



Факультет экономических наук

Экономика и статистика

Москва

АНАЛИЗ ЭКСПОРТА И ИМПОРТА ГАЗА НА СТРАНОВОМ УРОВНЕ В УСЛОВИЯХ ГЛУБОКОЙ НЕОПРЕДЕЛЕННОСТИ

Выполнил:

Студент группы БСТ213

Еськов Михаил Михайлович

Руководитель:

Профессор, Алескеров Фуад Тагиевич



Выгрузка данных

Данные были взяты с сайта [UN Comtrade](https://comtrade.un.org/) по коду 2711 - Petroleum gas and other gaseous hydrocarbons. В него входят природный газ, пропан, бутан и остальные углеводородные газы, в том числе и в сжиженном виде. Данные были выгружены с 2001 года по 2020 год. Они были объединены в 4 блока по пять лет: 2001-2005гг, 2006-2010гг, 2011-2015гг и 2016-2020гг.

В данных (отправитель и получатель) присутствуют зоны nes (not elsewhere specified): Areas, nes; Other Asia, nes и т. д. Также есть Special Categories, которые не раскрываются по просьбе партнёра страны. Все эти наблюдения были удалены, поскольку нас интересует сетевой анализ на страновом уровне.

Помимо этого, в данных присутствуют пропуски в столбце объёма (Qty в кг). Чтобы исказить выборку как можно меньше, каждый пропуск был заменён на частное стоимости экспорта (импорта) и средней цены за взятый период.

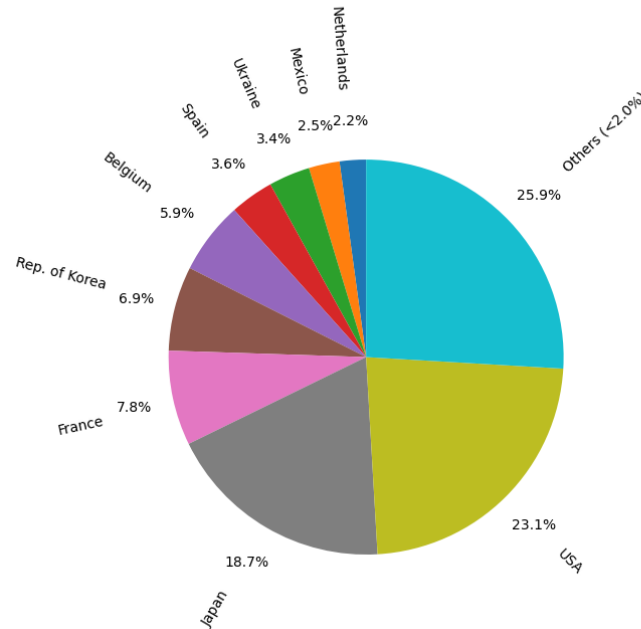
	ReporterDesc	PartnerDesc	PrimaryValue	Qty
0	Afghanistan	Azerbaijan	7.797488e+06	2.121909e+07
1	Afghanistan	China	2.392110e+05	5.980270e+05
2	Afghanistan	Iran	2.160327e+08	5.955653e+08
3	Afghanistan	Kazakhstan	2.327081e+08	6.401172e+08
4	Afghanistan	Pakistan	3.246000e+03	8.115000e+03
5	Afghanistan	Russian Federation	1.655448e+07	4.455844e+07
6	Afghanistan	Tajikistan	9.316000e+03	2.328900e+04
7	Afghanistan	Turkmenistan	4.274165e+08	1.127645e+09
8	Afghanistan	Türkiye	1.420800e+04	3.552000e+04
9	Afghanistan	United Arab Emirates	1.964578e+06	4.911445e+06



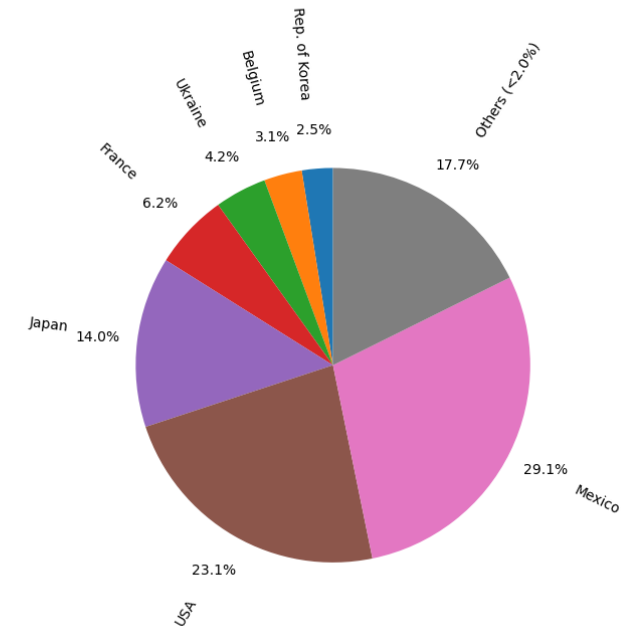
Проблема в данных

Как можно видеть на графиках, очень сильно выбивается пара Мексика - США. Если обычно разница между стоимостью и количеством (кг) долей от общего импорта (экспорта) составляет не более 1%, то для этой пары отклонение очень сильное. В блоке 2001-2005гг в мировом импорте по стоимости Мексика составляла долю в 2,5%, то по объёму 29.1%. Аналогичная ситуация у Америки только по экспорту. Это связано с тем, что в этой строке Мексика – США есть пропуск объёма, который был заполнен на основе средней цены. А истинная цена, вероятно, была выше средней. Таких несоответствий может быть несколько, поэтому проводить анализ по объёму будет некорректно. Из-за неоднородности экспортируемого газа, достаточно сложно будет более грамотно заполнить пропуски, поэтому я буду делать анализ по стоимости экспорта/импорта.

