

UNIVERSIDAD AUTONOMA DE MADRID

ESCUELA POLITECNICA SUPERIOR



Grado en Ingeniería Informática

TRABAJO FIN DE GRADO

**ANÁLISIS DE LOS DATOS DE LOS PAPELES DE PANAMÁ
USANDO MÉTODOS DE I.A**

**Autor: Miguel Mateos Robles
Tutor: David Renato Domínguez Carreta**

Junio 2017

ANÁLISIS DE LOS DATOS DE LOS PAPELES DE PANAMÁ USANDO MÉTODOS DE I.A

AUTOR: Miguel Mateos Robles
TUTOR: David Renato Domínguez Carreta

Grupo de la EPS
Dpto. Ingeniería Informática
Escuela Politécnica Superior
Universidad Autónoma de Madrid
Junio de 2017

Resumen

Los papeles de Panamá se tratan de la mayor filtración de documentos de la historia. Estos documentos se encontraban en el despacho de abogados de Mossack Fonseca cuya principal función era de hacer de intermediario para la creación de sociedades offshore. Los documentos desvelan como miles de personas ocultan su patrimonio en este tipo de sociedades, situadas en paraísos fiscales. Los documentos filtrados tienen un tamaño de 2,6 terabytes y contienen información sobre 214.488 entidades offshore que conecta a un total de 208 territorios diferentes.

Este Trabajo Fin de Grado consiste en la recopilación de datos de los Papeles de Panamá, para ello se ha usado la base de datos de Offshore Leaks. Esta base de datos ha sido analizada y reestructurada, para facilitar la lectura de datos con el objetivo de construir una red de conexiones de países basada en las relaciones entre diferentes territorios que aparecen en estos papeles. También se ha buscado el patrón de emisión que utiliza cada país. Este patrón ha sido comparado entre los países aplicando el coeficiente de correlación de Spearman. El propósito de esta comparación era identificar si siguen un patrón similar los paraísos fiscales y países vecinos.

También se ha analizado el comportamiento de las apariciones de los países en los papeles. Este comportamiento se puede dividir en dos funciones distintas, ajuste potencial si tienen pocas apariciones y exponencial en caso contrario. Posteriormente se ha querido comprobar si existe correlación entre las apariciones con datos como población, superficie y PIB.

Por último, se ha calculado la distancia existente entre los países aplicando el algoritmo de Dijkstra, el cual, mide la distancia mínima entre dos nodos de una red.

Se ha utilizado de ejemplo a España y los países con mayor número de apariciones en los papeles a la hora de comparar y analizar datos.

Palabras clave

Papeles de Panamá, sociedades offshore, paraísos fiscales, red de conexiones entre países, coeficiente de correlación de Spearman, algoritmo de Dijkstra.

Abstract

The Panama papers deal with the largest leak of documents in the history. These documents were in the Panamanian law firm Mossack Fonseca whose main function was to act as intermediary to create offshore societies. The documents reveal how thousands of people hid their assets in companies located in tax haven. The 2,6 terabyte sized leaked documents containing information of 214,488 offshore entities which connect 208 different territories.

Panama papers which have been provided with the database of Offshore Leak. The database has been analyzed and used to do new tables and to make easier to read data. The main objective of the project has been creating a connection networks of the countries involved based on the relations between different territories that appear in these papers. We also have researched the emission pattern used in each country. These emission patterns, have been compared applying Spearman correlation coefficient. The purpose of this comparison was identifying a similar pattern between tax havens and neighboring countries

We have also analyzed the performance of countries appearances in the papers and it was possible to divide these performances into two different functions: potential fitting for the those with low appearances and exponential fitting otherwise. Afterwards, we wanted to check if there is a correlation between appearances and other data such as population, area and GDP.

Finally, we have applied the Dijkstra algorithm which calculates the minimum distance between two nodes, with the purpose of calculating which countries are closer according to the paper's relations.

Spain and the countries with the largest number of appearances in the Panama papers have been used as an example for the comparison and analysis of the data.

Keywords

Panama Papers, offshore societies, tax haven, connection networks, Spearman correlation coefficient, algorithm Dijkstra.

Agradecimientos

Quisiera agradecer en primer lugar a mis padres porque gracias a su esfuerzo he podido estudiar la carrera que deseaba.

También a todos los amigos que he hecho en la universidad que han estado conmigo todo este camino, junto a los profesores que hemos tenido durante la carrera por todas las enseñanzas recibidas.

Por último, a mis amigos de toda la vida con los que he pasado duros días de biblioteca.

INDICE DE CONTENIDOS

1	Introducción.....	1
1.1	Motivación.....	1
1.2	Objetivos.....	1
1.3	Organización de la memoria.....	2
2	Estado del arte	3
2.1	Corrupción.....	3
2.1.1	Papeles de Panamá.....	3
2.2	Inteligencia Artificial.....	4
2.2.1	Grafos	4
2.2.1	Dijkstra	5
2.3	Correlación entre datos.....	5
2.3.1	Coefficiente de correlación de Spearman	5
3	Diseño.....	7
3.1	Recopilación de datos.....	7
3.2	Construcción de la base de datos.....	9
3.3	Diagrama de flujo	12
4	Desarrollo	13
4.1	Introducción.....	13
4.2	Construcción de la red de conexiones entre países.....	13
4.3	Patrón diccionario.....	17
4.4	Algoritmo Dijkstra.....	18
4.5	Correlación de los datos	19
5	Integración, pruebas y resultados	21
5.1	Red de conexiones entre países	21
5.2	Resultados Dijkstra.....	25
5.3	Correlación de datos	25
5.3.1	Correlación patrón diccionario	25
5.3.2	Apariciones en los Papeles de Panamá y su correlación	27
5.4	Datos de España	29
5.5	Recopilación de datos.....	30
6	Conclusiones y trabajo futuro.....	31
6.1	Conclusiones.....	31
6.2	Trabajo futuro	31
	Referencias	32
	Glosario	33
	Anexos.....	I
	Colección de datos de los países.....	I
	Tabla resultados Dijkstra sin filtrar	II
	Tabla resultados Dijkstra sin filtrar simplificada	II
	Tabla resultados Dijkstra con filtro	III
	Correlación del patrón diccionario	IV

