman Sachs and HSBC both country followers (MX, VN or MY, BN) are different from the country to which they belong. Certain other "small" countries are level 2 followers of level 1 banks. It seems that many countries with no powerful bank of their own depend on or are at least related to one or several big banks outside or in their region (for example MV for State Bank of India or MX for Goldman Sachs).

Networks built for other cases: other G_{R} components including $G_{rr} + G_{\mathrm{qr}}$, higher levels $j > 2$ and also the case of a pure bank network (without countries) are given in [26].

6. World bank influence on countries

In addition to the network structure on interactions of banks and countries, shown in Fig. 3, we can obtain a more direct and detailed characterization of bank influence on world countries via the sensitivity $D(b \to c, c)$ described in Section 2.

Thus in Fig. 4 we show the influence of Goldman Sachs on the world countries expressed by the sensitivity D presented on the world map. The strongest sensitivity is for Nigeria ($D = 1.03 \times 10^{-3}$) followed by Bangladesh, Vietnam, Denmark and S. Korea. For UK and USA we have $D = 6.7 \times 10^{-4}$ and the minimal sensitivity $D = 5.3 \times 10^{-4}$ is for Chad. On the world map we see also a significant sensitivity for Portugal, Egypt, Indonesia, Pakistan, Philippines, South Africa, Mexico taking positions from 6 to 12 from the most sensitive country of Nigeria. These countries belongs to the "Next Eleven" list marked by Goldman Sachs for countries with high macroeconomic stability, political maturity, openness of trade and investment policies and quality of education as criteria [36].

The world influence of Deutsche Bank is shown in Fig. 5. The highest sensitivity is for Libya ($D = 1.25 \times 10^{-3}$) even if there is no direct link from the wiki-article of Deutsche Bank to Libya. Thus this sensitivity appears due to indirect links between this country and Deutsche Bank. Libya is followed by Sudan and Slovakia. The world map shows that the influence of Deutsche Bank propagates mainly to East Europe, Turkey, Russia and China. We note that there is no direct link from Deutsche Bank to Russia, but there are links with VTB Bank in Russia that generates indirect links between Russia and Deutsche Bank. The lowest sensitivity is for East Timor ($D = 2.48 \times 10^{-4}$).

The world influence of ICB China is shown in Fig. 6. The most sensitive country is Luxembourg ($D = 4.75 \times 10^{-4}$) followed by Pakistan, Argentina, China, Japan, Djibouti. Indeed, in 2011 ICB China opened a branch in Luxembourg which became its European headquarters and its wiki-article has a direct link to Luxembourg. This article also directly points to Pakistan where ICB China established its branches. Argentina also has a direct link since ICB China acquired 80% of Standard Bank Argentina in 2012. The direct link to Japan results from the world's largest initial public offering (IPO) which surpassed the previous record from Japan. Djibouti has no direct links but the wiki-article of Djibouti points out that authorities have strengthened ties with China that generates certain indirect links. In global the world map of ICB China influence is marked by its propagation to Africa and Europe. The smallest sensitivity is for Chard ($D = 5.91 \times 10^{-5}$).

The sensitivity world maps and the sensitivity values for all 195 countries and all 60 banks are available at [26].

Above we presented the results only for English Wikipedia since it has the largest number of articles among all languages of Wikipedia. Thus in 2017 for English Wikipedia the number of articles is about 5 times larger compared to Chinese Wikipedia (see Table 1 in [22]). It is possible that English language provides a certain bias for USA and UK banks but in view of a sharp rank differences between USA/UK and China banks we expect that on average the advantage of USA/UK banks will be also present in other language editions. Also the international exchange between banks is mainly in English that supports the choice of English Wikipedia in these studies.

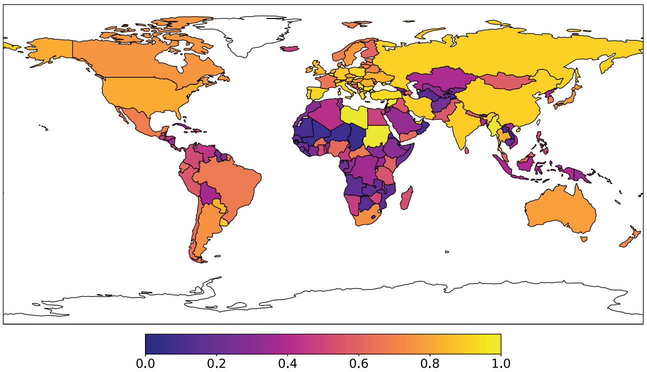

Fig. 5. Same as in Fig. 4 for sensitivity to Deutsche Bank with $D = 1.25 \times 10^{-3}$ (blue for zero) and $D = 2.48 \times 10^{-4}$ (yellow for unity).