World Import in 2001-2005 by value

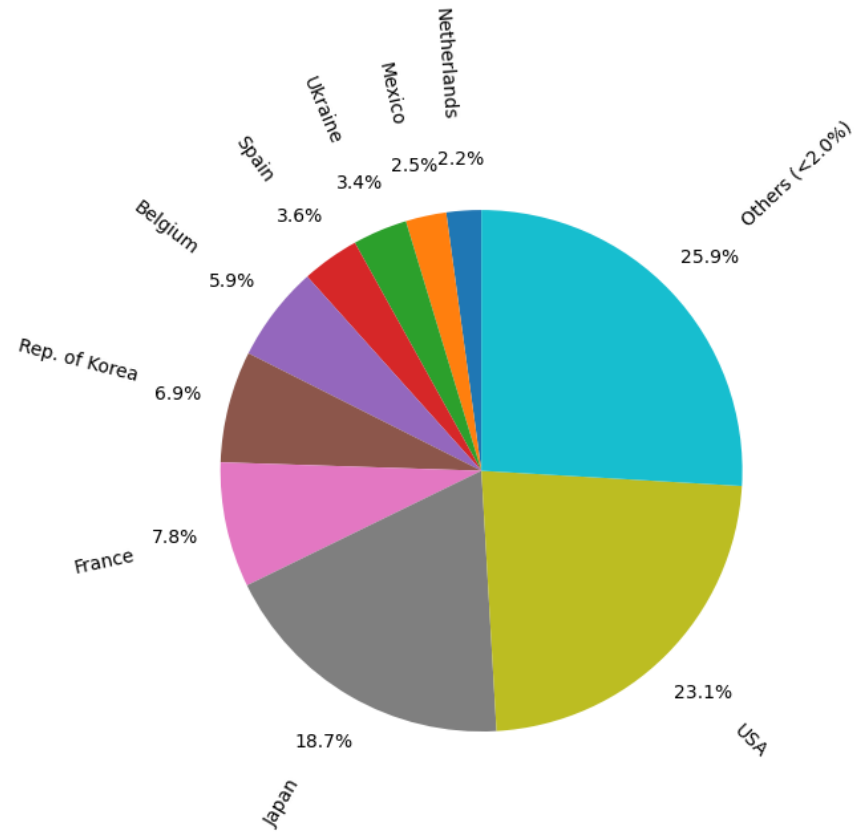


World Import in 2001-2005 by Qty

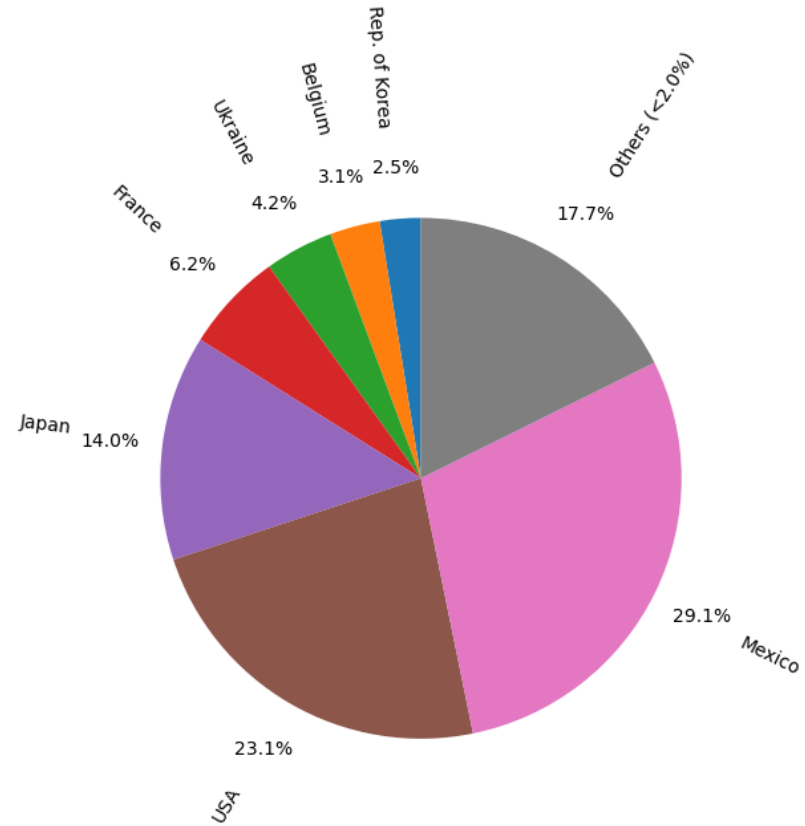




World Import in 2001-2005 by value



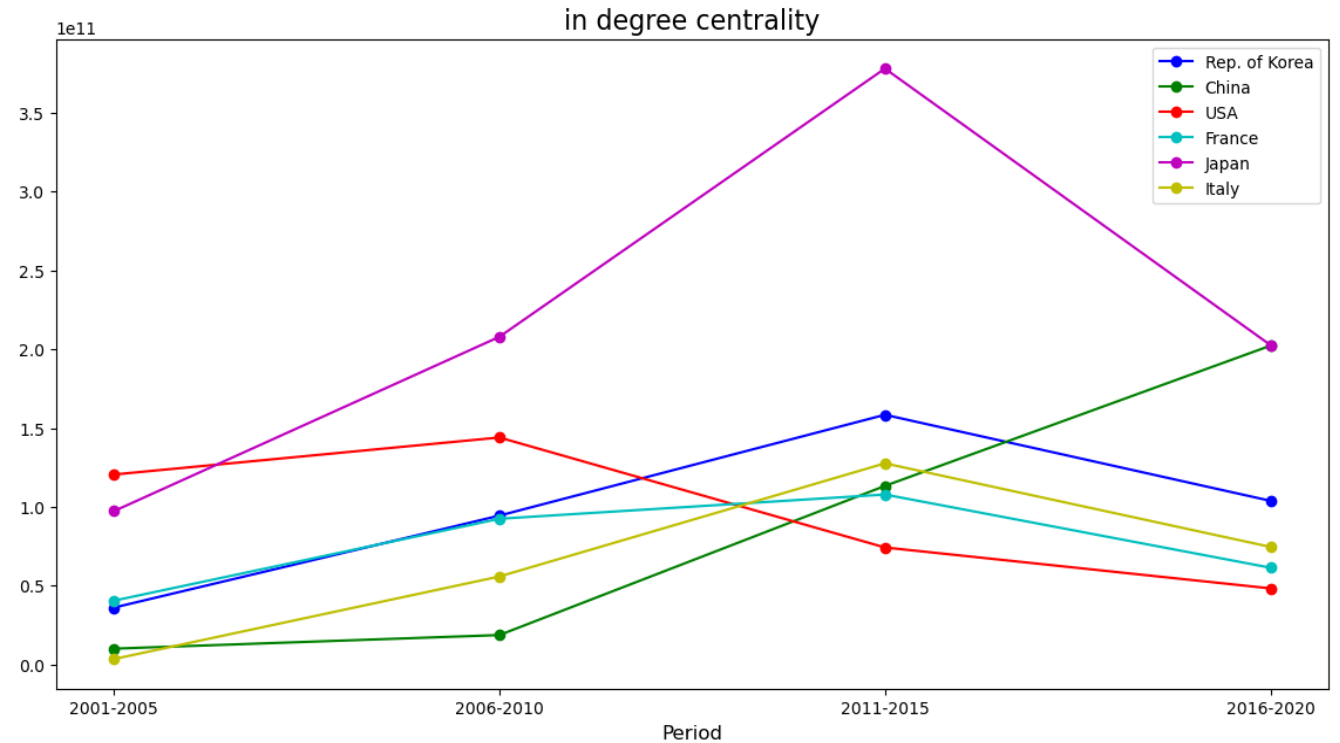
World Import in 2001-2005 by Qty



Weighted In-degree centrality

В нашем случае, Weighted In-degree центральность не что иное, как мировой импорт, т.е. сколько всего страна импортировала в данный период. Чтобы не загромождать работу таблицами со значениями индексов, я буду строить визуализацию по странам с наибольшими значениями индексов.

