Capturing the influence of geopolitical ties from Wikipedia with reduced Google matrix

Samer El Zant[1], Katia Jaffrès-Runser[1]*, Dima L. Shepelyansky[2]

1 Institut de Recherche en Informatique de Toulouse / Toulouse INP, Université de Toulouse, Toulouse, France, **2** Laboratoire de Physique Théorique du CNRS/IRSAMC, Université de Toulouse, Toulouse, France

* kjr@enseeiht.fr

Abstract

Interactions between countries originate from diverse aspects such as geographic proximity, trade, socio-cultural habits, language, religions, etc. Geopolitics studies the influence of a country's geographic space on its political power and its relationships with other countries. This work reveals the potential of Wikipedia mining for geopolitical study. Actually, Wikipedia offers solid knowledge and strong correlations among countries by linking web pages together for different types of information (e.g. economical, historical, political, and many others). The major finding of this paper is to show that meaningful results on the influence of country ties can be extracted from the hyperlinked structure of Wikipedia. We leverage a novel stochastic matrix representation of Markov chains of complex directed networks called the reduced Google matrix theory. For a selected small size set of nodes, the reduced Google matrix concentrates direct and indirect links of the million-node sized Wikipedia network into a small Perron-Frobenius matrix keeping the PageRank probabilities of the global Wikipedia network. We perform a novel sensitivity analysis that leverages this reduced Google matrix to characterize the influence of relationships between countries from the global network. We apply this analysis to two chosen sets of countries (i.e. the set of 27 European Union countries and a set of 40 top worldwide countries). We show that with our sensitivity analysis we can exhibit easily very meaningful information on geopolitics from five different Wikipedia editions (English, Arabic, Russian, French and German).

Introduction

Relationships between countries have always been of utmost interest to study for countries themselves as they have to be accounted for into any country's strategic and diplomatic plan. Studies are driven by observing the influence of a relationship between two countries on other countries from different perspectives listing economic exchanges, social changes, history, politics, religious, martial, regional as seen in [1]. The major finding of this paper is to show that meaningful results on geopolitics interactions could be extracted from Wikipedia for a given selection of countries. Therefore, it can be leveraged to provide a picture of countries relationships offering a new framework for geopolitical studies. In [2], Sara Javanmardi et al. show that even though anyone can edit a Wikipedia entry at any time, the average article quality

increases as it goes through various edits. Wikipedia's accuracy for its scientific entries has been proved by comparing it to Encyclopedia Britannica and to PDQ—NCI's Comprehensive Database in [3, 4]. To sum up, Wikipedia has become the largest accurate reliable free online open source of knowledge.

Wikipedia is an interesting target domain for network analysts due to the hyperlinked structure that provides a direct relationship between web pages and topics. Research on such networks has derived content-independent effective metrics to rank nodes and edges of the graph based on their relevance to a given criteria (clustering, importance ranking, etc.). In this study we concentrate on one of the most popular network analysis algorithms: the PageRank algorithm [5, 6]. PageRank algorithm could be seen as a Markov chain process relying on the definition of the so-called Google Matrix which describes the network interconnection. For various language editions of Wikipedia it has been shown that the PageRank vector produces a reliable ranking of historical figures over 35 centuries of human history [7–11] and a solid Wikipedia ranking of world universities (WRWU) [7, 12]. It has been shown as well that the Wikipedia ranking of historical figures is in a good agreement with the well-known Hart ranking [13], while the WRWU is in a good agreement with the Shanghai Academic ranking of world universities [14].

This paper analyses the networks extracted from 5 language editions of Wikipedia to study the influence of countries on each other. We proceed with this analysis for two sets of countries: *i*) the 27 member states of the European Union and *ii*) the top 40 countries according to English Wikipedia PageRank. For each Wikipedia language edition, we build a standard Wikipedia network representation as follows. Each webpage in Wikipedia is related to a clearly defined topic. On each page, there are hyperlinks pointing to other webpages of the same Wikipedia edition that are related to the topic of interest. As such, webpages are interconnected through directed links (i.e. hyperlinks), creating network of webpages. It is common to model this network as a directed graph where vertices represent all webpages and oriented edges represent the hyperlinks. This graph is complex as it can hold up to several millions of vertices and about ten times more edges.

This Wikipedia network graph models the direct links between topics. However, indirect links exist as well as two topics can be related by intermediary webpages. For instance, the webpage of France is indirectly related to Latvia because it has a direct link to the *Environmental Performance Index* webpage, that contains a link to Latvia. To correctly determine the interaction between the countries of interest, both direct and indirect interactions must be accounted for in our study.

A solution that captures the full contribution of direct and indirect interactions within a single stochastic matrix representation of the network of webpages has been proposed in [15]. This solution relies on the Google Matrix representation [5, 6] of the Wikipedia network of N nodes. Knowing a selection of N_r nodes, $N_r \ll N$, it calculates a *reduced Google matrix*. In this paper, the N_r nodes are the Wikipedia webpages whose topics are the countries of interest (e. g. we have https://en.wikipedia.org/wiki/France for France, https://en.wikipedia.org/wiki/United_States for US, etc.). The reduced Google matrix G_R is a N_r-by-N_r Perron-Frobenius matrix where each element $G_R(i, j)$ represents the probability that node i points onto node j using direct and indirect links in the complete Wikipedia network. Moreover, the reduced Google matrix theory offers a matrix decomposition of G_R that can be leveraged to distinguish the contribution of direct and indirect interactions from the overall $G_R(i, j)$ probability. Up to a constant multiplier, the PageRank probabilities of G_R are the same as for the Google matrix G of the global Wikipedia network.

Reduced Google matrix theory has been successfully leveraged in [16] and in [17]. Results in [16] highlight meaningful interactions between groups of political leaders from the

Wikipedia networks. Most relevant to the study presented in this paper, the work in [17] shows that reduced Google matrix is a perfect candidate for analyzing the geopolitics interactions between countries selected worldwide for 5 different Wikipedia language editions for two reasons: 1) Indirect interactions components of G_R capture reasonable and relevant information about hidden relationships between countries identified as hidden friends and followers 2) Part of the interactions are cross-cultural while others are clearly biased by the culture of the authors. This work has assessed the validity of the reduced Google matrix approach for the study of geopolitical interactions. It has extracted meaningful pieces of information from the intrinsic structure of the Wikipedia network by revealing the existence of indirect relationships between countries.

The work of this paper goes one step further as it quantifies the influence of a relationship between two countries on the rest of the reduced network using G_R. Previous work has identified the strongest ties, but this one focuses on capturing the impact of a change in the strength of a relationship between two countries on the overall network interactions of selected countries via the global network. The impact on the overall network structure is measured by calculating the variation of importance of the nodes in the network. We show that this sensitivity analysis renders a reasonable and meaningful idea of the influence of a given bilateral tie on the whole network.

More specifically, in this paper, we calculate G_R for the two groups of 27 EU and 40 world countries each. Thus, G_R reflects in a 40-by-40 or 27-by-27 matrix the complete (direct and indirect) relationships between countries. To identify the relative influence of one relationship between two nations, we propose in this paper to compute a logarithmic derivative of the PageRank probabilities calculated from G_R and $\tilde{G}_R$. PageRank probabilities are derived from G_R as explained later. They represent the importance of a node in the network. $\tilde{G}_R$ is almost equal to G_R. It only differs by the values of one column. If the relationship going from nation j to nation i is of interest in the study, only the values of column j are changed to relatively inflate the probability $\tilde{G}_R(i, j)$ of nation j ending in nation i compared to the other ones. This is done in practice by increasing $\tilde{G}_R(i, j)$ and then normalizing the column again to unity as it required by the definition of the Google matrix.

From our sensitivity analysis on both sets of countries, we extract reasonable and really interesting geopolitical influences. Indeed, for instance in the set of 27 EU countries, our data shows clearly that the Nordic group of nations (Sweden, Denmark, Finland) have strong relationships together. If one of them increases its ties to another EU country alone, the remaining ones see their importance drop. The same observation is made for the group created by Austria, Hungary and Slovenia nations. These observations have been made by geopolitical specialists as well in [18] and [19], respectively. Another striking result is the impact of the exit of Great Britain from EU on the other European countries. Our data shows that Ireland will be the most affected country, which is inline with a study delivered recently by the London School of Economics [20]. From our worldwide set of 40 countries, we show that strengthening the relationship between Russia and the United States of America would negatively impact the importance of Ukraine worldwide, which is identical to the interpretation represented by Francis Fukuyama in a recent article [21].

The paper is constructed as follows. At first we introduce the reduced Google matrix theory, together with a primer on Google Matrix and PageRank calculations. The reduced Google matrix is illustrated for both sets of 27 EU and 40 nations. Next, the methodology for our link sensitivity analysis is presented. A detailed analysis for the two groups of countries is given in the Results section that focuses on the sensitivity analysis of important relationships in the

group. Results are first given and discussed for the set of 27 EU countries and then for the set of 40 worldwide nations. Finally, conclusions are drawn in the last section.

Google matrix analysis of Wikipedia networks

Data description

Our study focuses on the networks representing 5 different Wikipedia editions (data collected mid February 2013) from the set of 24 analyzed in [10] [22]: EnWiki, ArWiki, RuWiki, DeWiki and FrWiki that contain 4.212, 0.203, 0.966, 1.533 and 1.353 millions of articles each. The selected countries are the 27 EU countries as of February 2013 (Croatia joined later, in July 2013) and the 40 countries selected from the EnWiki network as the top 40 countries of the PageRank probability for the complete network.

Countries that belong to the same region or having a common piece of history may probably exhibit stronger interactions in Wikipedia. For the set of 40 countries, we have created a color code that groups together countries that either belong to the same geographical region (e.g. Europe, South America, Middle East, North-East Asia, South-East Asia) or share a big part of history (former USSR; English speaking countries that are the legacy of the former British Empire) [17]. On the other hand, EU countries are grouped upon their accession date to the union (e.g. Founder, 1973, 1981-1986, 1995, 2004-2007). Color code for EU countries can be seen in Fig 1. Color code for the worldwide set of 40 countries is available in Table 1.
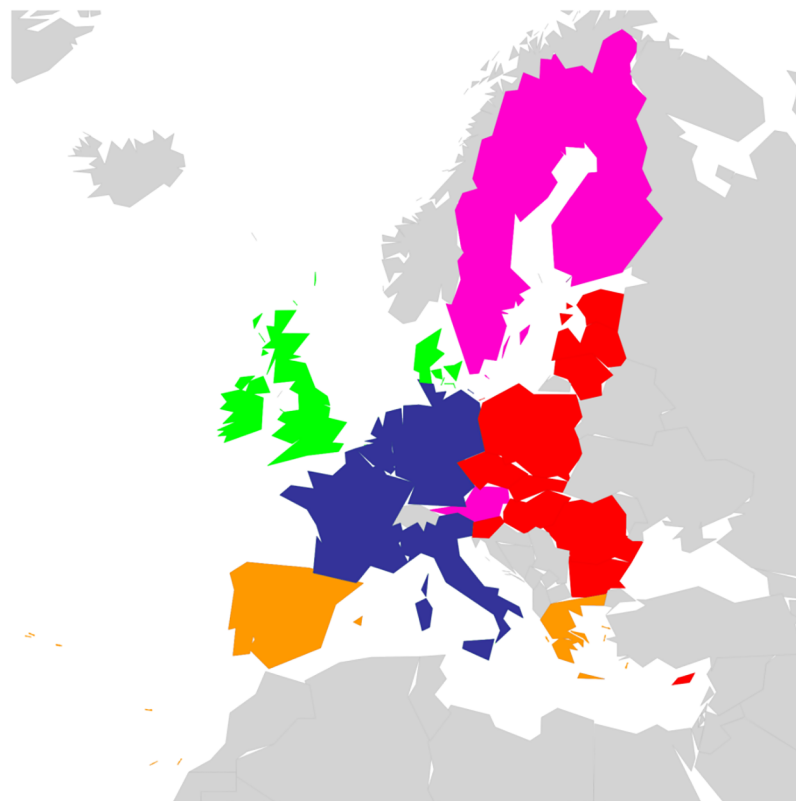

Fig 1. Geographical distribution of the EU countries. Color code groups countries into 5 subsets: Blue (BL) for Founders, Green (GN) for 1973 new member states, Orange (OR) for 1981 to 1986 new member states, Pink (PK) for 1995 new member states and Red (RD) for 2004 to 2007 new member states.

https://doi.org/10.1371/journal.pone.0201397.g001

Table 1. List of 40 selected countries.

Wikipedia edition			English		Arabic		Russian	
Countries	CC	Color	K	K^*	K	K^*	K	K^*
United States	US	OR	1	9	1	5	2	27
France	FR	RD	2	19	3	31	3	14
United Kingdom	GB	OR	3	25	6	20	7	3
Germany	DE	RD	4	33	8	14	4	24
Canada	CA	OR	5	26	13	19	12	26
India	IN	PK	6	23	9	25	13	8
Australia	AU	OR	7	35	16	22	18	12
Italy	IT	RD	8	15	5	1	6	32
Japan	JP	VT	9	4	11	9	11	7
China	CN	VT	10	8	12	17	9	21
Russia	RU	BL	11	6	7	2	1	2
Spain	ES	RD	12	30	4	8	8	15
Poland	PL	RD	13	12	26	32	10	17
Netherlands	NL	RD	14	37	18	33	15	31
Iran	IR	YL	15	2	14	15	30	22
Brazil	BR	GN	16	3	21	26	20	1
Sweden	SE	RD	17	22	22	7	19	5
New Zealand	NZ	OR	18	28	34	24	36	4
Mexico	MX	GN	19	40	23	38	22	37
Switzerland	CH	RD	20	38	20	34	16	18
Norway	NO	RD	21	32	35	16	27	11
Romania	RO	RD	22	10	19	6	32	36
Turkey	TR	YL	23	7	15	13	21	38
South Africa	ZA	OR	24	24	29	39	35	20
Belgium	BE	RD	25	18	27	37	29	30
Austria	AT	RD	26	39	28	28	14	28
Greece	GR	RD	27	21	10	36	25	25
Argentina	AR	GN	28	1	32	29	33	23
Philippines	PH	PK	29	17	36	21	39	33
Portugal	PT	RD	30	36	24	12	17	9
Pakistan	PK	PK	31	5	25	35	37	29
Denmark	DK	RD	32	16	33	10	31	19
Israel	IL	YL	33	20	17	18	28	6
Finland	FI	RD	34	14	37	4	26	16
Egypt	EG	YL	35	31	2	3	24	39
Indonesia	ID	PK	36	13	31	11	34	10
Hungary	HU	RD	37	11	40	40	23	40
Taiwan	TW	VT	38	27	39	27	40	34
South Korea	KR	VT	39	34	38	30	38	35
Ukraine	UA	BL	40	29	30	23	5	13

PageRank K and CheiRank K^* for EnWiki, FrWiki and RuWiki. Color code (CC) groups countries into 7 subsets: orange (OR) for English speaking countries, Blue (BL) for former Soviet union ones, Red (RD) for European ones, Green (GN) for South American ones, Yellow (YL) for Middle Eastern ones, Purple (VT) for North-East Asian ones and finally Pink (PK) for South-Eastern countries.