INDICE DE FIGURAS

FIGURA 1-1: GRAFO SIMPLE DE EJEMPLO.....	4
FIGURA 2-2: FÓRMULA COEFICIENTE DE CORRELACIÓN DE SPEARMAN	5
FIGURA 3-1: DIAGRAMA E-R OFFSHORE LEAKS	7
FIGURA 3-2: CONSULTA TODOS LOS CÓDIGOS.....	9
FIGURA 3-3: DIAGRAMA DE FLUJO	12
FIGURA 4-1: VIEW_RELACIONES_PAISES	14
FIGURA 4-2: CONSULTA RED DE CONEXIONES ENTRE PAÍSES SIN NORMALIZAR	14
FIGURA 4-3: FÓRMULA DE NORMALIZACIÓN RELACIONES PAÍSES.....	16
FIGURA 4-4: CONSULTA VALOR MÁXIMO DE CADA RELACIÓN DE LA RED DE CONEXIONES ENTRE PAÍSES	16
FIGURA 4-5: CONSULTA DICCIONARIO PAÍSES SIN NORMALIZAR	17
FIGURA 4-6: CONSULTA VALOR MÁXIMO DE CADA RELACIÓN DEL DICCIONARIO	17
FIGURA 4-7: FÓRMULA DE NORMALIZACIÓN DICCIONARIO PAÍSES.....	18
FIGURA 4-8: CONSULTA APARICIONES PAÍSES.....	19
FIGURA 5-1: CONEXIONES DE PAÍSES UMBRAL 3.....	21
FIGURA 5-2: CONEXIONES DE PAÍSES UMBRAL [2-3).....	22
FIGURA 5-3: CONEXIONES DE PAÍSES UMBRAL [1,5-2).....	23
FIGURA 5-4: CONEXIONES DE PAÍSES UMBRAL [1-1,5).....	23
FIGURA 5-5: CONEXIONES DE AGRUPACIONES	24
FIGURA 5-6: HISTOGRAMA APARICIONES PAÍSES	27
FIGURA 5-7: AJUSTE APARICIONES PAÍSES	28

INDICE DE TABLAS

TABLA 3-1 CODIGOS_PAISES	10
TABLA 3-2 AGRUPACION_ABREVIATURA.....	10
TABLA 3-3 AGRUPACIONES	11
TABLA 3-4 RELACIONES	11
TABLA 4-1 RELACIÓN ENTRE PAÍSES SIN NORMALIZAR	15
TABLA 4-2 VECTOR RELACIÓN ENTRE PAÍSES	16
TABLA 4-3 APARICIONES DICCIONARIO PAÍSES.....	17
TABLA 4-4 PATRÓN DICCIONARIO DE LOS PAÍSES	18
TABLA 5-1 CORRELACIÓN ENTRE PAÍSES CON MAYORES APARICIONES	26
TABLA 5-2 CORRELACIÓN APARICIONES.....	29

1 Introducción

1.1 Motivación

Los Papeles de Panamá se trata de uno de los grandes destapes de corrupción de nuestra época. En él aparecen grandes nombres como por ejemplo Leonel Messi, Pedro Almodóvar y Vladimir Putin.

La corrupción es uno de los problemas que más preocupa a la población actual, por ejemplo, continuamente en los medios de comunicación aparecen noticias relacionadas con este tema. Además, el tema de la corrupción tuvo una gran importancia en los debates políticos en las pasadas elecciones electorales, condicionando el voto de muchos ciudadanos.

Generalmente las grandes masas de dinero defraudadas han utilizado paraísos fiscales, con el objetivo de hacer pantalla y ocultar al verdadero titular de una cuenta que esté blanqueando dinero, invadiendo impuestos, etc. Para ello muchos de ellos han hecho uso de sociedades offshore que consisten en empresas constituidas en un país diferente de aquel en el que desarrolla su actividad. Muchas de estas empresas y encargados que participan en ellas se encuentran reflejadas en los Papeles de Panamá.

Por lo tanto, este tipo de corrupción se encuentra oculta, generalmente en los paraísos fiscales, por eso se pretende realizar una red de conexiones entre los países para ver las relaciones existentes entre diferentes regiones.

1.2 Objetivos

El primer objetivo planteado consiste en la recopilación de datos de los Papeles de Panamá en Internet, la principal fuente de información se ha conseguido de la web Offshore Leaks, que contiene una “base de datos del Consorcio Internacional de Investigación de Periodistas (ICIJ) con información sobre casi 500.000 entidades extraterritoriales que forman parte de las investigaciones de "Panama Papers", "Offshore Leaks" y las investigaciones de "Bahamas Leaks". Los datos cubren casi 40 años hasta principios de 2016 con enlaces a personas y empresas en más de 200 países y territorios”^[1].

También se han conseguido datos del Banco Mundial para comprobar si existe correlaciones entre las apariciones de los países en los Papeles de Panamá, con datos como el PIB, población, superficie.

Con todos estos datos se ha construido una base de datos con el objetivo principal del TFG, el cual es, realización de una red de conexión de los países mediante las relaciones en los Papeles de Panamá. Para el análisis de la red se ha decidido aplicar el algoritmo de Dijkstra que mide el camino mínimo entre dos nodos de una red, para comprobar la conectividad existente entre ellos.