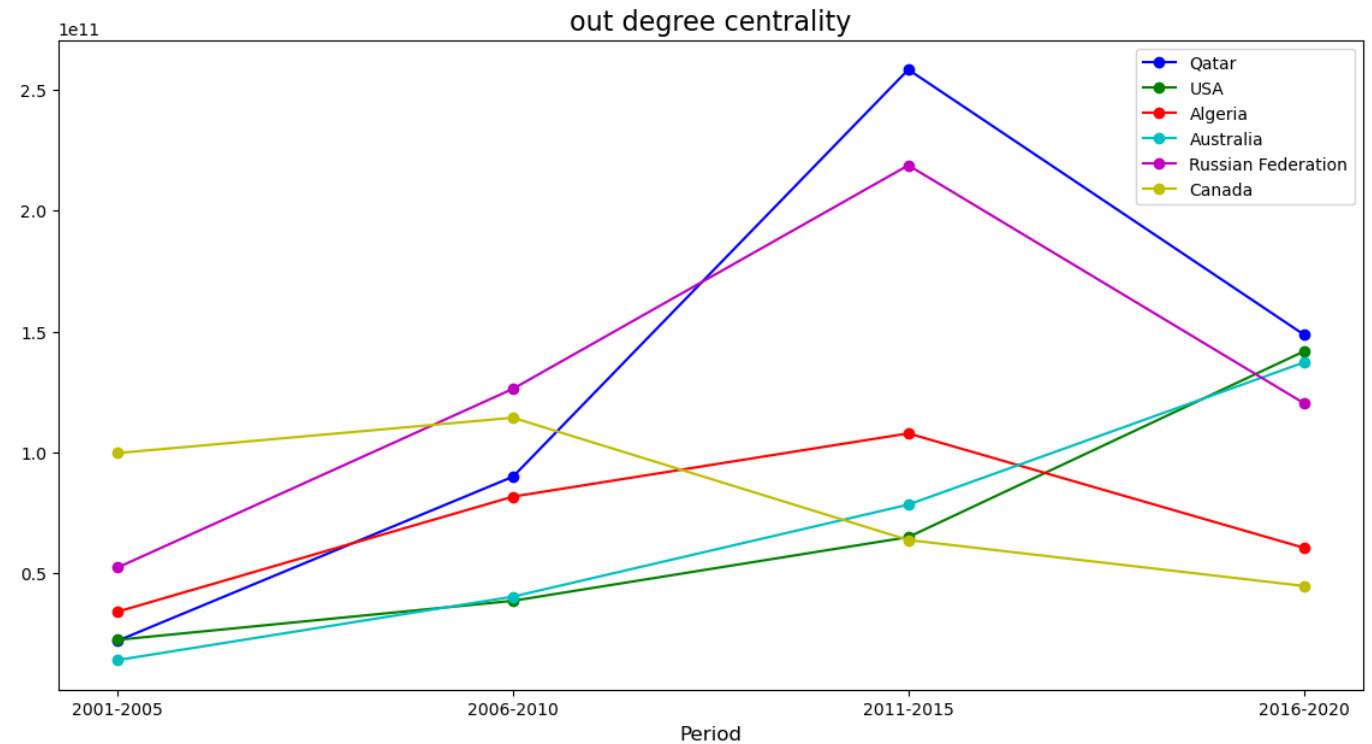
Таким образом, самыми большими импортёрами являются следующие страны: Южная Корея, Китай, США, Франция, Япония и Италия.



Weighted Out-degree centrality

В нашем случае, Weighted Out-degree центральность не что иное, как мировой экспорт, т.е. сколько всего страна экспортировала в данный период.

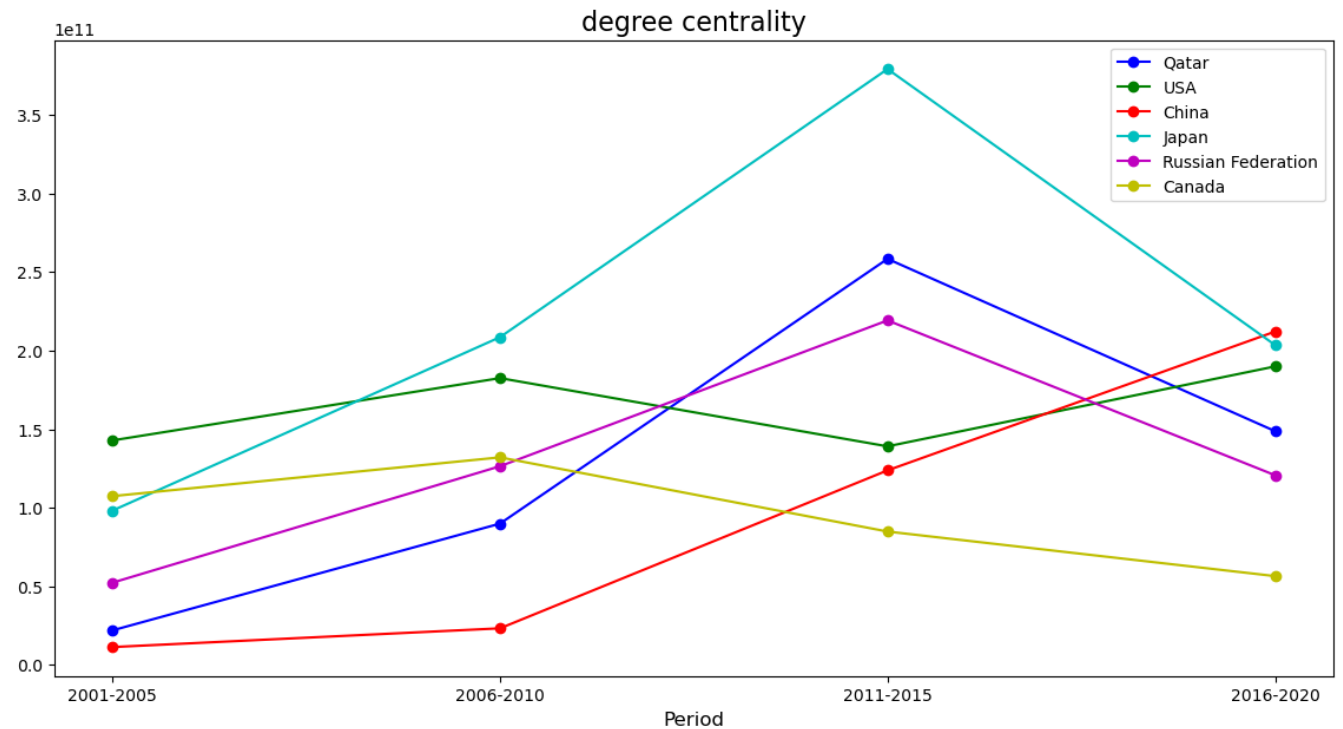
Как можно видеть, самыми большими экспортёрами являются следующие страны: Катар, США, Алжир, Австралия, Россия и Канада.