Google matrix, PageRank and CheiRank

It is convenient to describe the network of N Wikipedia articles by the Google matrix G constructed from the adjacency matrix A_{ij} with elements 1 if article (node) j points to article (node) i and zero otherwise. In this case, elements of the Google matrix take the standard form [5, 6]

$$G_{ij} = \alpha S_{ij} + (1 - \alpha)/N \quad , \tag{1}$$

where S is the matrix of Markov transitions with elements $S_{ij} = A_{ij}/k_{out}(j)$, $k_{out}(j) = \sum_{i=1}^{N} A_{ij} \neq 0$ being the node j out-degree (number of outgoing links) and with $S_{ij} = 1/N$ if j has no outgoing links (dangling node). Here $0 < \alpha < 1$ is the damping factor which for a random surfer determines the probability $(1 - \alpha)$ to jump to any node; below we use $\alpha = 0.85$. Element G_{ij} represents the probability of a random surfer to go from node j to node i. The right eigenvectors $\psi_i(j)$ of G are defined by:

$$\sum_{j'} G_{jj'} \psi_i(j') = \lambda_i \psi_i(j) \quad . \tag{2}$$

The PageRank eigenvector $P(j) = \psi_{i=0}(j)$ corresponds to the largest eigenvalue $\lambda_{i=0} = 1$ [5, 6]. It has positive elements which give the probability to find a random surfer on a given node in the stationary long time limit of the Markov process. All nodes can be ordered by a monotonically decreasing probability $P(K_i)$ with the highest probability at $K = 1$. The index K is the PageRank index. Left eigenvectors are biorthogonal to right eigenvectors of different eigenvalues. The left eigenvector for $\lambda = 1$ has identical (unit) entries due to the column sum normalization of G. In the following we use the notations ψ_L^T and ψ_R for left and right eigenvectors, respectively. Notation T stands for vector or matrix transposition.

In addition to the matrix G it is useful to introduce a Google matrix G^* constructed from the adjacency matrix of the same network but with inverted direction of all links [23]. The vector $P^*(K^*)$ is called the CheiRank vector [7, 23], with K^* being the CheiRank index obtained after numbering nodes in monotonic decrease of probability P^*. Thus, nodes with many ingoing (or outgoing) links have small values of $K = 1, 2, 3 \ldots$ (or of $K^* = 1, 2, 3, \ldots$) [6, 11].

The two sets of 27 EU and 40 world countries are listed in Tables 1 and 2 respectively. The set of 40 countries has been chosen by selecting the countries with the largest PageRank probabilities in the full EnWiki network. In Tables 1 and 2, a *local PageRank index K* is given whose values range between 1 and 27 for EU countries, and between 1 and 40 for the other set. This local ranking keeps the countries in the same sequence as the original ranking over the entire network of webpages. The most influential countries are the top ranked ones with $K = 1, 2, \ldots$. Similarly, the local CheiRank index K^* [11, 23] is given in both Tables for the two sets. At the top of K^* we have the most communicative countries. Both local K and K^* are given for EnWiki, ArWiki and RuWiki. Not surprisingly, the order of top countries changes with respect to the edition (for instance, the top country for K is US except for RuWiki whose top country is Russia).

It is convenient as well to plot all nodes in the (K, K^*) plane to highlight the countries that are the most influential ($K = 1, 2, \ldots$) and the most communicative ($K^* = 1, 2, \ldots$) at the same time. Fig 2 plots EU countries in the (K, K^*) plane for EnWiki, FrWiki and DeWiki editions. This plot is a bi-objective plot where K and K^* are to be minimized concurrently. It is interesting to look at the set of non-dominated countries which are the ones such that there is no other country beating them for both K and K^*.

Table 2. List of EU countries.

Wikipedia edition			English		French		German	
Countries	CC	Color	K	K*	K	K*	K	K*
France	FR	BL	1	10	1	6	2	9
United Kingdom	GB	GN	2	14	4	13	24	27
Germany	DE	BL	3	20	2	7	1	1
Italy	IT	BL	4	6	3	9	4	14
Spain	ES	OR	5	19	5	17	5	15
Poland	PL	RD	6	3	8	5	6	6
Netherlands	NL	BL	7	25	7	12	7	21
Sweden	SE	PK	8	13	11	25	8	18
Romania	RO	RD	9	1	18	4	17	20
Belgium	BE	BL	10	9	6	1	9	4
Austria	AT	PK	11	27	9	23	3	3
Greece	GR	OR	12	11	13	10	14	8
Portugal	PT	OR	13	24	12	2	11	2
Ireland	IE	GN	14	16	19	14	16	26
Denmark	DK	GN	15	7	14	20	10	10
Finland	FI	PK	16	4	17	18	15	7
Hungary	HU	RD	17	2	10	3	13	12
Czech Republic	CZ	RD	18	5	15	24	12	17
Bulgaria	BG	RD	19	22	20	11	20	13
Estonia	EE	RD	20	8	24	15	22	23
Slovenia	SI	RD	21	18	23	21	23	22
Slovakia	SK	RD	22	12	16	8	18	5
Lithuania	LT	RD	23	21	22	27	21	19
Cyprus	CY	RD	24	17	27	26	27	25
Latvia	LV	RD	25	23	25	22	25	24
Luxembourg	LU	BL	26	26	21	19	19	11
Malta	MT	RD	27	15	26	16	26	16

PageRank K and CheiRank K^* for EnWiki, FrWiki and DeWiki. Fig 1 gives color correspondence details. Color code groups countries into 5 subsets: Blue (BL) for Founders, Green (GN) for 1973 new member states, Orange (OR) for 1981 to 1986 new member states, Pink (PK) for 1995 new member states and Red (RD) for 2004 to 2007 new member states. Standard country codes (CC) are given as well.
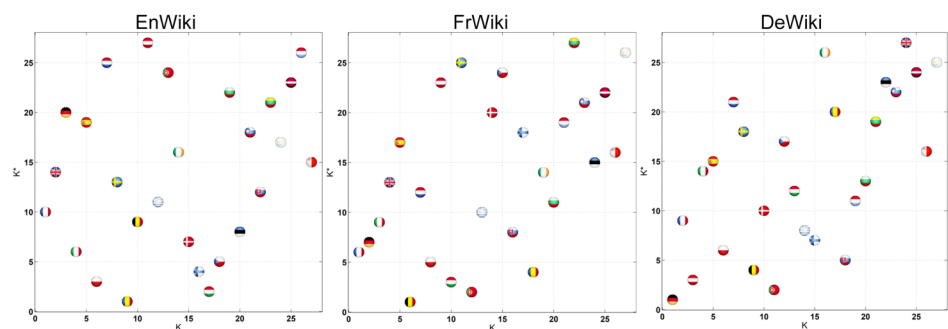
Fig 2. Position of EU countries in the local (K, K^*) plane. EnWiki (left), FrWiki (middle) and DeWiki (right) networks. Countries are marked by their flags.

To summarize, PageRank and CheiRank capture the relative importance of nodes in the network. They are extracted from the Google matrix representation of the network of web-pages. The Google matrix lists for each link the probability for directly transitioning from one webpage to the other one. The PagerRank probability $P(K)$ represents the probability of ending on a webpage, eventually. In the following, we introduce the Reduced Google Matrix that offers a complementary analysis that extracts the importance of the indirect interactions between a set of nodes of the original network.

Reduced Google matrix analysis

Let G be the typical Google matrix of Eq (1) for a network of N nodes such that $G_{ij} \geq 0$ and the column sum normalization $\sum_{i=1}^{N} G_{ij} = 1$ is verified for each column. We consider a sub-network with $N_r < N$ nodes, called "reduced network". In this case we can write G in a block form:

$$G = \begin{pmatrix} G_{rr} & G_{rs} \\ G_{sr} & G_{ss} \end{pmatrix} \tag{3}$$

where the index "r" refers to the nodes of the reduced network and "s" to the other $N_s = N - N_r$ nodes which form a complementary network which we will call the "scattering network". PageRank vector of the full network is given by:

$$P = \begin{pmatrix} P_r \\ P_s \end{pmatrix} \tag{4}$$

which satisfies the equation $GP = P$ or in other words P is the right eigenvector of G for the unit eigenvalue. This eigenvalue equation reads in block notations:

$$(\mathbf{1} - G_{rr}) P_r - G_{rs} P_s \quad = \quad 0, \tag{5}$$

$$-G_{sr} P_r + (\mathbf{1} - G_{ss}) P_s \quad = \quad 0. \tag{6}$$

Here $\mathbf{1}$ is the unit matrix of corresponding size N_r or N_s. Assuming that the matrix $\mathbf{1} - G_{ss}$ is not singular, i.e. all eigenvalues G_{ss} are strictly smaller than unity (in modulus), we obtain from Eq (6) that

$$P_s = (\mathbf{1} - G_{ss})^{-1} G_{sr} P_r \tag{7}$$

which gives together with (5):

$$G_R P_r = P_r \quad , \quad G_R = G_{rr} + G_{rs} (\mathbf{1} - G_{ss})^{-1} G_{sr} \tag{8}$$

where the matrix G_R of size $N_r \times N_r$, defined for the reduced network, can be viewed as an effective reduced Google matrix. Here the contribution of G_{rr} accounts for direct links in the reduced network and the second matrix inverse term corresponds to all contributions of indirect links of arbitrary order. The matrix elements of G_R are non-negative since the matrix inverse in Eq (8) can be expanded as:

$$(\mathbf{1} - G_{ss})^{-1} = \sum_{l=0}^{\infty} G_{ss}^l \ . \tag{9}$$

In Eq (9) the integer l represents the order of indirect links, i. e. the number of indirect links which are used to connect indirectly two nodes of the reduced network. We refer the reader to

[16] to get the proof that G_R also fulfills the condition of column sum normalization being unity.

Numerical evaluation of G_R

We can question how to evaluate practically the expression of Eq (8) of G_R for a particular sparse and quite large network when $N_r \sim 10^2\text{-}10^3$ is small compared to N and $N_s \approx N \gg N_r$. If N_s is too large (e. g. $N_s > 10^5$) a direct naive evaluation of the matrix inverse $(\mathbf{1} - G_{ss})^{-1}$ in Eq (8) by Gauss algorithm is not efficient. In this case we can try the expansion of Eq (9) provided it converges sufficiently fast with a modest number of terms. However, this is most likely not the case for typical applications since G_{ss} is very likely to have at least one eigenvalue very close to unity.

Therefore, we consider the situation where the full Google matrix has a well defined gap between the leading unit eigenvalue and the second largest eigenvalue (in modulus). For example if G is defined using a damping factor α in the standard way, as in Eq (1), the gap is at least $1 - \alpha$ which is 0.15 for the standard choice $\alpha = 0.85$ [6]. In order to evaluate the expansion of Eq (9) efficiently, we need to take out analytically the contribution of the leading eigenvalue of G_{ss} close to unity which is responsible for the slow convergence.

Below we denote by λ_c this leading eigenvalue of G_{ss} and by ψ_R (ψ_L^T) the corresponding right (left) eigenvector such that $G_{ss} \psi_R = \lambda_c \psi_R$ (or $\psi_L^T G_{ss} = \lambda_c \psi_L^T$). Both left and right eigenvectors as well as λ_c can be efficiently computed by the power iteration method in a similar way as the standard PageRank method. Vectors ψ_R are normalized with $E_s^T \psi_R = 1$ and ψ_L with $\psi_L^T \psi_R = 1$. It is well known (and easy to show) that ψ_L^T is orthogonal to all other right eigenvectors (and ψ_R is orthogonal to all other left eigenvectors) of G_{ss} with eigenvalues different from λ_c. We introduce the operator $\mathcal{P}_c = \psi_R \psi_L^T$ which is the projector onto the eigenspace of λ_c and we denote by $\mathcal{Q}_c = \mathbf{1} - \mathcal{P}_c$ the complementary projector. One verifies directly that both projectors commute with the matrix G_{ss} and in particular $\mathcal{P}_c G_{ss} = G_{ss} \mathcal{P}_c = \lambda_c \mathcal{P}_c$. Therefore we can derive:

$$(\mathbf{1} - G_{ss})^{-1} \quad = \quad \mathcal{P}_c \frac{1}{1 - \lambda_c} + \mathcal{Q}_c \sum_{l=0}^{\infty} \bar{G}_{ss}^l \qquad (10)$$

with $\bar{G}_{ss} = \mathcal{Q}_c G_{ss} \mathcal{Q}_c$ and using the standard identity $\mathcal{P}_c \mathcal{Q}_c = 0$ for complementary projectors. The expansion in Eq (10) converges rapidly since $\bar{G}_{ss}^l \sim |\lambda_{c,2}|^l$ with $\lambda_{c,2}$ being the second largest eigenvalue which is significantly lower than unity.

The combination of Eqs (8) and (10) provides an explicit algorithm feasible for a numerical implementation for modest values of N_r, large values of N_s and of course if sparse matrices G, G_{ss} are considered. We refer the reader to [16] for more advanced implementation considerations.