Además de la red, se ha obtenido un patrón que nos indica el tipo de relaciones que utiliza cada país cuando se trata del emisor, al que se le ha denominado patrón diccionario. Se ha comparado el patrón de estos países utilizando el coeficiente de correlación de Spearman.

1.3 Organización de la memoria

La memoria consta de los siguientes capítulos:

- **Capítulo 1 (Introducción):** en este punto se encuentra las motivaciones del TFG y los objetivos planteados.
- **Capítulo 2 (Estado del Arte):** en este apartado se dará información de los aspectos relacionados con el TFG para poder comprender mejor las tareas realizadas.
- **Capítulo 3 (Diseño):** en este capítulo, se explicará cómo se han obtenido los datos, además de las decisiones tomadas para la realización de la base de datos de la cual se ha basado el TFG. También se incluye un diagrama de flujo que muestra el procedimiento de cómo se ha trabajado.
- **Capítulo 4 (Desarrollo):** en este apartado se detallarán los aspectos más importantes a la hora de obtener los datos. Con estos datos se ha conseguido realizar una red de conexiones de países y agrupaciones de los Papeles de Panamá. También se ha obtenido el patrón de emisión de cada uno de ellos. Además, se comentan los detalles de cómo se ha realizado la correlación de los datos y la aplicación del algoritmo de Dijkstra, para calcular los caminos mínimos entre los países.
- **Capítulo 5 (Integración, pruebas y resultados):** en este punto se comentarán resultados obtenidos de cada una de las pruebas realizadas.
- **Capítulo 6 (Conclusiones y trabajo futuro):** para finalizar se comentarán conclusiones propias del proyecto realizado además de posibles trabajos futuros que se pueden realizar a partir de los datos obtenidos.

2 Estado del arte

2.1 Corrupción

La corrupción se define como un acto de un individuo inapropiado respecto a unos valores sociales, para obtener un tipo de beneficio propio.

La corrupción se encuentra muy incrustada en la sociedad actual, pudiéndola encontrar en un gran abanico de ámbitos diferentes, como puede ser en política, policial, comercial, etc...Donde las personas aprovechan su situación actual en la que se encuentren para obtener una mayor ganancia de ello.

Un tipo de corrupción consiste en el impago de impuestos, muchas de las veces, se consigue este propósito gracias a las fiscalías tributarias de otros países, donde los impuestos son menores. Este tipo de movimientos es uno de los motivos del gran número de documentos que aparecen en los papeles de Panamá. Disponer tus cuentas en este tipo de países, también puede realizarse de manera lícita.

2.1.1 Papeles de Panamá

Los Papeles de Panamá se trata de "una filtración de más de 11,5 millones de documentos (2,6 TB) extraídos en 2015 de los servidores de Mossack Fonseca, un bufete de abogados panameño dedicado principalmente a la gestión de sociedades offshore. Aunque la oficina principal es panameña, se trata de una firma internacional en la que trabajan más de 600 personas en delegaciones para 42 países" [3].

En los papeles de Panamá se pueden encontrar un gran número de personajes famosos, además, de las grandes cantidades de dinero que se ha movido a las sociedades offshore ha hecho que esta filtración tenga un gran impacto en la sociedad actual.

“Sociedad offshore es una empresa constituida en un país diferente de aquel en el que desarrolla su actividad. Constituir una empresa de ese tipo no es ilegal” [3].

El despacho de Mossack Fonseca se encargaba de hacer de intermediario de estas entidades offshore, para facilitar su creación. Las jurisdicciones de estas sociedades no se encuentran únicamente en Panamá, de hecho, están distribuidas entre diferentes paraísos fiscales.

“La base de datos de Mossack Fonseca fue entregada por un denunciante anónimo al periódico alemán Süddeutsche Zeitung. La magnitud de la base de datos hizo que el diario decidiera compartir el descubrimiento con el Consorcio Internacional de Periodistas de Investigación (ICIJ). La organización reclutó a un equipo internacional de 370 periodistas pertenecientes a un centenar de medios de comunicación de 76 países” [3]. Del trabajo de estos periodistas se ha conseguido crear una base de datos con la que se ha realizado este TFG.

2.2 Inteligencia Artificial

La inteligencia artificial se define como métodos cuya solución son aproximados para problemas complejos minimizando al máximo el error en sus procedimientos.

2.2.1 Grafos

Un grafo consiste en un conjunto de nodos (vértices) que son unidos mediante enlaces (aristas) que permiten establecer relaciones binarias entre ellos.

Se representan de la siguiente manera $G(V, E)$:

V = conjunto de nodos

E = conjunto de aristas

Las aristas generalmente disponen de un peso W que consiste en el coste de unir ambos nodos.

Un grafo dirigido consiste en un grafo donde las aristas indican el nodo emisor y el de destino de cada enlace. Generalmente se representan mediante su lista de adyacencias, un nodo adyacente a otro consiste en un nodo que dispone de conexión directa al otro mediante una arista.

Por ejemplo, el siguiente grafo se podría representar de la siguiente forma:

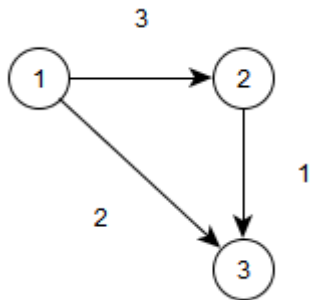


Figura 1-1: Grafo simple de ejemplo

$V = (1, 2, 3)$

$E = ($

1	2	3
1	3	2
2	3	1
Nodo	Nodo	Peso
Emisor	Destino	

)

Los grafos han servido para representar problemas complejos y dar solución a ellos, uno de los más conocidos se trata del problema del viajante. Este problema consiste en una lista de ciudades y las distancias que existen entre ellas si se encuentran conectadas. Se plantea la

siguiente pregunta, ¿cuál es la ruta más corta posible para visitar todas las ciudades y regresar a la ciudad origen pasando una única vez por cada ciudad?