Weighted Degree centrality

Степенная центральность представляет собой сумму in и out-degree центральность, то есть сумма экспорта и импорта для каждой страны.

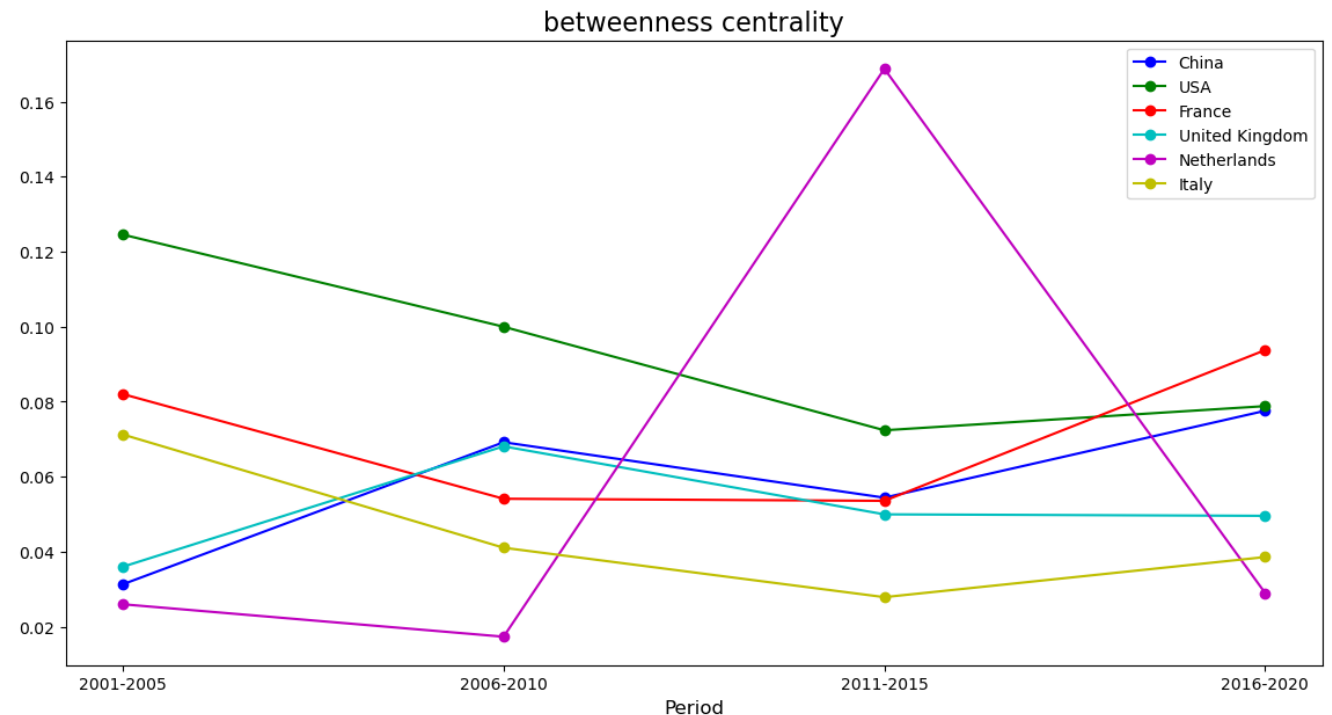
В топ стран по индексу Degree вошли Катар, США, Китай, Япония, Россия и Канада.



Betweenness centrality

Центральность по посредничеству показывает, насколько часто определенный узел выступает в качестве посредника или промежуточного элемента на кратчайших путях между другими парами узлов в сети. Узлы с высокой центральностью по посредничеству играют важную роль в связи и передаче информации между другими узлами в сети.

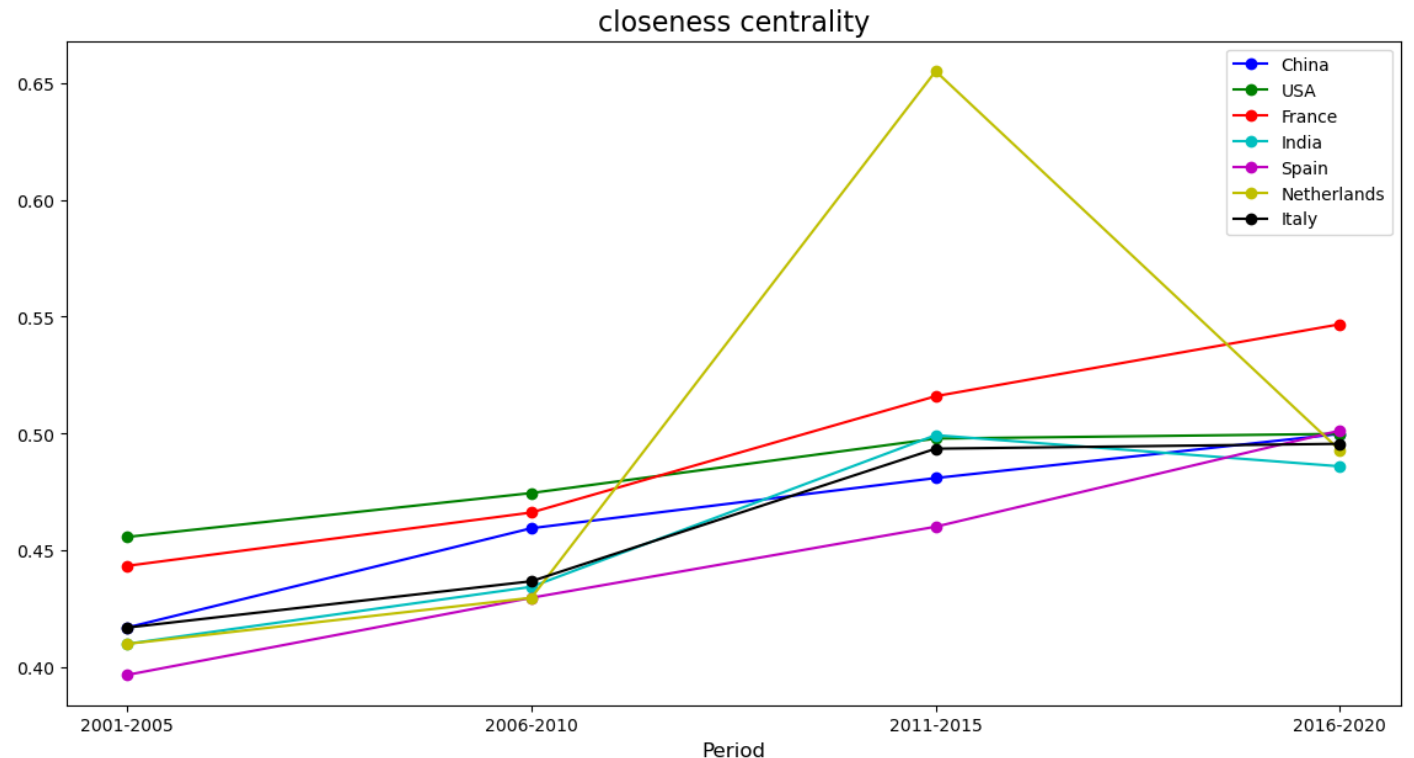
Наиболее центральными по посредничеству стали страны Китай, США, Франция, Великобритания, Нидерланды и Италия