Decomposition of G_R

On the basis of Eqs (8)–(10), the reduced Google matrix can be presented as a sum of three components:

$$G_R = G_{rr} + G_{pr} + G_{qr}, \qquad (11)$$

with the first component G_{rr} given by direct matrix elements of G among the selected N_r

nodes. The second projector component G_{pr} is given by:

$$G_{pr} = G_{rs} \mathcal{P}_c G_{sr}/(1 - \lambda_c), \quad \mathcal{P}_c = \psi_R \psi_L^T .$$ (12)

The third component G_{qr} is of particular interest in this study as it characterizes the impact of indirect or hidden links. It is given by:

$$G_{qr} = G_{rs}[\mathcal{Q}_c \sum_{l=0}^{\infty} \bar{G}_{ss}^{\,l}]G_{sr}, \quad \mathcal{Q}_c = \mathbf{1} - \mathcal{P}_c, \quad \bar{G}_{ss} = \mathcal{Q}_c G_{ss} \mathcal{Q}_c.$$ (13)

We characterize the strength of these 3 components by their respective weights W_{rr}, W_{pr}, W_{qr} given respectively by the sum of all matrix elements of G_{rr}, G_{pr}, G_{qr} divided by N_r. By definition we have $W_{rr} + W_{pr} + W_{qr} = 1$.

Results: G_R properties

Reduced Google matrix of country networks

Reduced Google matrix has been computed, together with its components G_{rr}, G_{pr} and G_{qr}, for the English language edition of Wikipedia (EnWiki) and for the 2 selected sets of 27 and 40 countries listed in Tables 1 and 2. We recall that the 40 countries in the first set are the ones with top PageRank K in the network of EnWiki. Countries are ordered by increasing K value in all subsequent matrix representations. The weight of the three matrix components of G_R are listed in Table 3. Predominant component is clearly G_{pr} but as we will explain next, it is not the most meaningful.

The meaning of G_{rr} is clear as it is directly extracted from the global Google matrix G. It gives the direct links between the selected nodes and more specifically the probability $G_{rr}(i, j)$ for the surfer to go directly from column j country to line i country.

The sum of G_{pr} and G_{qr} represents the contribution of all indirect links through the scattering matrix G_{ss}. The projector component G_{pr} is rather close to nearly identical columns given by the PageRank probabilities of N_r nodes (see Fig 3(B)). Fig 3 shows the matrix density plots for G_R and G_{pr} for the 27 EU countries where lines and columns are ordered by increasing K values. For both matrices, column values are proportional to their PageRank probabilities. As detailed in [16], we observe numerically that $G_{pr} \approx P_r E_r^T/(1 - \lambda_c)$, meaning that each column is close to the normalized vector $P_r/(1 - \lambda_c)$. As such, G_{pr} transposes essentially in G_R the contribution of the first eigenvector of G. We can conclude that even if the overall column sums of G_{pr} account for $\sim$ 95-97% of the total column sum of G_R, G_{pr} doesn't offer innovative information compared to the legacy PageRank analysis.

A way more interesting contribution is the one of G_{qr}. This matrix captures higher-order indirect links between the N_r nodes due to their interactions with the global network environment. We will refer to these links as *hidden links*. We note that G_{qr} is composed of two parts $G_{qr} = G_{qrd} + G_{qrnd}$ where the first diagonal term-only matrix G_{qrd} represents the probabilities to stay on the same node during multiple iterations of $\bar{G}_{ss}$ in (13) while the second matrix only captures non-diagonal terms in G_{qrnd}. As such, G_{qrnd} represents indirect (hidden) links between the N_r nodes appearing via the global network. We note that a few matrix elements of

Table 3. Weights of the three matrices components of G_R.

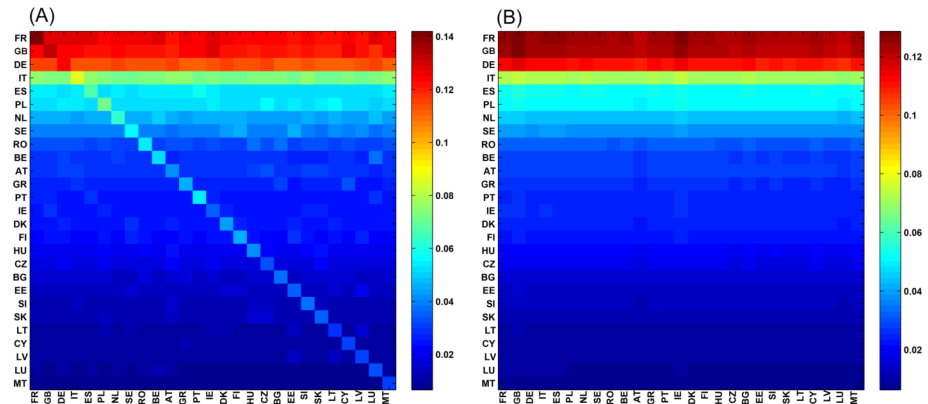|	W_{pr}	W_{qr}	W_{rr}	Sum
40	0.96120	0.029702	0.009098	1
EU	0.95332	0.038346	0.008334	1

Fig 3. Density plots of matrices G_R and G_{pr} extracted from EnWiki for 27 EU countries. (A): G_R, (B): G_{pr}. The colors represent maximum (red), intermediate (green) and minimum (blue) values.

G_{qr} can be negative, which is possible due to the negative terms in $\mathcal{Q}_c = \mathbf{1} - \mathcal{P}_c$ appearing in (13). The total weight of negative elements is however much smaller than W_{qr} (at least 6 times smaller and even non-existing in ArWiki for the network of 40 countries).

For the three EnWiki, FrWiki and DeWiki editions, Fig 4 plots the density of matrices G_R, G_{qrnd} and G_{rr}. We keep for all plots the same order of countries extracted from the EnWiki network. This is meant to highlight cultural differences among Wikipedia editions. From the first line of Fig 4, it is clear that G_R matrix is dominated by the projector G_{pr} contribution, which is proportional to the global PageRank probabilities. Several cultural biases can be extracted from G_R. For instance, France is the top country in EnWiki and FrWiki, while Germany is the top country in DeWiki.

The information from hidden links between countries is provided by G_{qrnd}. It shows, for the three selected languages editions, the strong hidden links connecting Finland to Sweden. Other interesting hidden links are between Ireland and United Kingdom in DeWiki or in EnWiki linking Luxembourg to France. The reduced Google matrix density plots for the network of 40 worldwide countries are to be found in reference [17].

Networks of friends and followers

As proposed in [17], it is possible to extract from G_R and G_{qrnd} a network of friends and followers to easily illustrate direct and hidden links in the network. Direct links are given by G_{rr} while hidden (i. e. indirect) are given by G_{qrnd}. For the sake of simplicity, we refer next to G_{qrnd} using G_{qr} notation.

To create these networks of friends and followers, we divide the set of N_r nodes into representative groups as shown in Fig 1 for 27 EU country set. EU countries are grouped upon their accession date to the union (e.g. Founder, 1973, 1981-1986, 1995, 2004-2007). One leading country per EU member state group has been selected as follows:

- France for Founders,

- United Kingdom for countries having joined in 1973,

- Spain for countries having joined between 1981 and 1986,

- Sweden for countries having joined in 1995,

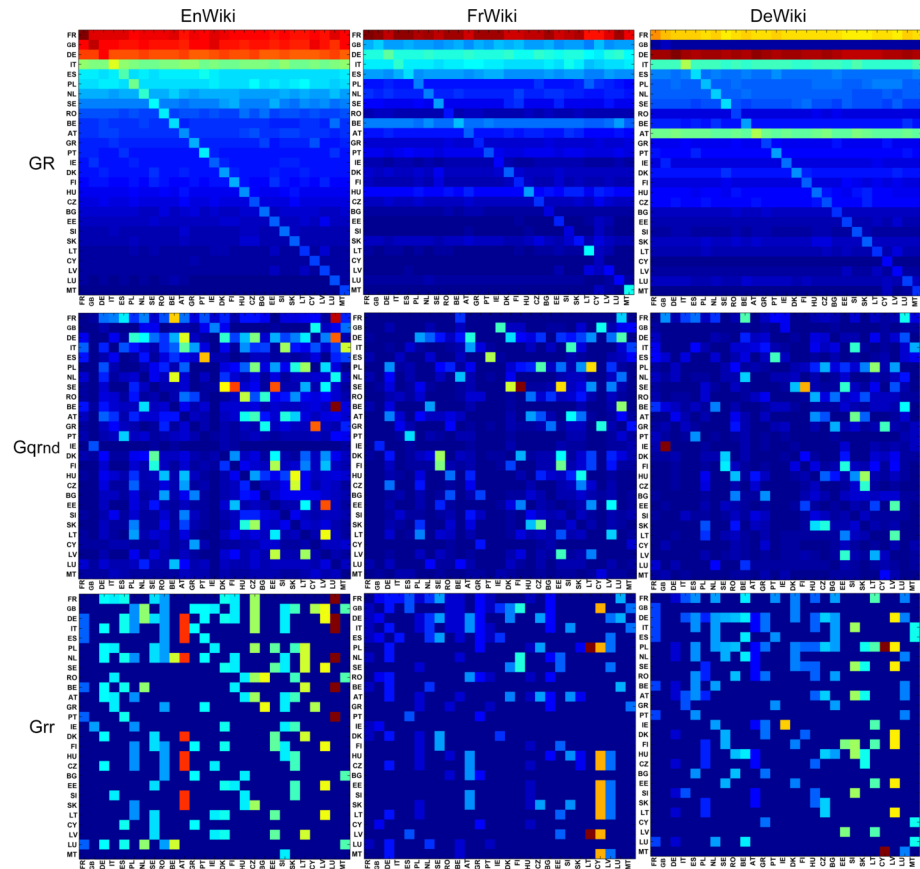- Poland for countries having joined between 2004 and 2007.

Fig 4. Density plots of G_R, G_{qrnd} and G_{rr}. G_R (first line), G_{qrnd} (second line) and G_{rr} (third line) for the reduced network of EU countries of EnWiki (left column), FrWiki (middle column) and DeWiki (right column). The nodes N_r are ordered in lines and columns by the reference PageRank of EnWiki. The colors represent maximum (red), intermediate (green) and minimum (blue) values.

https://doi.org/10.1371/journal.pone.0201397.g004

For each leading country j, we extract from both matrices G_{qr} and G_R the top 4 *Friends* (resp. *Followers*) given by the 4 best values of the elements of column j (resp. of line j). In other words, it corresponds to destinations of the 4 strongest outgoing links of j and the countries at the origin of the 4 strongest ingoing links of j. It is true that the word *friend* usually represents a symmetrical relationship. But we have chosen this denomination for its ease of use. Clearly, in this paper, a friend represents a node that is an attractor for the node of interest. These networks of top 4 friends and followers have been calculated for the five editions of Wikipedia.

Top 4 friends and top 4 followers of EU leading countries are extracted from G_R and G_{qr} to plot the graphs of Figs 5 and 6. Results for EnWiki, FrWiki and DeWiki are presented here. Note that Fig 6 pictures hidden links. The black thick arrows identify the top 4 friends and top 4 followers interactions. Red arrows represent the friends of friends (respectively the followers of followers) interactions that are computed recursively until no new edge is added to the graph. All graphs are visualized with the Yifan Hu layout algorithm [24] using Gephi [25].

The vertices of the network of friends obtained from G_R concentrate, for each Wiki, to about 7 countries, 5 of which being the leading ones. The other vertices are top PageRank countries such as Italy, Germany or Spain. This is due to the predominance of PageRank probabilities in the structure of G_R. A more valuable information could be extracted from the
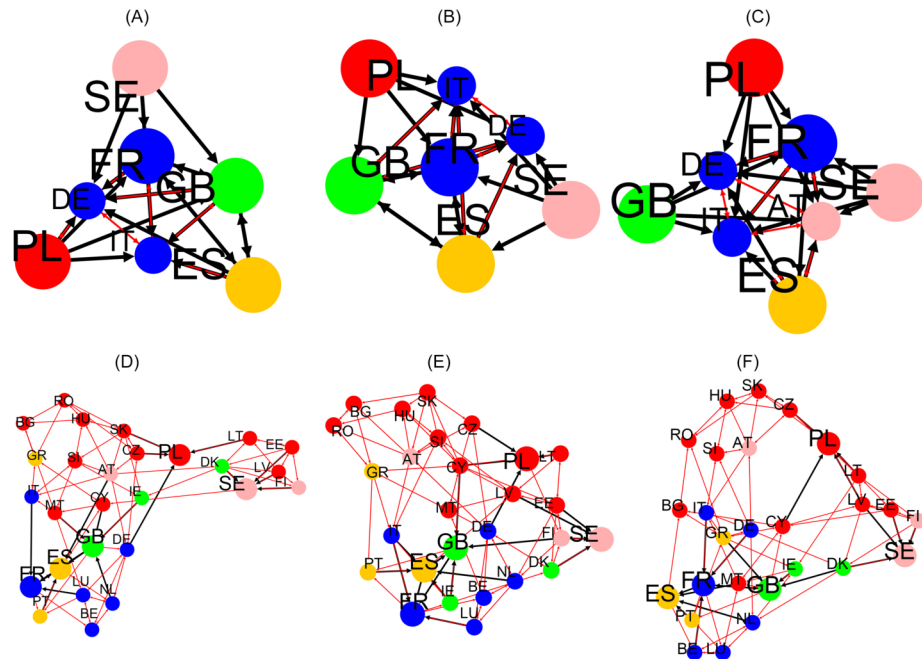
Fig 5. Relationships structure extracted from G_R for the network of EU countries. Friends (top line) and followers (bottom line) induced by the 5 top countries of each group (FR, GB, ES, SE, PL). Results are plotted for EnWiki (A and D), FrWiki (B and E) and DeWiki (C and F). Node colors represent geographic appartenance to a group of countries (cf. Fig 1 and Table 2 for details). Top (bottom) graphs: a country node with higher PageRank probability has a bigger size and points (is pointed by) with a bold black arrow to its top 4 friends (followers). Red arrows show friends of friends (resp. followers of followers) interactions computed until no new edges are added to the graph. Drawn with Yifan Hu algorithm [24] provided by the Gephi software [25].

network of followers. In all editions, Benelux and Nordic countries create a cluster densely interconnected. The networks of followers end up spanning the full set of EU countries in this representation. On this representation, it can be noticed that the order of arrival of member states is meaningful. Indeed, nodes of the same color are closely interconnected.