Este problema es uno de los más estudiados en optimización combinatoria y puede que no exista solución dependiendo del problema planteado. En el problema se aplican “heurísticas cuyo objetivo es generar soluciones de buena calidad en tiempos de cómputo pequeños” [4]. Este tipo de problemas se puede reflejar en un gran número de ámbitos diferentes como: economía, circuitos electrónicos, etc...

2.2.1 Dijkstra

El algoritmo de Dijkstra, encuentra el camino mínimo entre dos nodos de una red. A continuación, se muestra su pseudocódigo:

```
Dijkstra(n, G):  
v[ ] = F; p[ ] = NULL  
iniPQ(Q)  
distancia[n]=0; addPQ(n,Q)  
while emptyPQ(Q)≠F:  
    extPQ(w,Q)  
    if v[w]≠F:  
        v[w]=T  
        for all z adjacent to w:  
            p[z]=w; addPQ(z,Q)
```

El algoritmo explora las conexiones que le van surgiendo desde el nodo origen y marcándolas como vistas una vez exploradas, si la distancia del nodo origen al nodo destino es menor que la distancia actual del destino, se sustituye por el valor que sea más pequeño.

2.3 Correlación entre datos

La correlación mide la relación que existe entre dos vectores. Según su variación, existen diferentes tipos de correlación:

La correlación es directa cuando un valor en un punto crece en un vector y en el otro sigue el mismo comportamiento, inversa cuando tienen comportamiento distinto en el mismo punto o nula cuando no existe ningún tipo de dependencia entre ellos. El coeficiente de correlación de Spearman mide el grado de dependencia entre dos vectores.

2.3.1 Coeficiente de correlación de Spearman

El coeficiente de correlación de Spearman ρ , consiste en una medida de la relación lineal entre dos variables aleatorias cuantitativas. Se calcula de la siguiente forma:

$$\rho = 1 - \frac{6 \sum d^2}{n(n^2 - 1)}$$

Figura 2-2: fórmula coeficiente de correlación de Spearman

Donde d se trata de la diferencia entre los rangos (X menos Y) que se obtiene ordenando los datos y n es el número de datos.

El coeficiente ρ varía entre $[-1,1]$, cuanto más cerca se encuentre de 1 nos indicará que existe una relación directa, si es próximo a -1 se tratará de una relación inversa y cuanto más se aproxime a 0 nos indicará la ausencia de relación entre los datos.

Se puede interpretar de la siguiente forma:

- 1) Perfecta $\rho = 1$
- 2) Excelente $\rho = 0.9 \leq \rho < 1$
- 3) Buena $\rho = 0.8 \leq \rho < 0.9$
- 4) Regular $\rho = 0.5 \leq \rho < 0.8$
- 5) Mala $\rho < 0.5$

3 Diseño

3.1 Recopilación de datos

Los datos para la creación de la base de datos de los Papeles de Panamá, se han obtenido principalmente de la página web Offshore Leaks. En esta web se han descargado 5 ficheros en formato CSV que contiene las relaciones de los archivos encontrados en estos papeles. Estos ficheros se encuentran divididos en entities (entidades), intermediaries (intermediarios), officers (encargados), addresses (direcciones) y otro fichero all_edges (conexiones) que relaciona todos los anteriores.

Se ha decidido crear una base de datos en postgresQL, siguiendo el siguiente diagrama E-R, al que se le añadirán nuevas tablas como se indica en el apartado 3.2 Construcción de la base de datos.