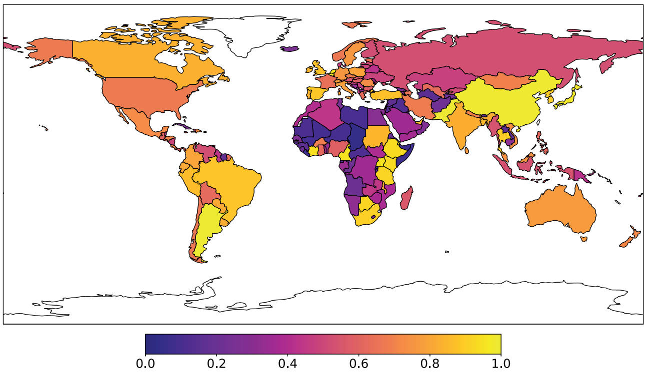

Fig. 6. Same as in Fig. 4 for sensitivity to ICB China with $D = 4.75 \times 10^{-4}$ (blue color for zero) and $D = 5.91 \times 10^{-5}$ (yellow color for unity).

7. Discussion

We performed a detailed analysis of interactions of the largest world banks [7,8] using the reduced Google matrix algorithm and the network of English Wikipedia with more than 5 million articles. While the top assets rank positions are occupied by the banks of China with ICB China at the first position the Wikipedia network analysis shows that the most influential banks at the PageRank top positions are US banks with the first one being Goldman Sachs bank. The performed REGOMAX analysis allows to establish direct and hidden interactions between the largest world banks and countries determining the closest friends and followers for each bank. The sensitivity analysis gives the world map of countries with their influence to a given bank. This geographical analysis demonstrates a clear tendency of banks expansions to other countries like e.g. to East European countries for Deutsche Bank, African countries, Pakistan and Argentina for ICB China. Since Wikipedia accumulates a huge amount of human knowledge we argue that the Wikipedia network analysis captures the real financial, social and historical interactions between world banks and countries.

We note that the financial networks of banks have a relatively small size. Thus the whole Federal Reserve has only about 6000 nodes [37] that is about 1000 times smaller compared than the Wikipedia network considered here. Thus we argue that the REGOMAX method will allow to perform an efficient analysis of such financial networks with new possibilities of crisis prevention.

We expect that the REGOMAX algorithm, developed from the physical problems of scattering [19], can become a useful tool for research in the field of econophysics [38].

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

This work was supported in part by the Programme Investissements d'Avenir, France ANR-11-IDEX-0002-02, reference ANR-10-LABX-0037-NEXT (project THETRACOM). This work was granted access to the HPC GPU resources of CALMIP

(Toulouse) under the allocation 2018-P0110. The development of the VexCL library was partially funded by the state assignment to the Joint supercomputer center of the Russian Academy of Sciences, Russia for scientific research.