The hidden friends and followers relationships are extracted from G_{qr} and illustrated in Fig 6. As discussed earlier, G_{qr} is not dominated by PageRank, and as such, the resulting network of friends includes more nodes and shows more diversity. It is worth noting that Germany, as one of the Founders, bridges the group of Founders to Sweden (the leader of the countries that have joined EU in 1995) and Poland (the leader of the countries that have joined EU between 2004 and 2007) in FrWiki and EnWiki. From EnWiki and DeWiki, strong ties are seen between Italy and France, while it is not the case from FrWiki authors. This is another example of cultural bias. However, lots of links are to seen in all three editions: GB-IE, SE-FI, ES-PT, PL-LT, IT-GR and many others. To underline this constant presence of links, we give in Table 4 the list of friends (resp. followers) that are among the top 4 ones in all 5 editions, in 4 out of 5 and in 3 out of 5 for G_{qr} analysis. For each leading country, around 2 to 3 top friends and followers exist accross all editions.

For the 40 worldwide countries set, networks of top 4 friends and followers are to be found in [17], calculated for the same 5 editions of Wikipedia as well. Similar observations have been made as for the set of 27 EU countries.

Fig 6. Relationship structure extracted from G_{qrnd} for the network of EU countries. Same legend as Fig 5 except friends and followers are computed from G_{qrnd}.

https://doi.org/10.1371/journal.pone.0201397.g006

Results: *GR* link sensitivity

Influence analysis of geopolitical ties using G_R

We have now established the global mathematical structure G_R and presented how it can be leveraged to extract meaningful geopolitical interactions among countries for the two sets of interest, naming 27 EU and 40 worldwide countries. These interactions are extracted from Wikipedia and thus stem from all links covering this very rich network of webpages. As such, they encompass not only interactions related to economics or politics, but from any possible domain (arts, history, entertainment, etc.). The strength of this study is to show that just from the structure of the network, relevant and timely information can be extracted. The hyper-linked structure of Wikipedia itself contains an important part of the universal knowledge stored in details on the webpages.

Table 4. Cross-edition friends and followers extracted from G_{qr} of EU countries per leading country.

Top country	G_{qr} Wiki friends present in			G_{qr} Wiki followers present in		
	all 5 editions	4 out of 5 editions	3 out of 5 editions	all 5 editions	4 out of 5 editions	3 out of 5 editions
FR	BE -ES	IT		BE	LU—ES	
GB	IE		DK—FR	IE—MT	CY	
ES	IT—PT	FR	BE	MT—PT		LU
SE	DK—FI		EE	DK—EE—FI	LV	
PL	CZ		DE—HU—LT—SK	CZ—LT—SK		LV

For each top country, we list the friends and followers that are identical accros all five Wikipedia editions, in 4 editions out of 5 and in 3 editions out of 5.

https://doi.org/10.1371/journal.pone.0201397.t004

Previous study has shown that G_R captures essential interactions between countries. The point is now to see how some ties between countries influence the whole network structure. More specifically, we focus here on capturing the impact of a change in the strength of a relationship between two countries on the importance of the nodes in the network. Therefore we have designed a sensitivity analysis that measures a logarithmic derivative of the PageRank probability when the transition probability of only one link is increased for a specific couple of nodes in G_R, relatively to the other ones.

Our sensitivity analysis is performed for a directed link where the relationship going from country i to j is increased. We investigate in the last part of this Section the imbalance between the influence of two opposite direction interactions. In other words, we conduct the aforementioned sensitivity analysis for the link going from country i to j, and for the link going in the opposite direction from j to i. For each pair of countries, we derive from this two-way sensitivity the relationship imbalance to identify the most important player in the relationship.

Sensitivity analysis

We define δ as the relative fraction to be added to the relationship from nation j to nation i in G_R. Knowing δ, a new modified matrix $\tilde{G}_R$ is calculated in two steps. First, element $\tilde{G}_R(i, j)$ is set to $(1 + \delta) \cdot G_R(i, j)$. Second, all elements of column j of $\tilde{G}_R$ are normalized to 1 (including element i) to preserve the unity column-normalization property of the Google matrix. Now $\tilde{G}_R$ reflects an increased probability for going from nation j to nation i.

It is now possible to calculate the modified PageRank eigenvector $\tilde{P}$ from $\tilde{G}_R$ using the standard $\tilde{G}_R \tilde{P} = \tilde{P}$ relation and compare it to the original PageRank probabilities P calculated with G_R using $G_R P = P$. The same process can be applied to the transposed version of $\tilde{G}_R$ to calculate the modified CheiRank probabilities $\tilde{P}^*$. Due to the relative change of the transition probability between nodes i and j, steady state PageRank and CheiRank probabilities are modified. This reflects a structural modification of the network and entails a change of importance of nodes in the network. These changes are measured by a logarithmic derivative of the PageRank probability of node a:

$$D_{(j \to i)}(a) = (dP_a/d\delta_{ij})/P_a = (\tilde{P}_a - P_a)/(\delta_{ij}P_a) \tag{14}$$

Notation $(j \to i)$ indicates that the link from node j to node i has been modified. Element $D_{(j \to i)}(a)$ gives the logarithmic variation of PageRank probability for country a if the link from j to i has been modified. We will refer to this variation as the *sensitivity* of nation a to the relationship from nation i to nation j. If this sensitivity is negative, country i has lost importance in the network. On the opposite, a positive sensitivity expresses a gain in importance. The computation has been tested for values of $\delta = \pm 0.01, \pm 0.03, \pm 0.05$. The result is not sensitive to δ and following results are given for $\delta = 0.03$.

Relationship imbalance analysis

As introduced earlier, sensitivity $D_{(j \to i)}(k)$ of Eq (14) measures the change of importance of node a if the link from nation j to i has been changed. The sensitivity of node a to a change in one direction is not necessarily the same as its sensitivity to the change in the opposite direction. We define as such the *2-way sensitivity* of node a which is simply the sum of the sensitivities calculated for both directions:

$$D_{(i \leftrightarrow j)}(a) = D_{(i \to j)}(a) + D_{(j \to i)}(a) \tag{15}$$

The two-way sensitivity can be leveraged to find out, for a pair of countries a and b, which one has the most influence on the other one. Therefore, we define the following metric:

$$F(a, b) = D_{(a \leftrightarrow b)}(a) - D_{(a \leftrightarrow b)}(b) \tag{16}$$

Here, we measure the 2-way sensitivity for nodes a and b when the link between them is modified both ways in G_R. If $F(a, b)$ is positive, it means that the 2-way sensitivity of a is larger than the 2-way sensitivity of b. In this case, a is more influenced by b than b by a. We can say that b is the *strongest* country. If $F(a, b)$ is negative, we can say that a is the strongest country.

Sensitivity results

Sensitivity analysis results are shown first for the 27 EU network and then for the 40 worldwide network. For each network, we have identified a set of meaningful links between countries to be modified and observed resulting sensitivity of other nations. We perform as well for each network the relationship imbalance analysis for each pair of nations. Note that if the modified link is clearly identified, we will drop the index $i \rightarrow j$ in our sensitivity measure notation for clarity.

27 EU network of countries

In order to better capture the countries' sensitivities from a multicultural perspective, we have calculated the sensitivities for 3 Wikipedia editions: EnWiki, FrWiki and DeWiki. All sensitivity results shown for 27 EU network have been averaged over the three editions as follows:

$$\bar{D} = \frac{1}{3} \sum_{i=1}^{3} D^i \tag{17}$$

where index i refers to the Wikipedia edition.

Sensitivity analysis. We start this analysis by introducing a first simple example where Italy increases its relationship with France (cf. Figs 7 and 8). Then, we analyze the impact on the EU countries of Great Britain's exit (i.e. Brexit) from European Union (cf. Figs 9 and 10). Next, we highlight the sensitivity of Luxembourg to the increase of Germany and France's cooperation with other member states (cf. Figs 11 and 12). Finally, we present the results that underline the strong ties that exist between groups of countries that function together in Europe (cf. Figs 13, 14, 15 and 16).

For each sensitivity analysis, we show two types of figures: *i*) an axial representation of the sensitivity $\bar{D}$ and *ii*) a colored map of Europe where countries' color indicate the sensitivity $\bar{D}$ as well. Color scale for these maps plots lower values of $\bar{D}$ in red, median in green and larger in blue. Each map represents the sensitivity values obtained for a given link variation.

Italy to France relationship. Italy is the second top export and import country of Slovenia with \$3.05B and \$3.84B respectively. In 1992, diplomatic relations began between the two countries and in 2012, Foreign Minister of Italy, Giulio Terzi, described the bilateral relationship between Italy and Slovenia as fruitful and dynamic [26]. Politically, Slovenia relies on
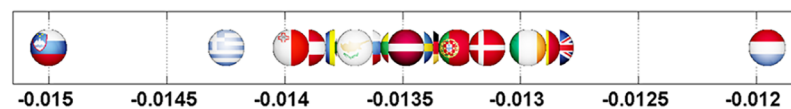
Fig 7. Axial representation of $\bar{D}$ for a link modification from {IT} to {FR}. Here $\bar{D}(IT) = -0.0159$ and $\bar{D}(FR) = 0.0701$ are not shown.
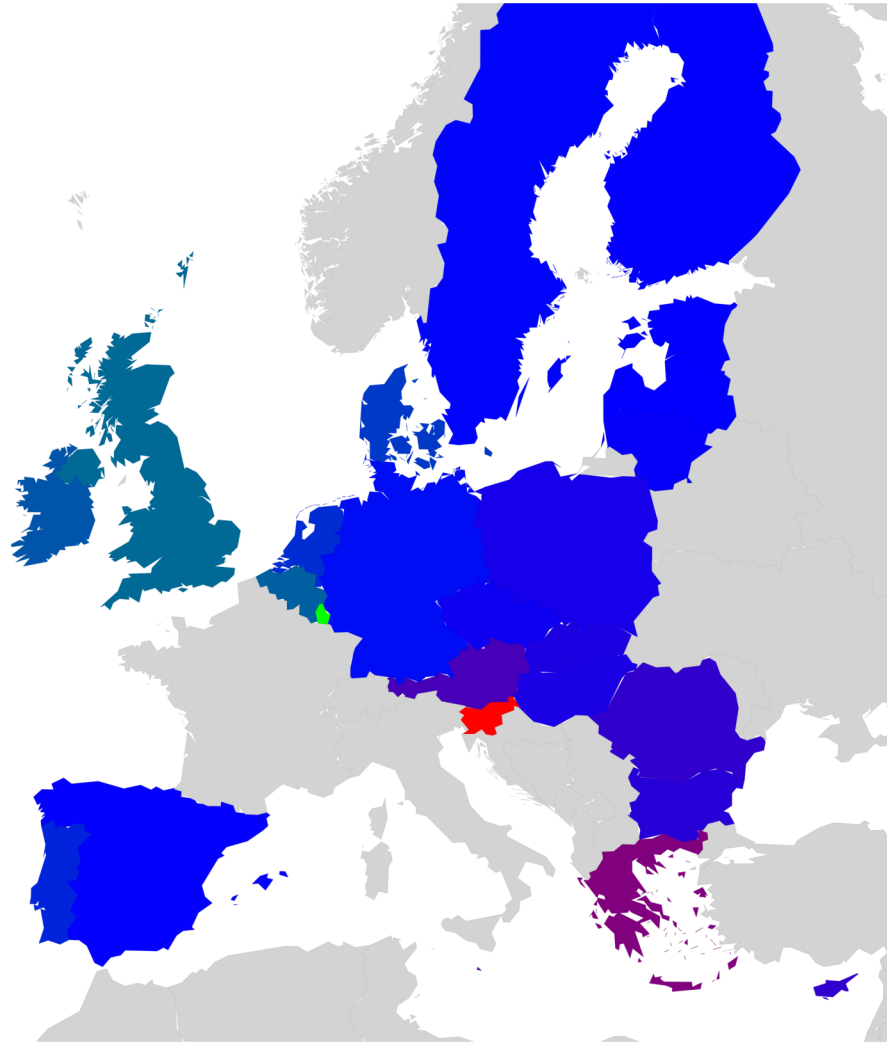
https://doi.org/10.1371/journal.pone.0201397.g007

Fig 8. Map representation of $\bar{D}$ for links modifications from {IT} to {FR}. Lower values of $\bar{D}$ in red, median in green and larger in blue (IT and FR are not shown).

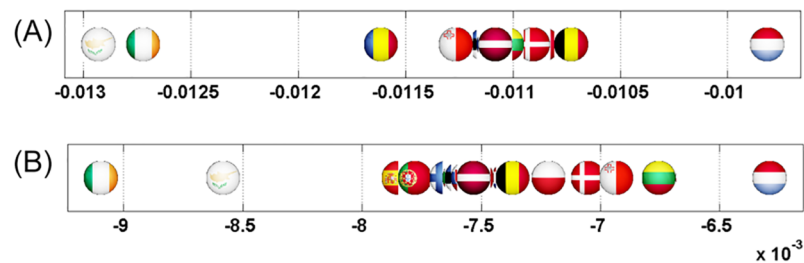https://doi.org/10.1371/journal.pone.0201397.g008

Fig 9. Axial representation of $\bar{D}$ for link modifications from {GB} to {FR or DE}. (A): GB to FR (not shown $\bar{D}(GB) = -0.0124$ and $\bar{D}(FR) = 0.0577$). (B): GB to DE (not shown $\bar{D}(GB) = -0.0087$ and $\bar{D}(DE) = 0.0606$).
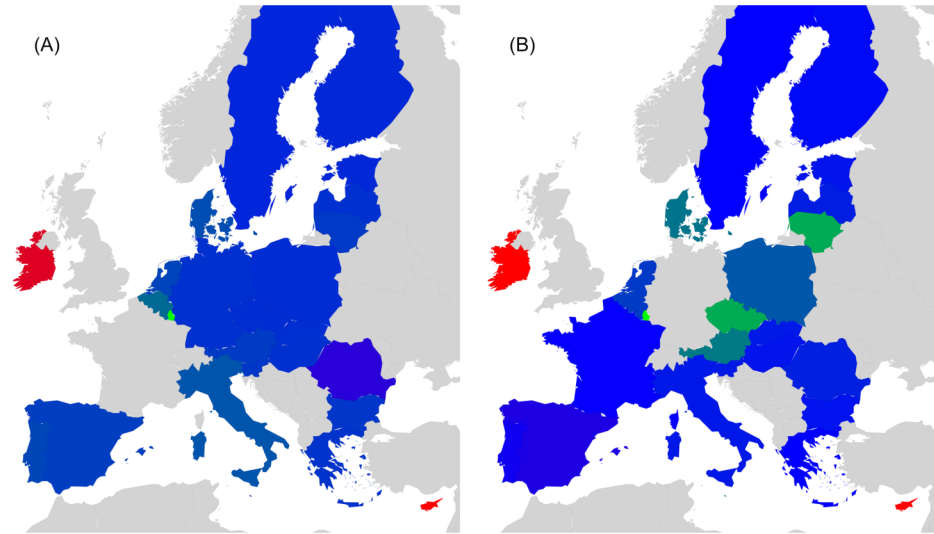
https://doi.org/10.1371/journal.pone.0201397.g009

Fig 10. Map representation of $\bar{D}$ for link modifications from {GB} to {FR or DE}. (A): GB to FR ($\bar{D}$ are not shown for GR, FR); (B): GB to DE($\bar{D}$ are not shown for GR, DE). Lower values of $\bar{D}$ in red, median in green and larger in blue.