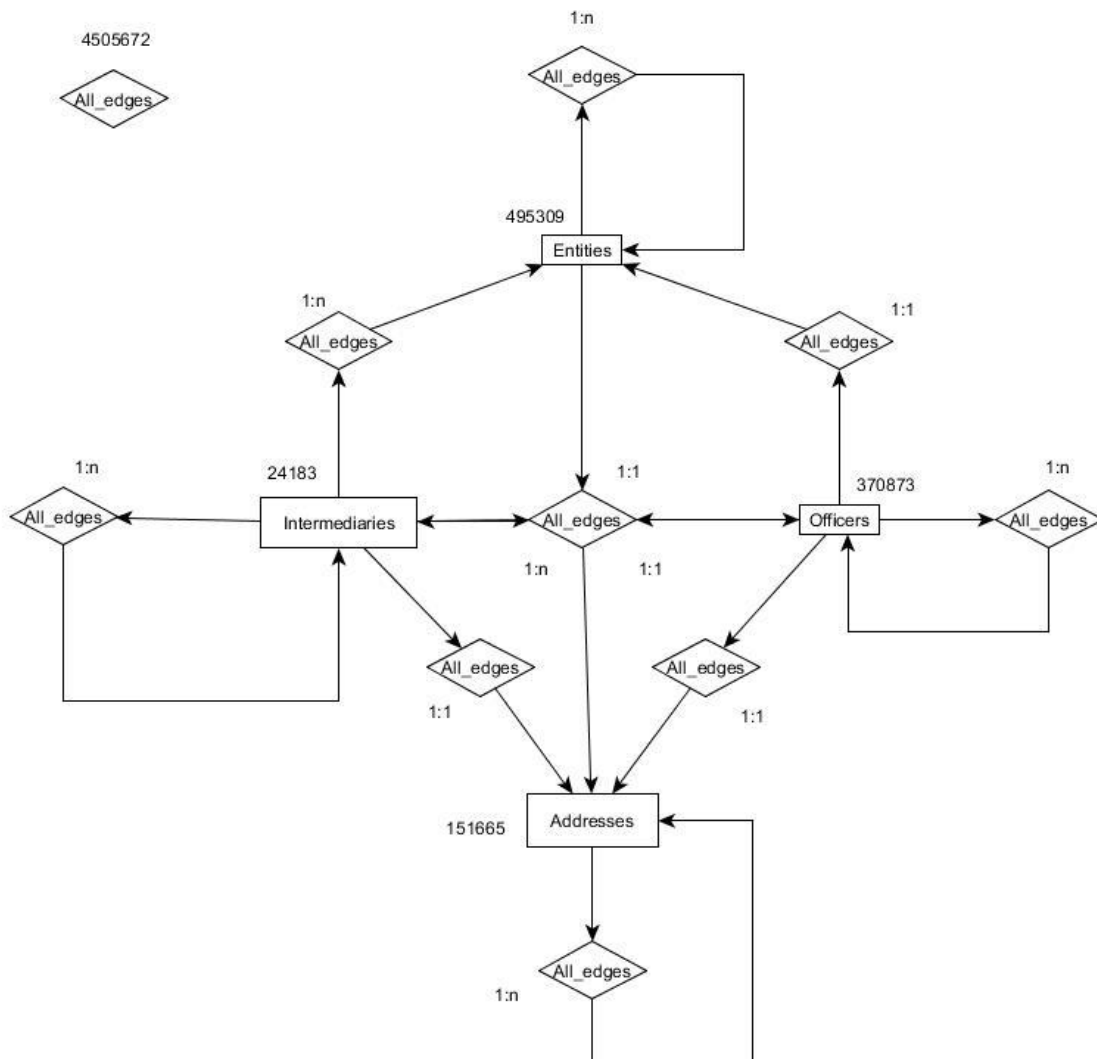


Figura 3-1: diagrama E-R Offshore Leaks

Entidades, corresponde a una empresa o un fondo creado en jurisdicción de baja tributación en el extranjero por un encargado. Hay un total de 495.309 entidades. Está conformada por los siguientes atributos: name, original_name, former_name, jurisdiction, jurisdiction_description, company_type, address. Internal_id, incorporattion_date, inactivation_date, struck_off_date, dorm_date, status, service_provider, ibcRUC, countr_codes, countries, note, valid_until, node_id, sourceID.

Entidades está relacionada consigo misma por similitudes en direcciones o nombres de otras entidades (1 entidad se relaciona con n entidades) y con direcciones donde se indica su dirección (1 entidad se relaciona con 1 dirección). Estas relaciones se consiguen mediante la relación Conexiones (All_edges) con su primary key node_id.

Intermediarios, es una relación entre alguien que busca una corporación offshore y un proveedor de servicios offshore, normalmente una firma de abogados o un intermediario proveedor de este tipo de servicios. El objetivo de los intermediarios es de crear una firma offshore para un cliente. Hay un total de 24.183 intermediarios. Está conformada por los siguientes atributos: name, internal_id, address, valid_until, country_codes, countries, status, node_id, sourceID, note.

Intermediarios está relacionada consigo misma por similitudes en direcciones o nombres de otros intermediarios (1 intermediario se relaciona con n intermediarios), con entidades y encargados debido a que el intermediario puede tratarse de uno de ellos (1 intermediario se relaciona con n agentes o n encargados) y con direcciones donde se indica su dirección (1 intermediario se relaciona con 1 dirección). Estas relaciones se consiguen mediante la relación Conexiones (All_edges) con su primary key node_id.

Encargados, se trata de una persona o compañía que desempeña un papel en una entidad offshore. Hay un total de 370.873 encargados. Está conformada por los siguientes atributos: name, icij_id, valid_until, country_codes, countries, node_id, sourceID, note.

Encargados está relacionada consigo misma por la correspondencia con otros encargados (1 encargado se relaciona con n encargados), su relación con Entidades e Intermediarios se debe a que son encargados en estas empresas (1 encargado se relaciona con 1 entidad o 1 intermediario) y con direcciones donde se indica su dirección (1 encargado se relaciona con 1 dirección). Estas relaciones se consiguen mediante la relación Conexiones (All_edges) con su primary key node_id.

Direcciones, se tratan de direcciones postales que aparecen originalmente en la base de datos. Hay un total de 151.665 direcciones. Está conformada por los siguientes atributos: address, icij_id, valid_until, country_codes, countries, node_id, sourceID, note.

Direcciones está relacionada consigo misma por similitudes con otras direcciones (1 dirección se relaciona con n direcciones), además el resto de entidades se relaciona con Direcciones mediante la relación Conexiones (All_edges) con su primary key node_id.

La relación Conexiones, contiene todas las relaciones existentes entre las entidades. Hay un total de 4.505.672 conexiones. Está formada por los siguientes atributos: node_1, rel_type, node_2, sourceID, valid_until, start_date, end_date, donde node_1 hace referencia a la entidad emisora y nodo_2 a la entidad que recibe la acción por la que se conecta con el nodo_1.

Con esta recopilación de datos se ha creado una base de datos como se indica en la Figura 3-1: diagrama E-R Offshore Leaks.

Es importante señalar que todas las entidades disponen del atributo countries, pero muchos de estos códigos se encuentran con valores nulos o sin identificar.

Para poder realizar las consultas donde la entidad del nodo 1 se trata de la misma entidad del nodo 2, se ha decidido duplicar las tablas de direcciones, entidades, encargados e intermediarios. Se han duplicado estas tablas con el objetivo de poder asignar cada una a un nodo, por ejemplo, un intermediario que tiene la misma dirección que otro intermediario se realizaría de la siguiente manera: al nodo 1 se le asignaría la tabla intermediaries y al nodo 2 se le asignaría la tabla duplicada intermediaries _aux.

Además de los datos de Offshore Leaks, se ha complementado esta información con los datos del Banco Mundial, donde se pueden obtener el PIB, población, superficie de los países.