Closeness centrality

Центральность по близости показывает, насколько близко (или насколько быстро) узел может достичь всех других узлов в сети. Узлы с более высоким значением центральности по близости считаются более "близкими" к другим узлам и имеют более быстрый доступ к информации или ресурсам в сети.

По графику видно, что самыми центральным по близости являются страны Китай, США, Франция, Индия, Испания, Нидерланды и Италия.

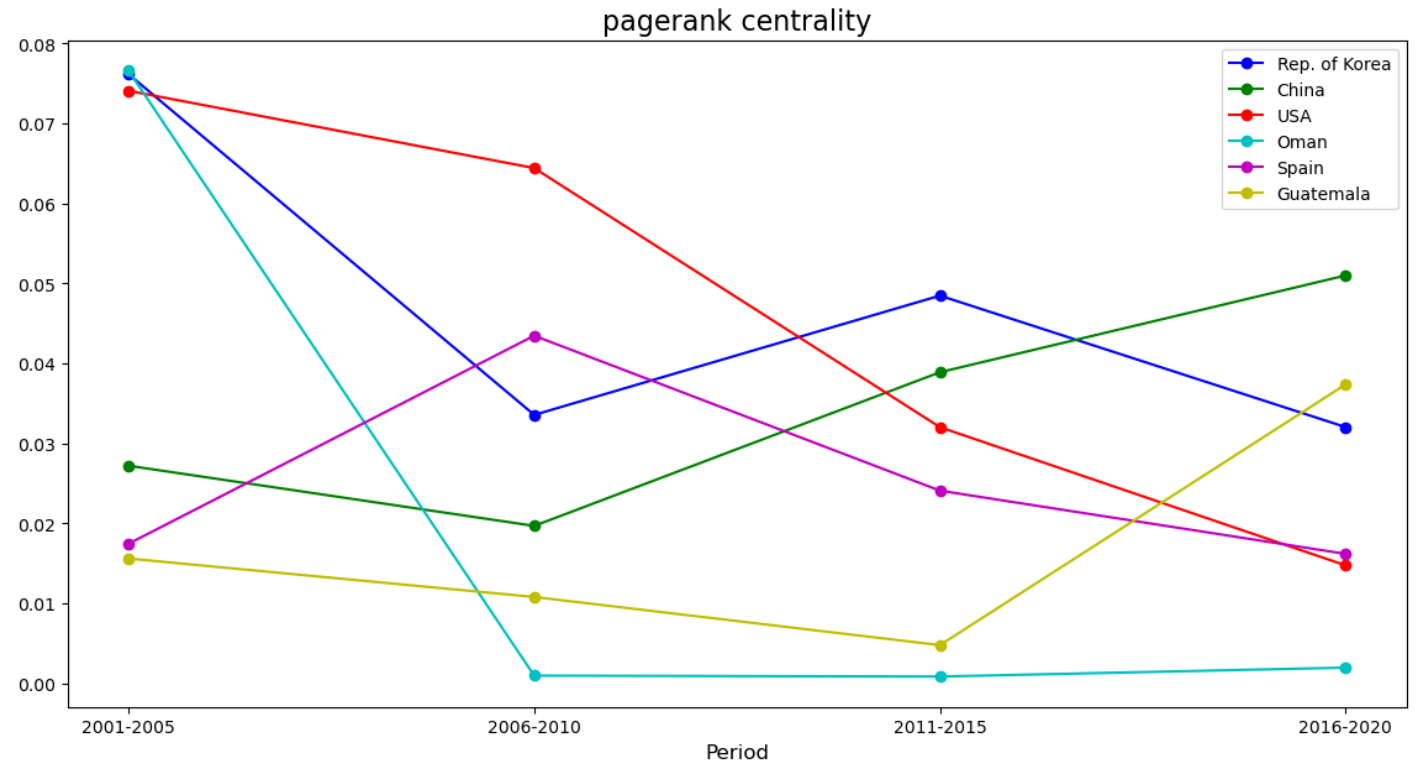


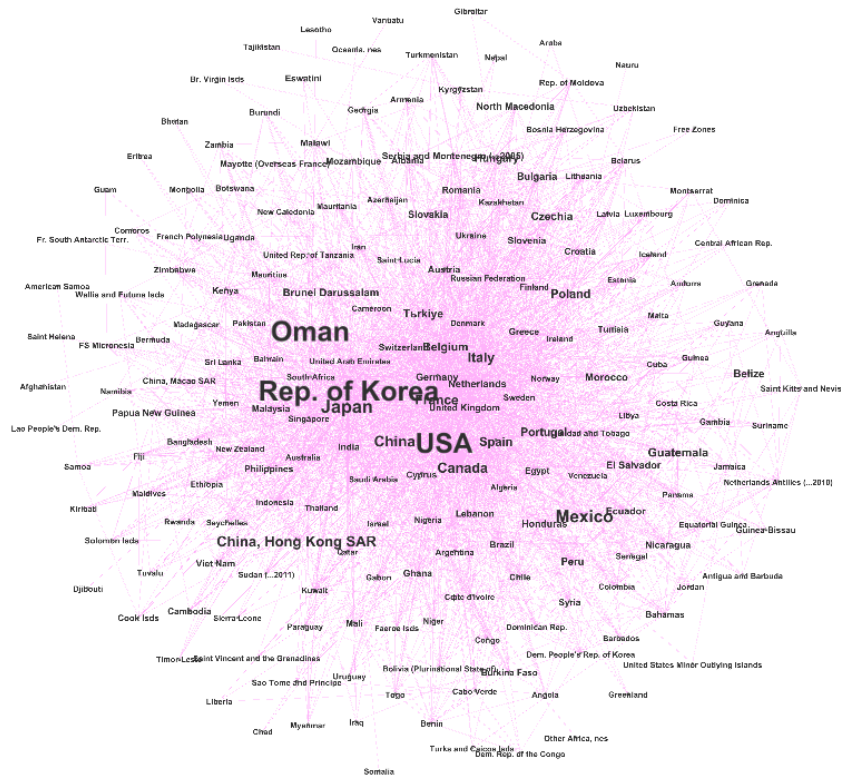
PageRank

Также был посчитан индекс PageRank для данных сетей.

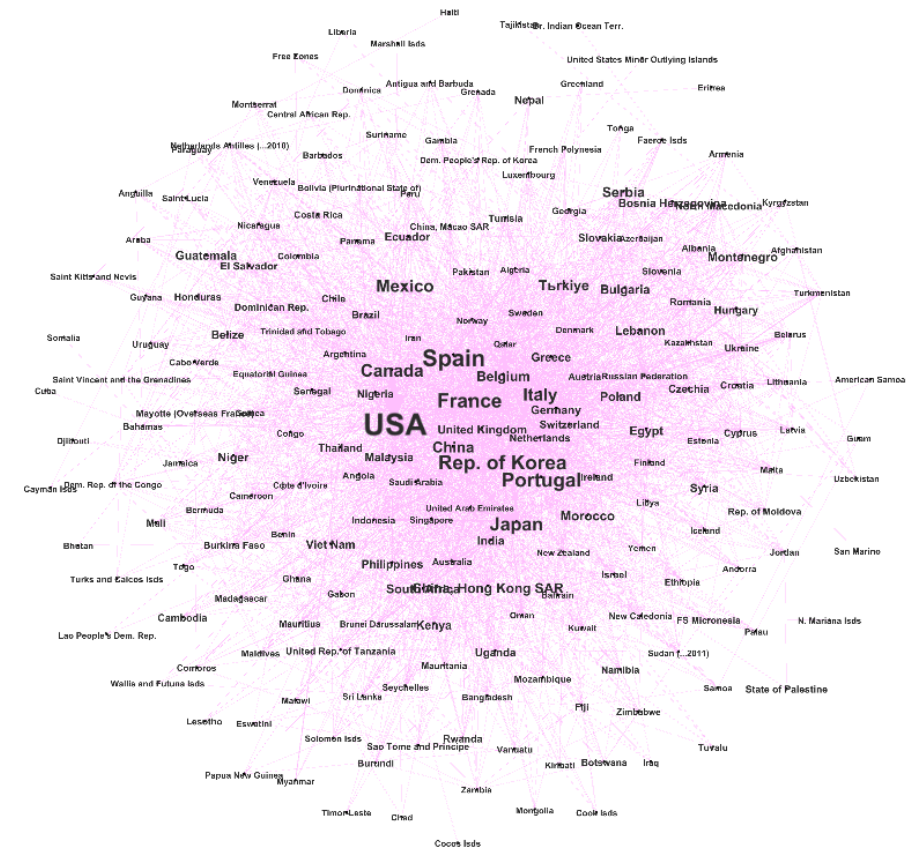
По индексу PageRank наиболее важными оказались страны: Южная Корея, Китай, США, Оман, Испания и Гватемала.

Дополнительно, в Gephi была построена сеть на основе PageRank. Чем больше значение PageRank, тем больше название страны.