References

[1] Financial crisis of 2007-2008, 2019, https://en.wikipedia.org/w/index.php?title=Financial_crisis_of_2007%E2%80%932008&oldid=882711856 (Accessed February (2019)).
[2] Three weeks that changed the world, The Guardian (2008) https://www.theguardian.com/business/2008/dec/28/markets-credit-crunch-banking-2008 (Accessed February (2019)).
[3] M.L. Bech, J.T.E. Chapman, R. Garratt, Which bank is the central bank?, J. Monet. Econ. 57 (2010) 352.
[4] G. Hale, Bank relationships, business cycles, and financial crises, J. Int. Econ. 88 (2012) 312.
[5] C. Minoiu, J.A. Reyes, A network analysis of global banking: 1978-2010, J. Fin. Stab. 9 (2013) 168.
[6] B. Craig, G. von Peter, Von peter g interbank tiering and money center banks, J. Finan. Intermediation 23 (3) (2014) 322.
[7] List of largest banks, wikipedia, 2018, https://en.wikipedia.org/w/index.php?title=List_of_largest_banks&oldid=871310902 (Accessed November (2018)).
[8] The World's 100 Largest Banks, S & P Global Market Intelligence report, 6, 2018, https://platform.mi.spglobal.com/web/client?auth=inherit#news/article?id=44027195&cdid=A-44027195-11060 (Accessed February (2019)).
[9] S. Brin, L. Page, The anatomy of a large-scale hypertextual web search engine, Comput. Netw. ISDN Syst. 30 (1998) 107.
[10] A.M. Langville, C.D. Meyer, Google's PageRank and beyond: The Science of Search Engine Rankings, Princeton University Press, Princeton, 2006.
[11] L. Ermann, K.M. Frahm, D.L. Shepelyansky, Google matrix analysis of directed networks, Rev. Modern Phys. 87 (2015) 1261.
[12] L. Ermann, D.L. Shepelyansky, Google matrix of the world trade network, Acta Phys. Polon. A 120 (2011) A158.
[13] L. Ermann, D.L. Shepelyansky, Google matrix analysis of the multiproduct world trade network, Eur. Phys. J. B 88 (2015) 84.
[14] J. Giles, Internet encyclopaedias go head to head, Nature 438 (2005) 900.
[15] J.M.Jr. Reagle, Good Faith Collaboration: The Culture of Wikipedia, in: History and Foundations of Information Science, MIT Press, 2012.
[16] F.A. Nielsen, Wikipedia research and tools: review and comments, SSRN Electron. J. (2012) Available https://doi.org/10.2139/ssrn.2129874 (Accessed February (2019)).
[17] W. Lewoniewski, K. Wecel, W. Abramowicz, Relative quality and popularity evaluation of Multilingual Wikipedia Articles, Informatics 4 (4) (2017) 43.
[18] K.M. Frahm, D.L. Shepelyansky, Wikipedia networks of 24 editions of 2017, 2017, http://www.quantware.ups-tlse.fr/QWLIB/24wiki2017 (Accessed February (2019)).
[19] K.M. Frahm, D.L. Shepelyansky, Reduced google matrix, 2016, arXiv:1602.02394 [physics.soc].
[20] K.M. Frahm, K. Jaffres-Runser, D.L. Shepelyansky, Wikipedia mining of hidden links between political leaders, Eur. Phys. J. B 89 (2016) 269.
[21] K.M. Frahm, S. El Zant, K. Jaffres-Runser, D.L. Shepelyansky, Multi-cultural wikipedia mining of geopolitics interactions leveraging reduced Google matrix analysis, Phys. Lett. A 381 (2017) 2677.
[22] C. Coquide, J. Lages, D.L. Shepelyansky, World influence and interactions of universities from Wikipedia networks, Eur. Phys. J. B 92 (2019) 3.
[23] J. Lages, D.L. Shepelyansky, A. Zinovyev, Inferring hidden causal relations between pathway members using reduced Google matrix of directed biological networks, PLoS One 13 (1) (2018) e0190812.
[24] M.C. Munnix, R. Schaefer, T. Guhr, A random matrix approach to credit risk, PLoS One 9 (5) (2014) e98030.
[25] M. Bardoscia, G. Livan, M. Marsili, Statistical mechanics of complex economies, J. Stat. Mech.: Theor. Exp. 2017 (2017) 043402.
[26] D. Demidov, K.M. Frahm, D.L. Shepelyansky, Wikipedia bank network, 2019, http://http://www.quantware.ups-tlse.fr/QWLIB/wikibanknet (Accessed February (2019)).
[27] A.D. Chepelianskii, Towards physical laws for software architecture, 2010, arXiv:1003.5455 [cs.SE].
[28] K.M. Frahm, D.L. Shepelyansky, REGOMAX, 2019, https://github.com/kmfrahm/regomax (Accessed February (2019)).
[29] D. Demidov, K. Ahnert, K. Rupp, P. Gottschling, Programming CUDA and OpenCL: A case study using modern C++ Libraries, SIAM J. Sci. Comput. 35 (5) (2013) C453.
[30] D. Demidov, VEXCL, 2018, https://github.com/ddemidov/vexcl (Accessed December (2018)).
[31] E. Cuthill, J. McKee, Reducing the bandwidth of sparse symmetric matrices, in: Proc. 24th Nat. Conf, ACM, 1969, p. 157.
[32] Olympe, Calmip, 2019, https://www.calmip.univ-toulouse.fr/spip.php?article582 (Accessed December (2018)).
[33] D. Demidov, REGOMAX-GPU, 2019, https://gitlabcom/ddemidov/regomax-gpu (Accessed February (2019)).
[34] G. Rollin, J. Lages, D.L. Shepelyansky, World influence of infectious diseases from wikipedia network analysis, IEEE Access (2018) http://dx.doi.org/10.1101/424465.
[35] S. Dorogovtsev, Lectures on Complex Networks, Oxford University Press, Oxford, 2010.
[36] Goldman sachs, 2019, https://en.wikipediaorg/wiki/Goldman_Sachs#cite_note-34 (Accessed February (2019)).
[37] K. Soramaki, M.L. Bech, J. Arnold, R.J. Glass, W.E. Beyeler, The topology of interbank payment flows, Physica A 379 (2007) 317.
[38] R.N. Mantegna, H.E. Stanley, Introduction to Econophysics: Correlations and Complexity in Finance, Cambridge University Press, Cambridge, 1999.