3.2 Construcción de la base de datos

La base de datos mencionada anteriormente ha sido necesario modificarla añadiéndole nuevas tablas para facilitar la lectura de los datos. A continuación, se detallarán el motivo y la manera en la que se han creado las nuevas tablas.

A la hora de realizar consultas se comprobó que pueden aparecer varios países en el atributo countries en un mismo nodo. Esto significa que el nodo emisor o destino pertenece a todos los países que aparezcan en esa tupla, por ejemplo, una entidad que se encuentra en Tailandia y Hong Kong se representará de la siguiente manera, countries: Thailand;Hong Kong. Esto complicaba la construcción de la red de conexiones entre países, por lo que se ha decidido crear una nueva tabla Codigos_paises. Esta tabla está formada por estas tuplas duplicadas por tantos países aparezcan en ella, pero añadiéndole un atributo clasificación, donde en este atributo aparece un único país. De esta manera permanece la información intacta y se facilita la lectura de datos. Este proceso se ha realizado de la siguiente manera:

- Se ha realizado la siguiente consulta que detecta todos los códigos de los países que aparecen en la base de datos.

```
136 L-- Consulta para averiguar todos los country_codes diferentes que aparecen en los papeles de panama
137 COPY(
138     SELECT country_codes, countries FROM (
139         (SELECT DISTINCT addr_country_codes as country_codes, addr_countries as countries FROM Addresses) UNION
140         (SELECT DISTINCT ent_country_codes, ent_countries FROM Entities) UNION
141         (SELECT DISTINCT int_country_codes, int_countries FROM Intermediaries) UNION
142         (SELECT DISTINCT ofi_country_codes, ofi_countries FROM Officers)
143     )AS ordenado ORDER BY country_codes
144 )TO '/tmp/consulta.csv' WITH DELIMITER ',' csv header;
```

Figura 3-2: consulta todos los códigos

La consulta consiste en seleccionar los códigos de todas las entidades de la base de datos mediante UNION.

- Los datos son leídos por un programa en java donde se duplican las tuplas pero se clasifican con un único país como se muestra en la siguiente tabla:

Codigos_largo	Paises_largo	Codigo	Pais
THA;HKG	Thailand;Hong Kong	THA	Thailand
THA;HKG	Thailand;Hong Kong	HKG	Hong Kong
...

Tabla 3-1 codigos_paises

De esta manera se consigue la Tabla Codigos_paises con primary key (codigos_largo, código) que se relaciona con el resto de entidades mediante el atributo codigos_largo.

Así mismo, ha sido necesario duplicar esta tabla para poder asignar una a la entidad del nodo emisor y la otra a la del nodo destino.

De este modo se puede identificar de una forma más clara todas las relaciones que existen en la base de datos.

También se ha decidido realizar otra tabla donde se le asigna a los códigos de los países una región geográfica para de esta manera ver los resultados de una forma más global. Estas regiones son las siguientes:

nombre_agrupacion	codigo_agrupacion
África Magreb	AFM
África Subsahariana	AFS
América Antillas	AMA
América Central	AMC
América del Norte	AMN
América del Sur	AMS
Asia del Sur	ASS
Asia Oriental	ASO
Asia Sudeste	ASE
Europa Eslava	EUS
Europa Germana	EUG
Europa Latina	EUL
Oceanía	OCE
Oriente Medio	ORM
Sin identificar	XXX

Tabla 3-2 agrupacion_abreviatura

A la hora de realizar las agrupaciones también se encontraba el problema de varios códigos que pueden corresponderse a agrupaciones diferentes. Para solucionarlo se ha hecho un proceso similar al anterior, duplicando estas tuplas y añadiéndolas una clasificación diferente. Este proceso se ha realizado a mano por cada uno de los 716 códigos de países que aparecen.

De esta manera se ha conseguido la tabla agrupaciones donde aparecen los países con sus respectivas agrupaciones.

codigo_pais	nombre_pais	codigo_agrupacion	agrupación
ABW	Aruba	AMA	América Antillas
AND	Andorra	EUL	Europa Latina
AGO	Angola	AFS	África Subsahariana
ALB	Albania	EUS	Europa Eslava
...

Tabla 3-3 agrupaciones

Por ultimo también ha sido necesario crear la tabla relaciones formada por los diferentes tipos de conexión existentes en la base de datos. En la base de datos se encontraban un total de 274 Relaciones (SELECT DISTINCT con_rel_type FROM All_edges), las cuales se han agrupado en únicamente 19 por ser semejantes. Las 19 relaciones se encuentran en el glosario de la memoria. Se decidió crear la relación others para añadir las relaciones que tenían un menor número de apariciones.

Para agrupar las relaciones se ha realizado un proceso similar al de los códigos de los países, duplicando las tuplas y añadiéndoles una clasificación única. Este proceso se ha realizado a mano.

La tabla inferior se trata de una pequeña muestra de la tabla relaciones donde se ven tuplas duplicadas por tener dos clasificaciones diferentes. Muchas de las relaciones consistían en la unión de abreviaciones de estas, por ejemplo, sec,pre, abreviación de secretary y president.

conexión	conexion_clasificada
same name as	same name as
intermediary of	intermediary of
shareholder of	shareholder
President,secretary	President
President,secretary	Secretary
dir. and sec.	Director
dir. and sec.	Secretary
...	...

Tabla 3-4 relaciones

3.3 Diagrama de flujo

En este apartado se muestra el procedimiento de cómo se ha trabajado a lo largo del TFG.

El primer paso, fue la creación de la base de datos como se ha explicado en el punto anterior. Con la base de datos creada el siguiente paso ha sido la creación de consultas y vistas para la obtención de los datos.