Italy to become a member of the principal UN, EU and NATO bodies [26]. No doubt Slovenia would suffer if Italy decided to go away from it and increase its relationships with France. The 27 EU network exactly shows the negative impact of Italy increasing its link in G_R with France: Slovenia is the nation with lowest sensitivity on Figs 7 and 8.

Impact of Brexit. The United Kingdom has triggered article 50 on March 27, 2017 to leave the European Union as a consequence of the referendum of June 23rd, 2016 [27] (Brexit is an abbreviation for Britain exit [27]). To understand its impact on EU countries with our dataset, we have reduced (and not increased as done in other studies) the G_R transition probability UK towards France or Germany. We remind that our network is dated by 2013 but it captures the strong UK influence. Results are shown in Figs 9 and 10 and indicate that Ireland and Cyprus are by far the most negatively affected countries in both cases. Moreover, the sensitivity of UK is negative as it benefits less from France's or Germany's influence. These facts
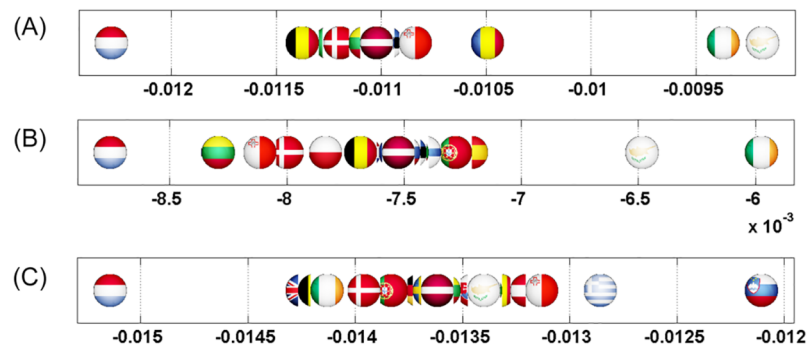
Fig 11. Axial representation of $\bar{D}$ for link modifications from {FR or DE} to {GB or IT}. (A): FR to GB (not shown $\bar{D}(FR) = -0.0117$ and $\bar{D}(GB) = 0.1572$). (B): DE to GB (not shown $\bar{D}(DE) = -0.0081$ and $\bar{D}(GB) = 0.1248$). (C) FR to IT (not shown $\bar{D}(FR) = -0.0143$ and $\bar{D}(IT) = 0.1508$).
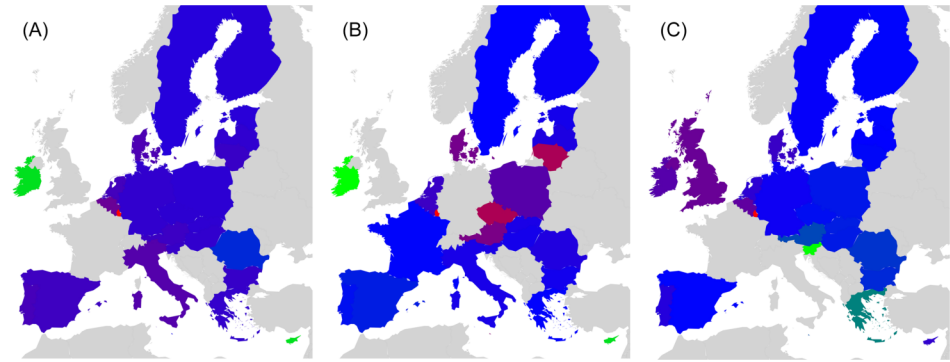
Fig 12. Map representation of $\bar{D}$ for link modifications from {FR or DE} to {IT or GB}: Luxembourg is negatively impacted here. (A): FR to GB. (B): DE to GB. (C) FR to IT. Lower values of $\bar{D}$ in red, median in green and larger in blue; for linked countries $\bar{D}$ is not shown.

have been recently backed up by specialists. In [20], a study delivered by the London School of Economics discussing the consequences of Brexit forecasts that UK will loose 2.8% of its GDP. We recall that the Gross Domestic Product is the monetary value of all the finished goods and services produced within a country's borders in a specific time period [28]. Similarly, [20] shows that Ireland will loose as well 2.3% of its GDP, which is the largest proportional loss caused by Brexit. Cyprus-UK Relations are strong as claimed by the official website of the Ministry of Foreign Affairs of Cyprus [29]. Referring to [30], UK is the 4th top export destination for Cyprus with $242M and the 2nd import origin with $508M. As such, this clear bond of UK
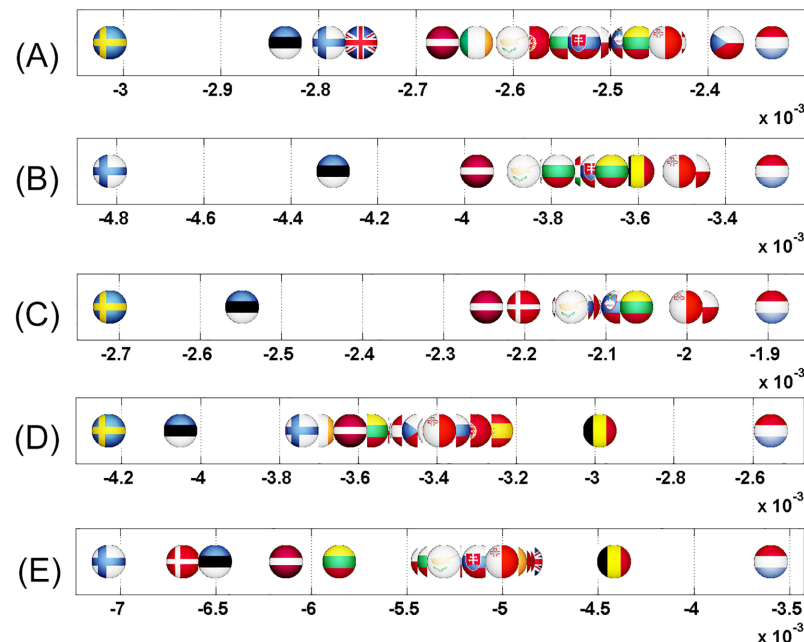
Fig 13. Axial representation of $\bar{D}$ for link modifications from Nordic countries to {FR or DE}. (A): DK to DE (not shown $\bar{D}(DK) = -0.0050$ and $\bar{D}(DE) = 0.0208$). (B): SE to DE (not shown $\bar{D}(SE) = -0.0064$ and $\bar{D}(DE) = 0.0313$). (C): FI to DE (not shown $\bar{D}(FI) = -0.0046$ and $\bar{D}(DE) = 0.0173$). (D): DK to FR (not shown $\bar{D}(DK) = -0.0077$ and $\bar{D}(FR) = 0.0197$). (E): SE to FR (not shown $\bar{D}(SE) = -0.0100$ and $\bar{D}(FR) = 0.0296$).
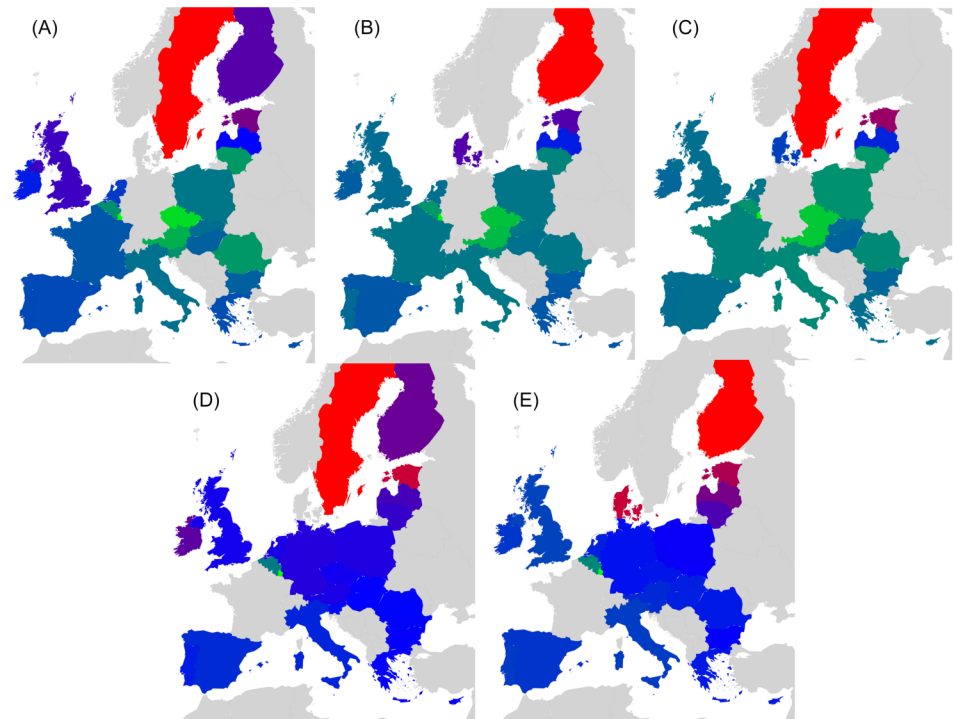
Fig 14. Map representation of $\bar{D}$ for link modifications from Nordic countries to {FR or DE}. (A): DK to DE. (B): SE to DE. (C): FI to DE. (D): DK to FR. (E): SE to FR. Lower values of $\bar{D}$ in red, median in green and larger in blue; for linked countries $\bar{D}$ is not shown.

https://doi.org/10.1371/journal.pone.0201397.g014

with Cyprus explains that if GB suffers from Brexit, Cyprus will do as well. Our data strikingly exhibits the same conclusion as shown in Figs 9 and 10.

Luxembourg's sensitivity to Germany and France. Luxembourg shares its borders with Belgium, Germany and France with whom it has strong and diverse relationships. Luxembourg has a very open economy. Together with Belgium, they position themselves as the 12th largest economy in the world. Two of the top three export and import countries of Belgium-Luxembourg are Germany ($44.6B, $50.4B) and France ($43.8B, $36.8B) [30]. Official languages in Luxembourg are Luxembourgish, French and German. Luxembourg has robust relationships with France [31, 32] and Germany [33] in various areas such as finance, culture, science, security or nuclear power. It is clear that Luxembourg will suffer if one of these European countries reduces its exchanges with it. In Figs 11 and 12, we clearly show with our sensitivity analysis that Luxembourg is strongly influenced by France and Germany. If France or Germany increases its relationships with Italy or Great Britain, Luxembourg is by far the most negatively impacted country.

Clusters of countries. By analyzing the sensitivity of countries to various 2-nation relationships, we have noticed that several groups of nations function together. These groups are strongly interconnected, and if anyone of these group members increases its relationship strength with a country outside of the group, all group members loose importance in the network. We highlight two meaningful examples next: the cluster of Nordic countries and the cluster Austro-Hungarian cluster. Other clusters we have identified in our network are for instance the cluster of Benelux countries (e.g. Belgium, the Netherlands and Luxembourg) or the cluster of the Iberian peninsula (e.g. Portugal and Spain).
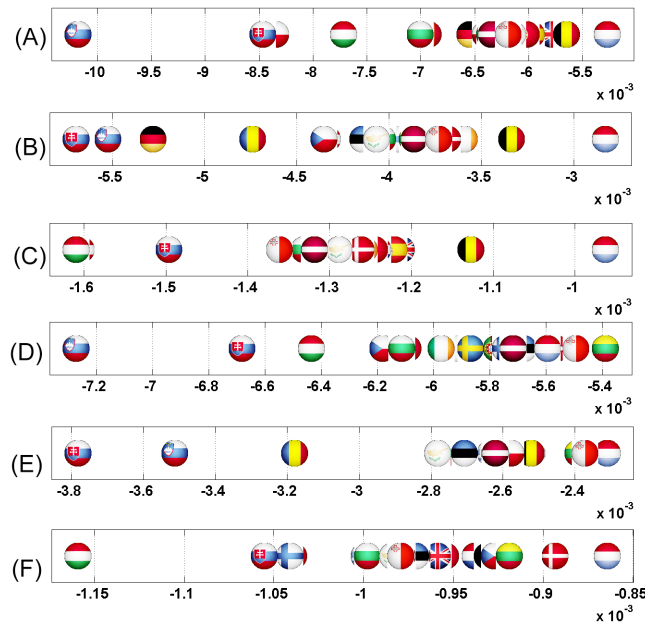
Fig 15. Axial representation of $\bar{D}$ for link modifications from {AT, HU and SI} to {FR or DE}. (A): AT to FR (not shown $\bar{D}(AT) = -0.0101$ and $\bar{D}(FR) = 0.0373$). (B): HU to FR (not shown $\bar{D}(HU) = -0.0080$ and $\bar{D}(FR) = 0.0205$). (C): SI to FR (not shown $\bar{D}(SI) = -0.0046$ and $\bar{D}(FR) = 0.0075$). (D): AT to DE (not shown $\bar{D}(AT) = -0.0070$ and $\bar{D}(DE) = 0.0393$). (E): HU to DE (not shown $\bar{D}(HU) = -0.0052$ and $\bar{D}(DE) = 0.0311$). (F): SI to DE (not shown $\bar{D}(SI) = -0.0034$ and $\bar{D}(DE) = 0.0081$).

https://doi.org/10.1371/journal.pone.0201397.g015

For both investigated groups, we test the influence of an increase in collaboration from one member of the group to France or to Germany. France and Germany have been chosen as they are central members of European Union.