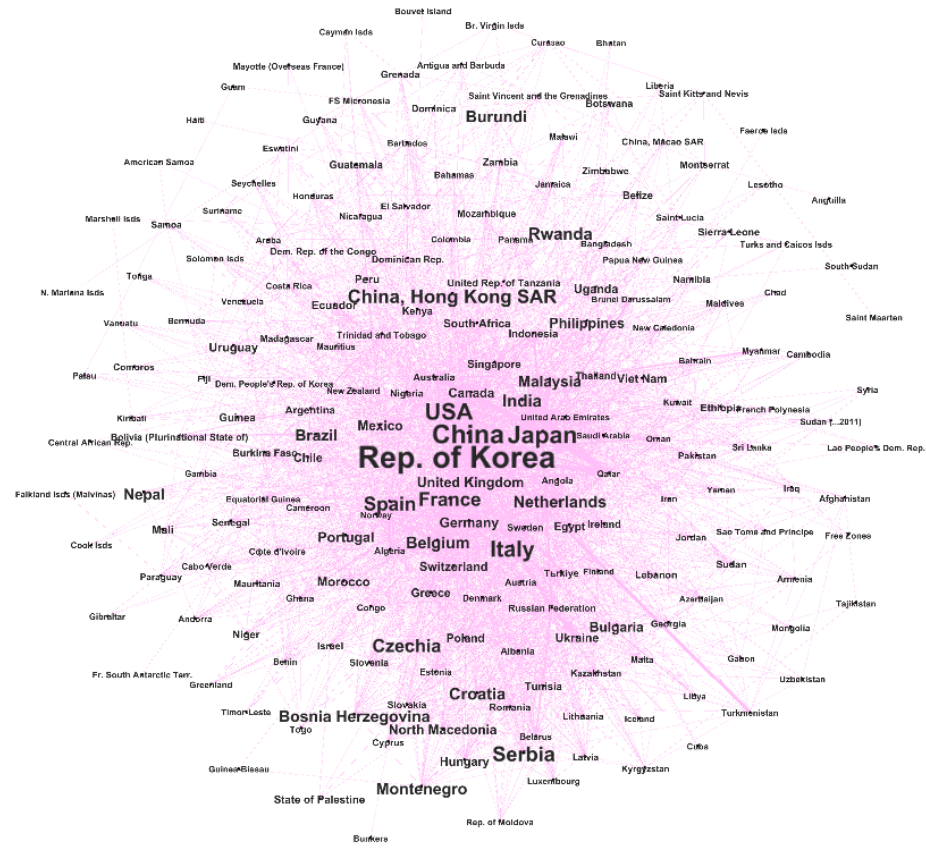




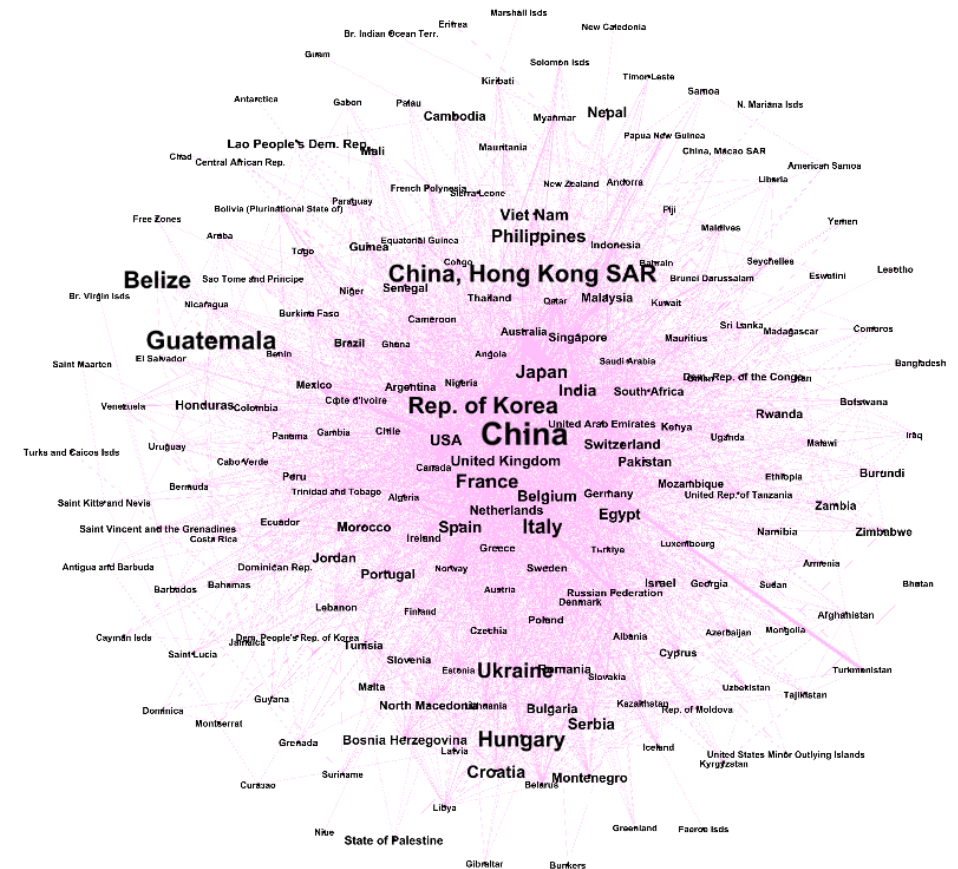
Сеть 2001-2005гг



Сеть 2006-2010гг



Сеть 2011-2015гг

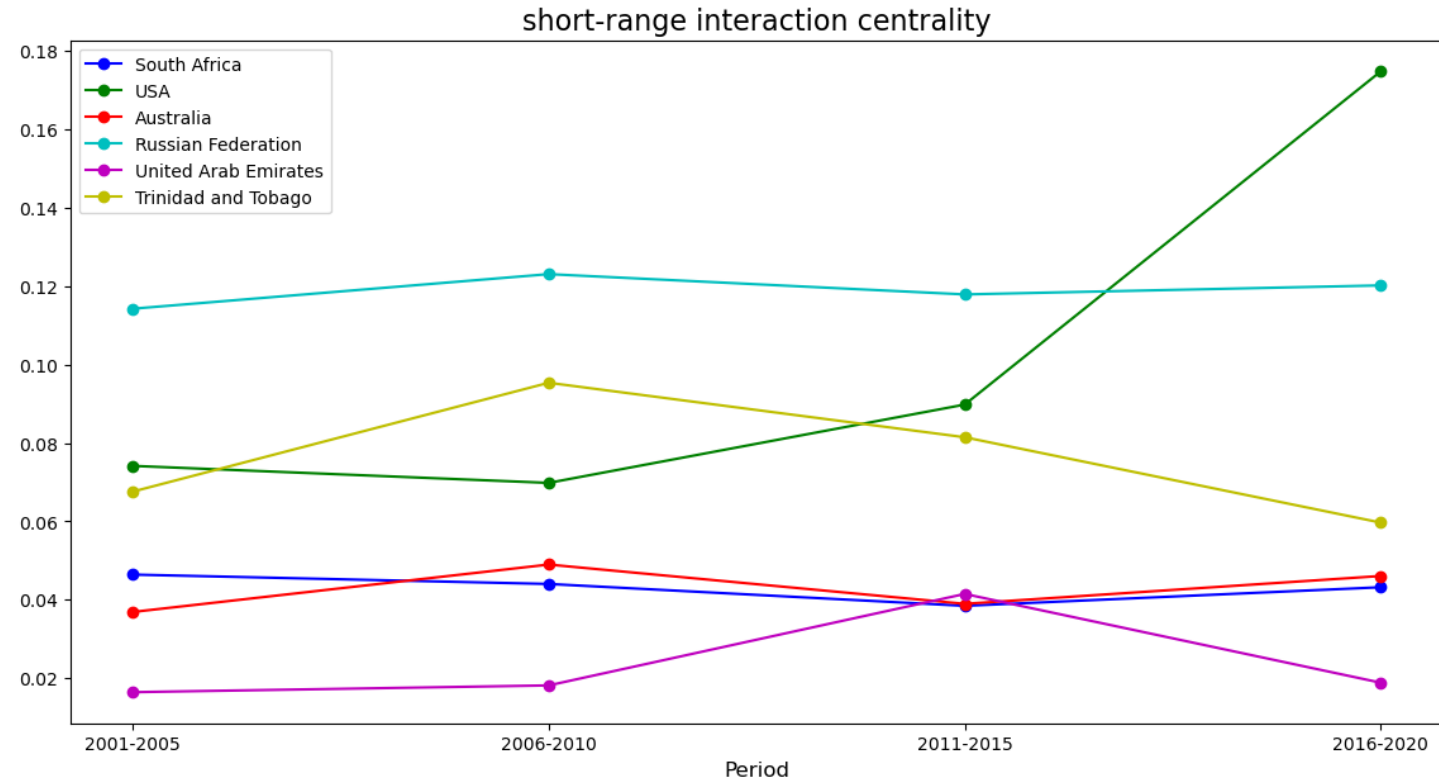


Сеть 2016-2020гг

Short-range interaction centrality

SRIC индекс отражающий важность вершины на основе влияния на соседние вершины, и, в отличие от других индексов, более точно показывает скрытые влияния в сети, но действует лишь на первом уровне. В качестве квоты было взято $q = 50\%$, поскольку эту долю будет достаточно сложно скомпенсировать с помощью других партнеров.