Los datos de las consultas serán procesados por diferentes programas en java. Estos programas siguen la siguiente estructura: lectura de datos de la consulta, transformación e interpretación de los datos, volcar en un fichero los resultados obtenidos.

La función de los programas ha sido la de construir la red de conexiones entre países y agrupaciones, además de lo que se ha llamado patrón diccionario de cada país que se explicará en el punto 4. Desarrollo. También hay un programa que comprueba si las relaciones son globales o locales, es decir si las relaciones se encuentran dentro de su agrupación o no.

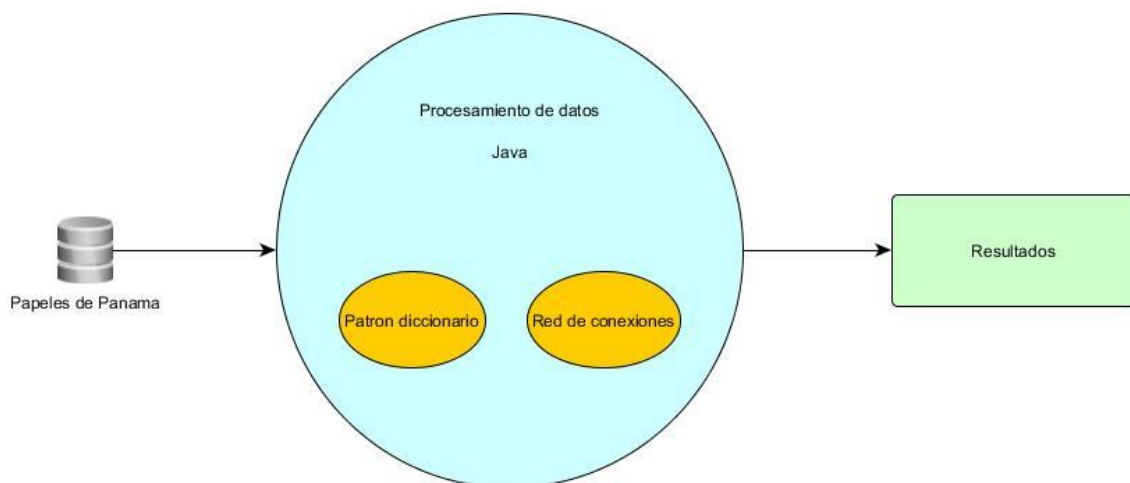


Figura 3-3: diagrama de flujo

La red de conexiones de países, posteriormente ha sido leída por un programa en Python que aplicará el algoritmo de Dijkstra. Este programa crea una matriz de distancia entre los países. También se ha utilizado un programa en R que realiza la correlación entre los patrones diccionario de cada país.

4 Desarrollo

4.1 Introducción

En este apartado se explicará cómo se han realizado las consultas SQL y los detalles más importantes de los programas realizados. Se ha decidido realizar en java los programas de procesamiento de datos, construcción de la red de países y del patrón diccionario; Python para el algoritmo de Dijkstra y en R se ha realizado un programa para comprobar si existen correlaciones entre los datos.

4.2 Construcción de la red de conexiones entre países

La red de conexiones entre países relaciona todos los pares de países según las apariciones que se encuentran en los Papeles de Panamá. Esta conectividad se muestra con un peso que se ha calculado normalizando las apariciones de cada tipo de relación entre los países. A continuación, se detalla todo este proceso:

En primer lugar, es necesario realizar una consulta a la base de datos que muestre todas las relaciones entre los países y el número de veces que se repiten. Al tratarse de una red de países, únicamente se han cogido las relaciones que unen dos países, es decir, se han omitido las relaciones con países NULL o se encuentran sin identificar (código XXX).