The Nordic countries Denmark, Finland, and Sweden have much in common: their way of life, history, language and social structure [18]. After World War II, the first concrete step into unity was the introduction of a Nordic Passport Union in 1952. Nordic countries co-operate in the Nordic Council, a geopolitical forum. In the Nordic Statistical Yearbook [18], Klaus Munch illustrates that "The Nordic economies are among the countries in the Western World with the best macroeconomic performance in the recent ten years". Nordic countries should keep cooperating to stay strong. Thus, if any Nordic country attempts to abandon these relationships in favor of other countries, it will negatively impact the remaining Nordic countries. Our sensitivity analysis illustrates this impact in Figs 13 and 14. In these figures, we show how the relationship increase between any Nordic country towards France or Germany induces a drop in sensitivity for Nordic countries.

Referring to [19], relations between Slovenia, Hungary and Austria are tight. Hungary has supported Slovenia for its NATO membership applications and Austria has assisted Slovenia in entering European Union. Relationships between Austria and Hungary are important for both countries in the economic, political and cultural fields [34]. Concerning economy [30], Austria is one of the top import origins for Hungary and Slovenia with $5.54B and $2.37B respectively. Similarly to the Nordic group of countries, if Austria, Slovenia or Hungary increases its relationships with another European country, the other two will be affected. Sensitivity analysis backs up this statement as seen in Figs 15 and 16.

Relationship imbalance analysis. Relationship imbalance analysis has been derived for all pairs of European countries following Eq (16). Fig 17 shows a density plot of $F(a, b)$. We
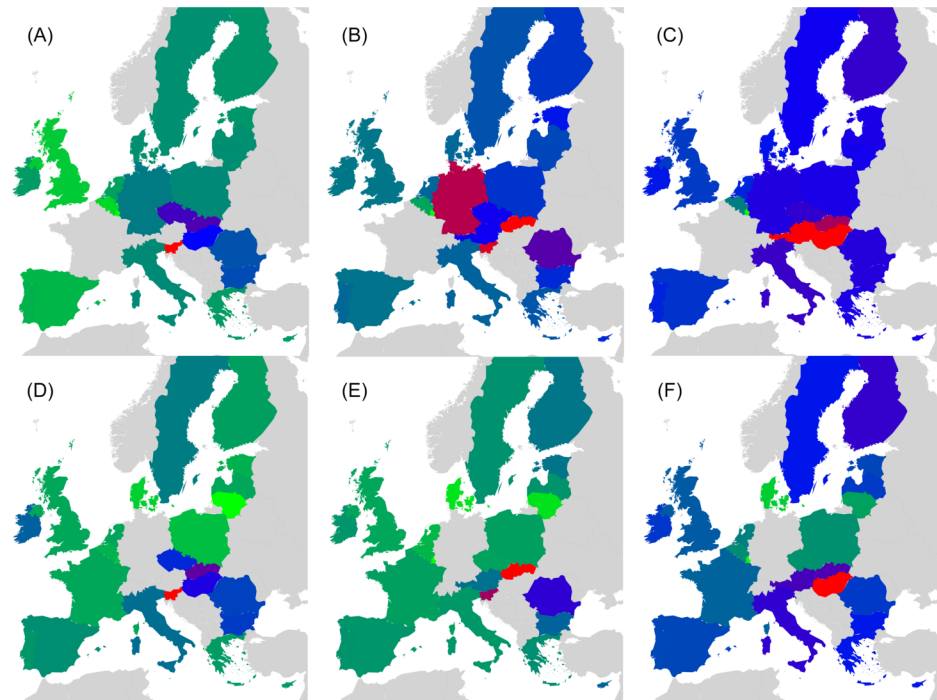
Fig 16. Map representation of $\bar{D}$ for link modifications from {AT, HU and SI} to {FR or DE}. (A): AT to FR. (B): HU to FR. (C): SI to FR. (D): AT to DE. (E): HU to DE. (F): SI to DE. Lower values of $\bar{D}$ in red, median in green and larger in blue; for linked countries $\bar{D}$ is not shown.
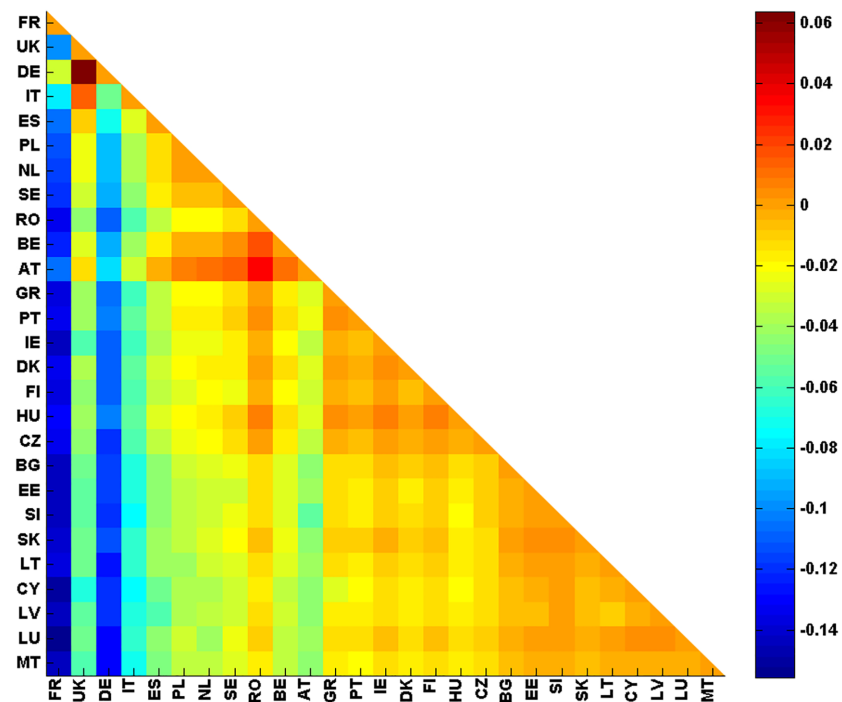
https://doi.org/10.1371/journal.pone.0201397.g016

Fig 17. Relationship imbalance analysis: F-representation for 27 EU network. $F(a, b)$ is given by the colorbar. X-axis and Y-axis represent a and b respectively. If $F(a, b)$ is negative, nation a has more influence on nation b than b on a.

https://doi.org/10.1371/journal.pone.0201397.g017
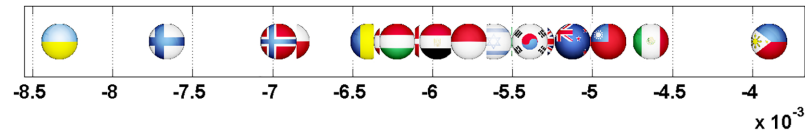
Fig 18. Axial representation of $\bar{D}$ for link modification from RU to US. (not shown $\bar{D}(RU) = -0.0089$ and $\bar{D}(US) = 0.0446$).

https://doi.org/10.1371/journal.pone.0201397.g018

recall that if $F(a, b)$ is negative, nation a has more influence on nation b than b on a. If $F(a, b)$ is positive, nation b dominates nation a. According to *The Globe of Economic Complexity* [35] and identical to our results in Fig 17, Germany and France are the two largest economies in Europe. From G_R we can clearly see the dominance of France and Germany on other EU countries. Another interesting result of Fig 17 is the equal influence between all pairs of countries created by one member of {GR, PT, IE, DK, FI, HU} and another of {BG, EE, SI, SK, LT, CY, LV, LU, MT}. These pairs have $F(a, b)$ close to zero and are plotted with orange color in Fig 17.

40 worldwide network of countries

Similarly to the 27 EU countries dataset, sensitivity results are averaged over 5 Wikipedia editions: ArWiki, EnWiki, FrWiki, RuWiki and DeWiki. We first show as well the sensitivity analysis for carefully selected links and then conclude this part with the sensitivity imbalance analysis for all pairs of countries.

Sensitivity analysis. In this worldwide set of countries, we have identified relationships whose impact on the network clearly shows how meaningful the sensitivity analysis proposed in this paper is.

US—Russia. As mentioned previously in the introduction, and according to the results in Figs 18 and 19, Ukraine would be the most affected country if Russia gets closer to US. This is
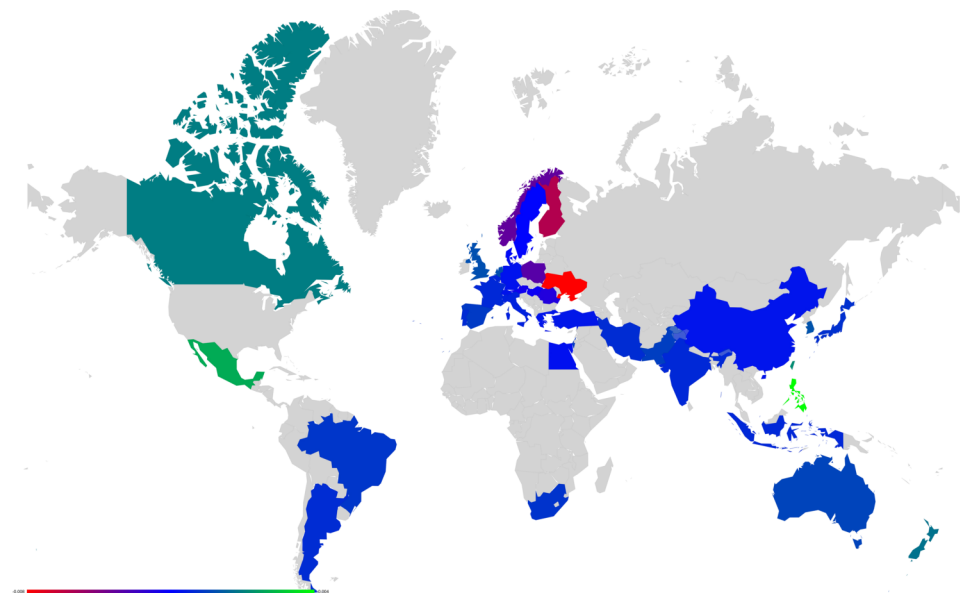

Fig 19. Map representation of $\bar{D}$ for link modification from RU to US. For linked countries $\bar{D}$ is not shown.

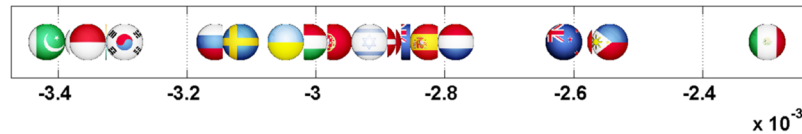https://doi.org/10.1371/journal.pone.0201397.g019

Fig 20. Axial representation of $\bar{D}$ for link modification from CN to US (discarding TW). (not shown $\bar{D}(CN) = -0.0056$, $\bar{D}(US) = 0.0210$ and $\bar{D}(TW) = -0.0087$).

due to the fact that Ukraine and Russia were both in the USSR and their economies are strongly interconnected. The next influenced country is Finland which also has strong economic relations with Russia being a part of Russian Empire till beginning of 20th century.

China to US. The effects of an increase in the relationship from China to US are shown in Figs 20 and 21. Taiwan and Pakistan are the most negatively affected countries. Taiwan is not pictures in Figs 20 and 21 as it greatly reduces readability of the plots. Indeed, sensitivity of Taiwan is $\bar{D}(TW) = -0.0087$, 4 times the one of Pakistan. BBC's article [36] on the division between China and Taiwan illustrates that US is the most important friend and the only ally of Taiwan. China claims Taiwan as its territory and Taiwan counts on US to establish its full independence to stand up against China. As such, if the ties between China and US get stronger, Taiwan will loose its best ally.

In 1951, Pakistan and China officially established their diplomatic relations and in 2016 they celebrated 65 years of friendship [37]. Regarding security strategy, China has always supported Pakistan in facing terrorism. Politically, Pakistan stands with China on many issues concerning China's core interests (e.g. Taiwan, Tibet, Xinjiang). The trade volume between the two countries reached $100.11B by 2015 and in 2016 the $46B China-Pakistan Economic Corridor (CPEC) [38] was constructed. If China strengthens its relationship with US, Pakistan
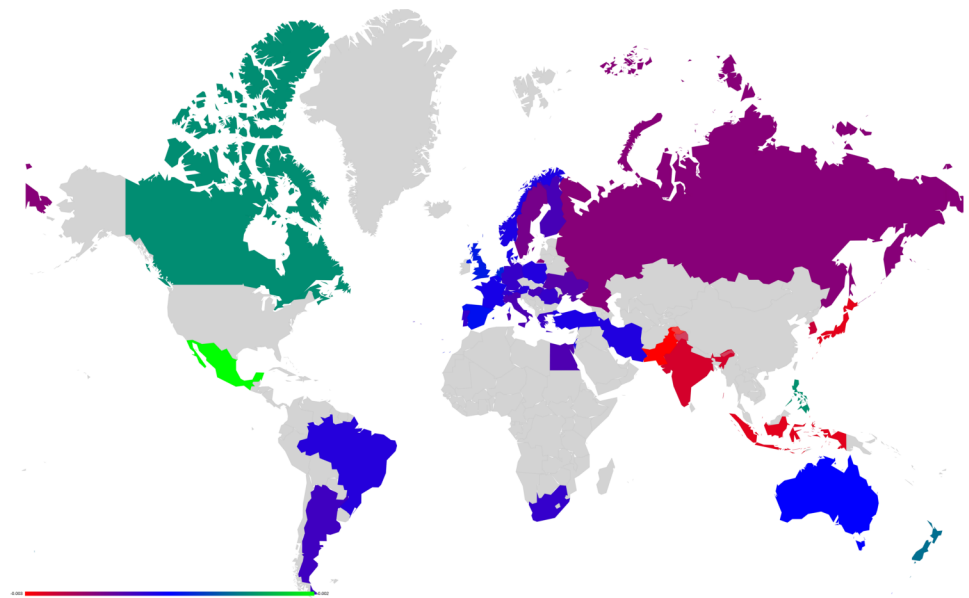

Fig 21. Map representation of $\bar{D}$ for link modification from CN to US (discarding TW). For linked countries $\bar{D}$ is not shown.
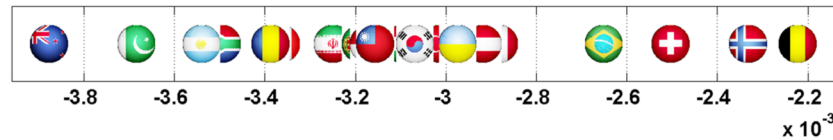
Fig 22. Axial representation of $\bar{D}$ for link modification from GB to FR. (not shown $\bar{D}(GB) = -0.00403$ and $\bar{D}(FR) = 0.0368$).

may clearly suffer from it. An article by Ian Price [39] raises a serious question on whether United States aims at sabotaging the CPEC in the near future.