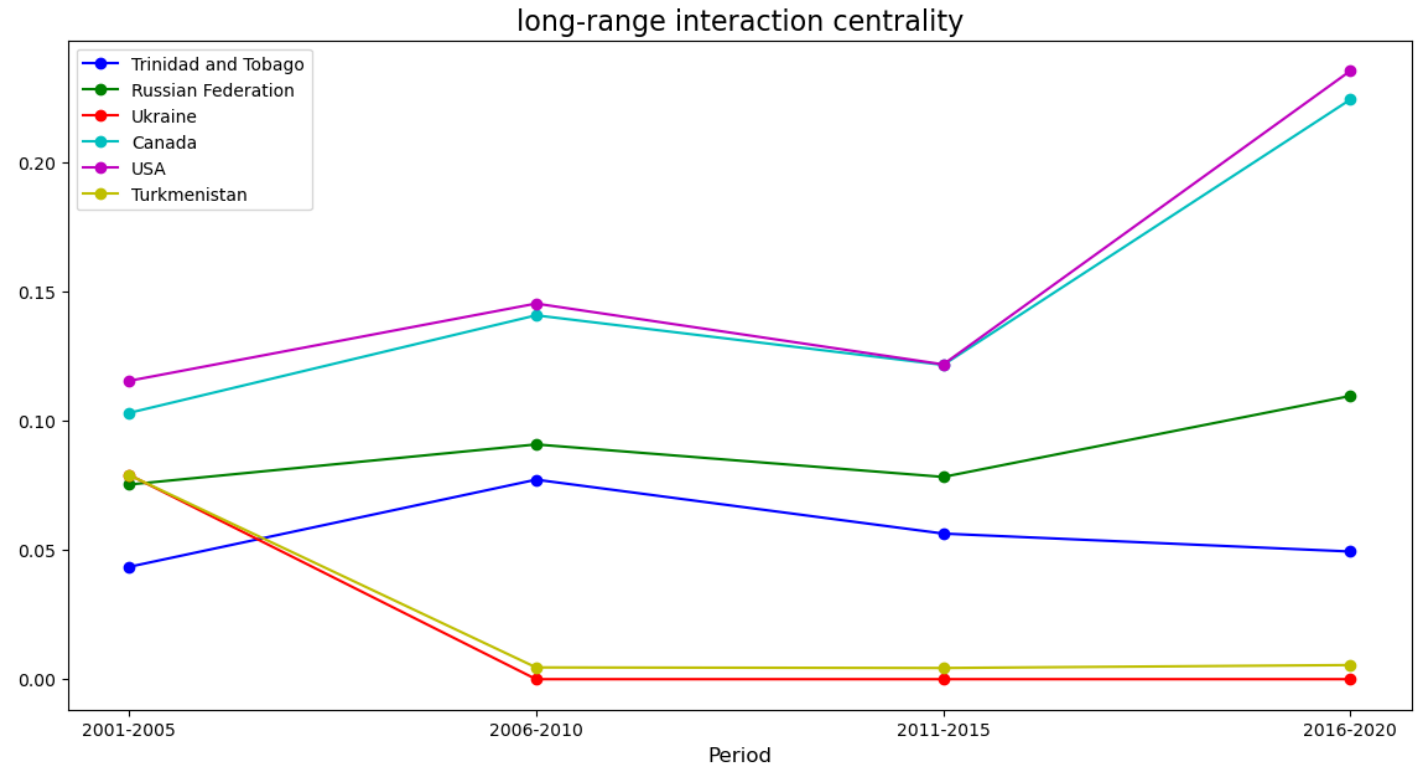
Как можно видеть, в топе стран по индексу SRIC оказались страны, которых не было в предыдущих индексах. Наиболее центральными странами на основе индекса SRIC оказались Южная Африка, США, Австралия, Тринидад и Табаго, Россия и ОАЭ



Long-range interaction centrality

LRIC индекс является усовершенствованной версией SRIC, но также учитывает дальние влияния. В качестве квоты было взято $q = 50\%$, модель MaxMin и максимальная длина пути равна 3.

Страны оказывающие наибольшее влияние оказались США, Канада, Россия, Тринидад и Тобаго, Туркменистан и Украина.





Выводы

Исходя из значения индекса LRIC, как наиболее подходящего для анализа, США занимает первое место по значимости на протяжении 20 лет (2001-2020гг), а Канада второе на протяжении 20 лет (2001-2020гг). Россия занимает 3 место по значимости на основе индекса LRIC с 2006г по 2020г. 4 место с 2006г по 2020г является страна Тринидад и Тобаго, что является очень неожиданным результатом. Таблица значения индекса LRIC выглядит следующим образом:

	country	2001-2005	2006-2010	2011-2015	2016-2020
130	USA	1.154315e-01	0.145357	0.121806	0.235437
167	Canada	1.030218e-01	0.140809	0.121608	0.224325
1	Russian Federation	7.531335e-02	0.090847	0.078267	0.109604
70	Trinidad and Tobago	4.344928e-02	0.077211	0.056345	0.049395
152	Mozambique	7.552327e-07	0.040874	0.039008	0.043743
115	Australia	3.951179e-02	0.040876	0.026400	0.038289
77	South Africa	3.160158e-02	0.040875	0.034674	0.038275
223	Germany	1.668541e-02	0.018202	0.017786	0.016616
22	China	4.789399e-03	0.022741	0.013021	0.016607
67	Belgium	4.776252e-03	0.040896	0.034926	0.016559

Значение индекса LRIC в период 2001-2020гг
(отсортированы по 2016-2020гг)



Список литературы

Freeman, L. C. (1979). Centrality in social networks conceptual clarification. В L. C. Freeman.

Fuad Aleskerov, N. M. (2016). CENTRALITY MEASURES IN NETWORKS BASED ON NODES ATTRIBUTES, LONG-RANGE INTERACTIONS AND GROUP INFLUENCE. В N. M. Fuad Aleskerov. препринт WP7/2016/04.

Rochat, Y. (2009). *Closeness Centrality Extended To Unconnected Graphs*.

Sergey Brin, L. P. (1998). The anatomy of a large-scale hypertextual Web search engine. В L. P. Sergey Brin.