United Kingdom to France.   The modification of this link gives the most strong effect on New Zealand (see Figs 22 and 23). Indeed, referring to New Zealand Ministry of Foreign Affairs and Trade [40], UK is the top destination for New Zealand's goods and services exports within the EU, and a base for New Zealand companies doing business in Europe. According to the statistics of March 2015, the total trade in goods between the two countries is $2,807 billion. New Zealand works closely with UK to face terrorism: strategic dialogue talks on security policy issues with UK are held every year. Also, New Zealand shares important cultural and historical links with UK. For New Zealand, UK is the key to Europe. This means intuitively that New Zealand will be strongly affected by the Brexit. These facts are totally in line with our sensitivity analysis conclusions plotted in Figs 22 and 23. In order to face the consequences of Brexit together, UK and NZ have started a serious discussion as mentioned in [41, 42].

US-Israel-Egypt.   The Arab-Israeli relationship has been conflicting ever since the Jewish community has shown interest in establishing a nation-state in Palestine. The 1917 Balfour Declaration favored the establishment of a Jewish national home in Palestine and US supported it [43]. On November 29 1947, the United Nations General Assembly adopted the partition resolution number 181 [44] that would divide Palestinian territory into Jewish and Arab
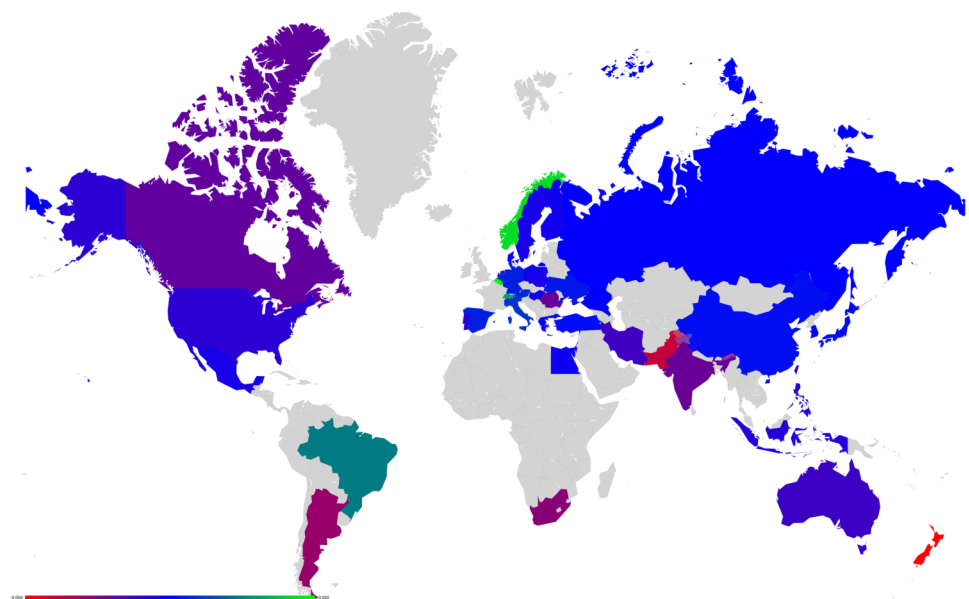

Fig 23. Map representation of $\bar{D}$ for link modification from GB to FR. For linked countries $\bar{D}$ is not shown.
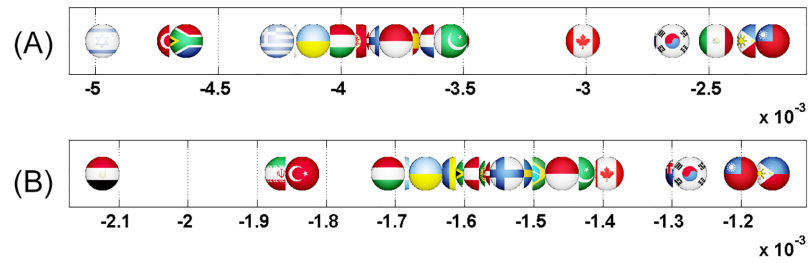
Fig 24. Axial representation of $\bar{D}$ for links modifications from {IL and EG} to US. (A):EG to US (not shown $\bar{D}(EG) = -0.0080$ and $\bar{D}(US) = 0.0252$). (B):IL to US (not shown $\bar{D}(IL) = -0.0041$ and $\bar{D}(US) = 0.0108$).
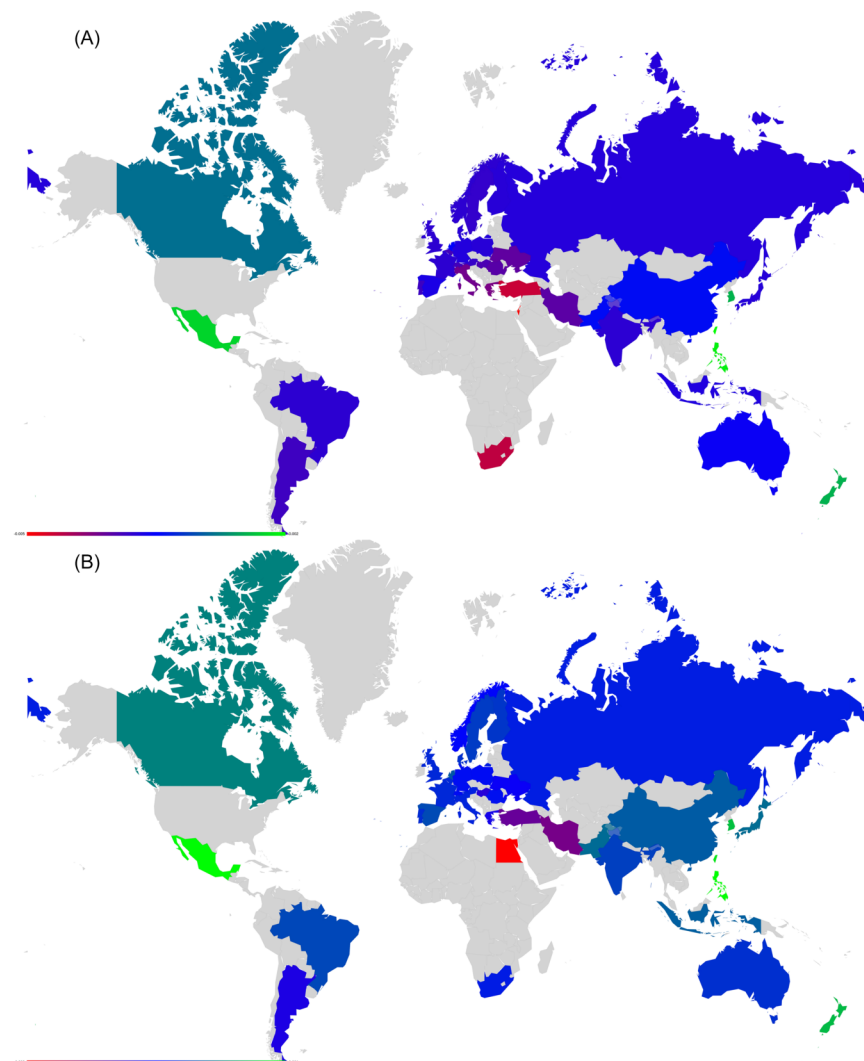
https://doi.org/10.1371/journal.pone.0201397.g024

Fig 25. Map representation of $\bar{D}$ for links modifications from {IL and EG} to US. (A):EG to US. (B):IL to US. For linked countries $\bar{D}$ is not shown.

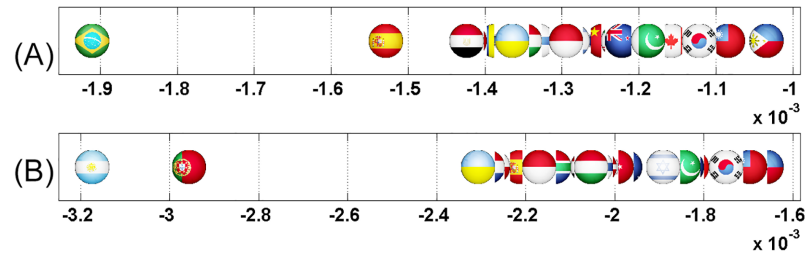https://doi.org/10.1371/journal.pone.0201397.g025

Fig 26. Axial representation of $\bar{D}$ for link modifications from {AR and BR} to US. (A): AR to US (not shown $\bar{D}(AR) = -0.0050$ and $\bar{D}(US) = 0.0094$). (B): BR to US (not shown $\bar{D}(BR) = -0.0074$ and $\bar{D}(US) = 0.0149$).
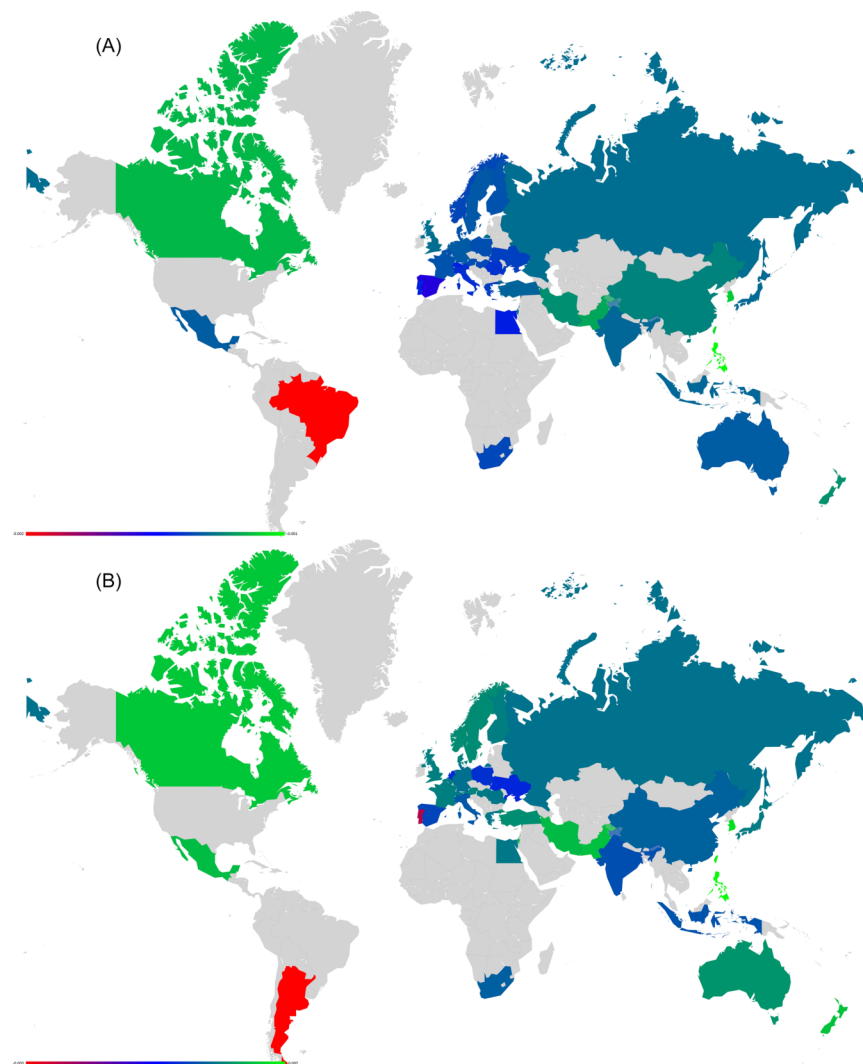
https://doi.org/10.1371/journal.pone.0201397.g026

Fig 27. Map representation of $\bar{D}$ for link modifications from {AR and BR} to US. (A): AR to US. (B): BR to US. For linked countries $\bar{D}$ is not shown.

https://doi.org/10.1371/journal.pone.0201397.g027

states. Again, US stood aside Israel in supporting the United Nations resolution. Palestinians (and Arabs in general) denounced the partition. Since then, Arab-Israeli did combat in five major wars (1948, 1956, 1967, 1973 and 1982) with Egypt the leader of Arab side in 3 out of 5 wars. Even though the Camp David Accords [45] between Egypt and Israel were signed on September 17, 1978 followed by a peace treaty on March 26, 1979 [46] (both being signed in US and witnessed by Jimmy Carter), the relationship is still conflicting. It has been called the "cold peace". On the other side, Israeli-US relations are getting stronger according to Jeremy M. Sharp [47]: Israel is the largest cumulative recipient of US foreign aid since World War II. Our results show in Figs 24 and 25 that Egypt and Israel will be the most affected countries if the other one gets closer to US.

Argentina and Brazil. Their relationship [48] includes all possible fields: economy, history, culture, trade and social structure. As members of the Mercosur sub-regional bloc, Argentina and Brazil relationship offers free trade and fluid movement of goods, people, and currency. Besides that, a Nuclear Cooperation between these two countries was signed on July 18, 1991 and the Brazilian-Argentine Agency for Accounting and Control of Nuclear Materials (ABACC) was created as a binational safeguard organization. Comparing our results (shown in Figs 26 and 27) with these facts of strong relationship between Argentina and Brazil, we find that any unilateral rapprochement between Argentina or Brazil to US will negatively affect the other country.

Relationship imbalance analysis. Relationship imbalance analysis has been derived for all pairs of 40 countries following Eq (16) as well. Fig 28 shows a density plot of $F(a, b)$. US is clearly the dominant country among all other 39 countries chosen worldwide. Also, Fig 28 shows that some countries have a strong influences such as France, Germany, Russia, China and Egypt. Germany and France are the two main players of European Union. Russia has an long history of sovereignty over eastern Europe and Northern Asia, economically, politically
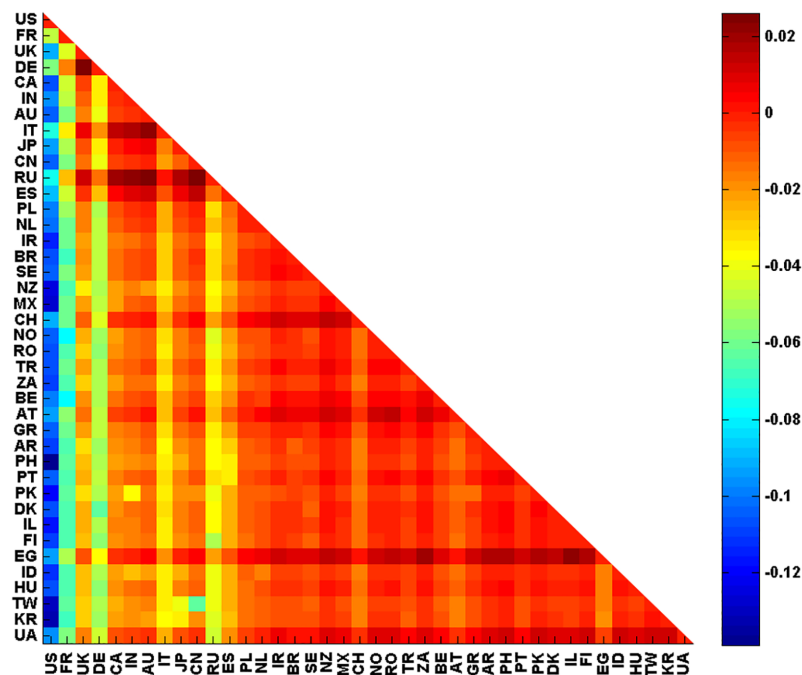

Fig 28. Relationship imbalance analysis: F-representation for 27 EU network. $F(a, b)$ is given by the colorbar. X-axis and Y-axis represent a and b respectively. If $F(a, b)$ is negative, nation a has more influence on nation b than b on a.

https://doi.org/10.1371/journal.pone.0201397.g028

and culturally. Egypt plays a central role in the middle east. China, with its large population and strong economy, is dominating several countries. However, its role may be underestimated since no Chinese Wikipedia edition is accounted for in our study.

Discussion

This work offers a new perspective for future geopolitics studies. It is possible to extract from multicultural Wikipedia networks a global understanding of the interactions between countries at a global, continental or regional scale. Reduced Google matrix theory has been shown to capture hidden interactions among countries, resulting in new knowledge on geopolitics. Results show that our sensitivity analysis captures the importance of relationships on network structure. This analysis relies on the reduced Google matrix and leverages its capability of concentrating all Wikipedia knowledge in a small stochastic matrix. We stress that the obtained sensitivity of geopolitical relations between two countries and its influence on other world countries is obtained on a pure mathematical statistical analysis without any direct appeal to political, economical and social sciences. We hope that our mathematical studies will be useful for further developments of artificial intelligence approaches tailored to analyze political relationships. Experts in geopolitics are welcome to leverage this framework to first validate our mathematical results and to second drive the search for new knowledge.

Acknowledgments

Author Contributions

Conceptualization: Dima L. Shepelyansky.

Data curation: Dima L. Shepelyansky.

Formal analysis: Dima L. Shepelyansky.

Funding acquisition: Katia Jaffrès-Runser, Dima L. Shepelyansky.

Investigation: Samer El Zant, Katia Jaffrès-Runser, Dima L. Shepelyansky.

Methodology: Samer El Zant, Katia Jaffrès-Runser, Dima L. Shepelyansky.

Project administration: Katia Jaffrès-Runser.

Resources: Katia Jaffrès-Runser.

Software: Samer El Zant.

Supervision: Katia Jaffrès-Runser, Dima L. Shepelyansky.

Validation: Samer El Zant.

Visualization: Samer El Zant.

Writing – original draft: Samer El Zant.

Writing – review & editing: Katia Jaffrès-Runser, Dima L. Shepelyansky.

References

1. Jones E. The European Miracle: Environments, Economies and Geopolitics in the History of Europe and Asia. Cambridge UK: Cambridge University Press; 2003.

2. Javanmardi S, Lopes C. Statistical Measure of Quality in Wikipedia. Proc. of 1st Workshop on Soc. Med. Anal. (SOMA'10), New York, NY, USA, 2010; 132-138.

3. Giles J. Internet encyclopaedias go head to head. Nature 2005; 438: 900–901. https://doi.org/10.1038/438900a PMID: 16355180

4. Rajagopalan MS, Khanna V, Stott M, Leiter Y, Showalter T, Dicker AP et al. Accuracy of Cancer Information on the Internet: A Comparison of a Wiki with a Professionally Maintained Database. Journal of Clinical Oncology. 2010; 28:15_suppl 6058–6058. https://doi.org/10.1200/jco.2010.28.15_suppl.6058

5. Brin Sergey, Page Lawrence. The anatomy of a large-scale hypertextual Web search engine. Computer Networks and ISDN Systems. 1998; 30:1–7 107–117. https://doi.org/10.1016/S0169-7552(98)00110-X

6. Langville AM and Meyer CD. Google's PageRank and beyond: the science of search engine rankings. Princeton: Princeton University Press; 2006.

7. Zhirov AO, Zhirov OV, Shepelyansky DL. Two-dimensional ranking of Wikipedia articles. Eur. Phys. J. B, 2010; 77: 523. https://doi.org/10.1140/epjb/e2010-10500-7

8. Eom Y-H, Frahm KM, Benczur A, Shepelyansky A. Time evolution of Wikipedia network ranking. Eur. Phys. J. B, 2013; 86: 492. https://doi.org/10.1140/epjb/e2013-40432-5

9. Eom Y-H, DL Shepelyansky DL. Highlighting entanglement of cultures via ranking of multilingual Wikipedia articles. PLOS ONE 2013; 8(10): e74554. https://doi.org/10.1371/journal.pone.0074554 PMID: 24098338

10. Eom Y-H, Aragon P, Laniado D. Interactions of cultures and top people of Wikipedia from ranking of 24 language editions. PLoS One 2015; 10(3):e0114825. https://doi.org/10.1371/journal.pone.0114825 PMID: 25738291

11. Ermann L, Frahm KM, Shepelyansky DL. Google matrix analysis of directed networks. Rev. Mod. Phys. 2015; 87: 1261. https://doi.org/10.1103/RevModPhys.87.1261

12. Lages J, Patt A, Shepelyansky DL. Wikipedia ranking of world universities. Eur. Phys. J. B, 2016; 89: 69. https://doi.org/10.1140/epjb/e2016-60922-0

13. Hart MH. The 100: ranking of the most influential persons in history. New-York: Citadel Press; 1992.

14. Academic Ranking of World Universities. Available from: http://www.shanghairanking.com/. Cited July 2018.

15. Frahm KM, Shepelyansky DM. Reduced Google matrix; 2016. Preprint. Available from: arXiv:1602.02394. Cited July 2018.

16. Frahm KM, Jaffrès-Runser K, Shepelyansky DL. Wikipedia mining of hidden links between political leaders. Eur. Phys. J. B 2016; 89:269. https://doi.org/10.1140/epjb/e2016-70526-3

17. Frahm KM, El Zant S, Jaffrès-Runser K, Shepelyansky DL. Multi-cultural Wikipedia mining of geopolitics interactions leveraging reduced Google matrix analysis. Physics Letters A; 2017; 381:33 2677–2685. https://doi.org/10.1016/j.physleta.2017.06.021

18. Munch Haagensen K. Nordic Statistical Yearbook 2014. [Internet] Nordic Council of Ministers, Copenhagen 2014. Available from: http://norden.diva-portal.org/smash/get/diva2:763002/FULLTEXT07.pdf. Cited July 2018.

19. Slovenia Business Law Handbook: Strategic Information and Laws. USA: International Business Publications, 2015.

20. Dhingra S, Ottaviano G, Sampson T, Van Reenen J. The consequences of Brexit for UK trade and living standards. London School of Economics and Political Science—CEP BREXIT Analysis 2016; 2.

21. Fukuyama F. Trump ou la fin de la démocratie libérale. Courrier International, Nov. 2016, N1360. [Internet] Available from https://www.courrierinternational.com/article/politique-trump-ou-la-fin-de-la-democratie-liberale. Cited July 2018.

22. Eom Y-H, Aragon P, Laniado D. 24 Wikipedia networks datasets—February 2013. [Internet] Available from http://www.quantware.ups-tlse.fr/QWLIB/topwikipeople/index.html. Cited July 2018.

23. Chepelianskii AD. Towards physical laws for software architecture. Preprint. Available from: arXiv:1003.5455 [cs.SE] (2010). Cited July 2018.

24. Hu Y. Efficient, High-Quality Force-Directed Graph Drawing. The Mathematica Journal. 2006; 10:1.

25. Bastian M, Heymann S, Jacomy M. Gephi: An Open Source Software for Exploring and Manipulating Networks. Proc. of ICWSM. 2009: 361-362.

26. Italy-Slovenia: 20 years of fruitful relations, says Terzi. Italy Ministry of Foreign Affairs. [Internet] Available from http://www.esteri.it/mae/en/sala_stampa/archivionotizie/approfondimenti/2012/03/20120327_itslov.html. Cited July 2018.

27. Hunt A, Wheeler B. Brexit: All you need to know about the UK leaving the EU. British Broadcasting Corporation (BBC), 27 March 2017. Available from http://www.bbc.com/news/uk-politics-32810887. Cited July 2018.

28. Investopedia website. Gross Domestic Product—GDP. [Internet] Available from http://www.investopedia.com/terms/g/gdp.asp. Cited July 2018.

29. Cyprus Ministry of Foreign Affairs. Cyprus—UK Relations. [Internet] Available from http://www.mfa.gov.cy/mfa/highcom/london.nsf/hcrcl10_en/hcrcl10_en?OpenDocument. Cited July 2018.

30. Website of The Observatory of Economic Complexity. OEC—Portugal (PRT) Exports, Imports, and Trade Partners. [Intenet] Available from http://atlas.media.mit.edu/en/profile/country/prt/. Cited July 2018.

31. French Ministry of Economy—Treasurer. Les échanges commerciaux entre la France et le Luxembourg en 2014. [Internet] Available from http://www.tresor.economie.gouv.fr/File/411730. Cited July 2018.

32. French Ministry of Foreign Affairs and International Development. La France et le Luxembourg. [Internet] Available from http://www.diplomatie.gouv.fr/fr/dossiers-pays/luxembourg/la-france-et-le-luxembourg/. Cited July 2018.

33. Federal Foreign Office. Bilateral relations—Luxembourg. [Internet] Available from http://www.auswaertiges-amt.de/EN/Aussenpolitik/Laender/Laenderinfos/01-Nodes/Luxemburg_node.html. Cited July 2018.

34. Kőrösi I. Austrian and Hungarian Relations since 1989, the Current Situation and Future Perspectives. Centre for Economic and Regional Studies HAS Institute of World Economics, August 01, 2013. [Internet] Available from http://kki.hu/assets/upload/FPR_Beliv_035.pdf. Cited July 2018.

35. Center for International Development. The Globe of Economic Complexity. [Internet] Available from http://globe.cid.harvard.edu/. Cited July 2018.

36. British Broadcasting Corporation (BBC). What's behind the China-Taiwan divide?. 3 Dec 2016. [Internet] Available from http://www.bbc.com/news/world-asia-34729538. Cited July 2018.

37. Pakistan Today. Pakistan and China relations: 65 years of friendship to strategic partnership. 15 May 2016. [Internet] Available from http://www.pakistantoday.com.pk/2016/05/15/pakistan-and-china-relations-65-years-of-friendship-to-strategic-partnership/. Cited July 2018.

38. Stevens A. Pakistan lands $46 billion investment from China. CNN Money, 20 April 2015. [Internet] Available from http://money.cnn.com/2015/04/20/news/economy/pakistan-china-aid-infrastucture/. Cited July 2018.

39. Price I. Is the US Trying to Sabotage the China-Pakistan Economic Corridor?. The Diplomat, 29 September 2016. [Internet] Available from http://thediplomat.com/2016/09/is-the-us-trying-to-sabotage-the-china-pakistan-economic-corridor/. Cited July 2018.

40. New Zealand Ministry of Foreign Affairs and Trade. United Kingdom. [Internet] Available from https://www.mfat.govt.nz/en/countries-and-regions/europe/united-kingdom/. Cited July 2018.

41. Singh A, Connolly S. UK will attempt to agree post-Brexit trade deal with New Zealand, Theresa May says. Independent, 14 Jan 2017. [Internet] Available from http://www.independent.co.uk/news/uk/politics/theresa-may-new-zealand-trade-deal-bill-english-brexit-downing-street-a7526956.html. Cited July 2018.

42. NewsHub. UK keen for NZ trade deal after Brexit. [Internet] Available from http://www.newshub.co.nz/home/politics/2017/01/uk-keen-for-nz-trade-deal-after-brexit.html. Cited July 2018.

43. The Arab-Israeli War of 1948. http://history.state.gov [Internet] Available from https://history.state.gov/milestones/1945-1952/arab-israeli-war. Cited July 2018.

44. United Nations. United Nations Resolution 181: Resolution adopted on the report of the AD HOC committee on the palestinian question. [Internet] Available from http://www.un.org/en/ga/search/view_doc.asp?symbol=A/RES/181(II). Cited July 2018.

45. Israel Ministry of Foreign Affairs. Camp David Accords. [Internet] Available from http://www.mfa.gov.il/mfa/foreignpolicy/peace/guide/pages/camp%20david%20accords.aspx. Cited July 2018.

46. Israel Ministry of Foreign Affairs. Israel-Egypt Peace Treaty. [Internet] Available from http://www.mfa.gov.il/mfa/foreignpolicy/peace/guide/pages/israel-egypt%20peace%20treaty.aspx. Cited July 2018.

47. Sharp JM. US Foreign Aid to Israel. December 22, 2016. [Internet] Available from https://fas.org/sgp/crs/mideast/RL33222.pdf. Cited July 2018.

48. Pasquariello Mariano KL. Two to Tango: An Analysis of Brazilian-Argentine Relations. Brazilian Political Science Review. 2013; 7:1.