Se ha decidido realizar una vista en la base de datos en la que aparecen todos estos datos, para poder realizar consultas sobre ella.

```

1  -- Relaciones entre paises
2  DROP VIEW IF EXISTS view_relaciones_paises;
3  CREATE VIEW view_relaciones_paises AS
4  SELECT codigo, pais, conexion_clasificada, codigo_2, pais_2, SUM(count) AS apariciones FROM (
5  (SELECT Codigos_paises.codigo, Codigos_paises.pais, conexion_clasificada, Codigos_paises_2.codigo_2, Codigos_paises_2.pais_2, count(*)
6  FROM Officers INNER JOIN All_edges ON ofi_node_id = con_node_1
7  INNER JOIN Intermediaries ON int_node_id = con_node_2
8  LEFT JOIN Relaciones ON con_rel_type = conexion
9  LEFT JOIN codigos_paises ON codigo_largo = ofi_country_codes
10 LEFT JOIN codigos_paises_2 ON codigo_largo_2 = int_country_codes
11 GROUP BY Codigos_paises.codigo, Codigos_paises.pais, conexion_clasificada, Codigos_paises_2.codigo_2, Codigos_paises_2.pais_2)
12 UNION ALL
13 (SELECT Codigos_paises.codigo, Codigos_paises.pais, conexion_clasificada, Codigos_paises_2.codigo_2, Codigos_paises_2.pais_2, count(*)
14 FROM Officers INNER JOIN All_edges ON ofi_node_id = con_node_1
15 INNER JOIN Entities ON ent_node_id = con_node_2
16 LEFT JOIN Relaciones ON con_rel_type = conexion
17 LEFT JOIN codigos_paises ON codigo_largo = ofi_country_codes
18 LEFT JOIN codigos_paises_2 ON codigo_largo_2 = ent_country_codes
19 GROUP BY Codigos_paises.codigo, Codigos_paises.pais, conexion_clasificada, Codigos_paises_2.codigo_2, Codigos_paises_2.pais_2)
20 UNION ALL
21 (SELECT Codigos_paises.codigo, Codigos_paises.pais, conexion_clasificada, Codigos_paises_2.codigo_2, Codigos_paises_2.pais_2, count(*)
22 FROM Officers INNER JOIN All_edges ON ofi_node_id = con_node_1
23 INNER JOIN Intermediaries ON int_node_id = con_node_2
24 LEFT JOIN Relaciones ON con_rel_type = conexion
25 LEFT JOIN codigos_paises ON codigo_largo = ofi_country_codes
26 LEFT JOIN codigos_paises_2 ON codigo_largo_2 = int_country_codes
27 GROUP BY Codigos_paises.codigo, Codigos_paises.pais, conexion_clasificada, Codigos_paises_2.codigo_2, Codigos_paises_2.pais_2)
28 UNION ALL
29 (SELECT Codigos_paises.codigo, Codigos_paises.pais, conexion_clasificada, Codigos_paises_2.codigo_2, Codigos_paises_2.pais_2, count(*)
30 FROM Officers INNER JOIN All_edges ON ofi_node_id = con_node_1
31 INNER JOIN Entities ON ent_node_id = con_node_2
32 LEFT JOIN Relaciones ON con_rel_type = conexion
33 LEFT JOIN codigos_paises ON codigo_largo = ofi_country_codes
34 LEFT JOIN codigos_paises_2 ON codigo_largo_2 = ent_country_codes
35 GROUP BY Codigos_paises.codigo, Codigos_paises.pais, conexion_clasificada, Codigos_paises_2.codigo_2, Codigos_paises_2.pais_2)
36 UNION ALL
37 (SELECT Codigos_paises.codigo, Codigos_paises.pais, conexion_clasificada, Codigos_paises_2.codigo_2, Codigos_paises_2.pais_2, count(*)
38 FROM Officers INNER JOIN All_edges ON ofi_node_id = con_node_1
39 INNER JOIN Intermediaries ON int_node_id = con_node_2
40 LEFT JOIN Relaciones ON con_rel_type = conexion
41 LEFT JOIN codigos_paises ON codigo_largo = ofi_country_codes
42 LEFT JOIN codigos_paises_2 ON codigo_largo_2 = int_country_codes
43 GROUP BY Codigos_paises.codigo, Codigos_paises.pais, conexion_clasificada, Codigos_paises_2.codigo_2, Codigos_paises_2.pais_2)
44 UNION ALL
45 (SELECT Codigos_paises.codigo, Codigos_paises.pais, conexion_clasificada, Codigos_paises_2.codigo_2, Codigos_paises_2.pais_2, count(*)
46 FROM Officers INNER JOIN All_edges ON ofi_node_id = con_node_1
47 INNER JOIN Entities ON ent_node_id = con_node_2
48 LEFT JOIN Relaciones ON con_rel_type = conexion
49 LEFT JOIN codigos_paises ON codigo_largo = ofi_country_codes
50 LEFT JOIN codigos_paises_2 ON codigo_largo_2 = ent_country_codes
51 GROUP BY Codigos_paises.codigo, Codigos_paises.pais, conexion_clasificada, Codigos_paises_2.codigo_2, Codigos_paises_2.pais_2)
52 UNION ALL
53 (SELECT Codigos_paises.codigo, Codigos_paises.pais, conexion_clasificada, Codigos_paises_2.codigo_2, Codigos_paises_2.pais_2, count(*)
54 FROM Officers INNER JOIN All_edges ON ofi_node_id = con_node_1
55 INNER JOIN Intermediaries ON int_node_id = con_node_2
56 LEFT JOIN Relaciones ON con_rel_type = conexion
57 LEFT JOIN codigos_paises ON codigo_largo = ofi_country_codes
58 LEFT JOIN codigos_paises_2 ON codigo_largo_2 = int_country_codes
59 GROUP BY Codigos_paises.codigo, Codigos_paises.pais, conexion_clasificada, Codigos_paises_2.codigo_2, Codigos_paises_2.pais_2)
60 UNION ALL
61 (SELECT Codigos_paises.codigo, Codigos_paises.pais, conexion_clasificada, Codigos_paises_2.codigo_2, Codigos_paises_2.pais_2, count(*)
62 FROM Officers INNER JOIN All_edges ON ofi_node_id = con_node_1
63 INNER JOIN Entities ON ent_node_id = con_node_2
64 LEFT JOIN Relaciones ON con_rel_type = conexion
65 LEFT JOIN codigos_paises ON codigo_largo = ofi_country_codes
66 LEFT JOIN codigos_paises_2 ON codigo_largo_2 = ent_country_codes
67 GROUP BY Codigos_paises.codigo, Codigos_paises.pais, conexion_clasificada, Codigos_paises_2.codigo_2, Codigos_paises_2.pais_2)
68 UNION ALL
69 (SELECT Codigos_paises.codigo, Codigos_paises.pais, conexion_clasificada, Codigos_paises_2.codigo_2, Codigos_paises_2.pais_2, count(*)
70 FROM Officers INNER JOIN All_edges ON ofi_node_id = con_node_1
71 INNER JOIN Intermediaries ON int_node_id = con_node_2
72 LEFT JOIN Relaciones ON con_rel_type = conexion
73 LEFT JOIN codigos_paises ON codigo_largo = ofi_country_codes
74 LEFT JOIN codigos_paises_2 ON codigo_largo_2 = int_country_codes
75 GROUP BY Codigos_paises.codigo, Codigos_paises.pais, conexion_clasificada, Codigos_paises_2.codigo_2, Codigos_paises_2.pais_2)
76 UNION ALL
77 (SELECT Codigos_paises.codigo, Codigos_paises.pais, conexion_clasificada, Codigos_paises_2.codigo_2, Codigos_paises_2.pais_2, count(*)
78 FROM Officers INNER JOIN All_edges ON ofi_node_id = con_node_1
79 INNER JOIN Entities ON ent_node_id = con_node_2
80 LEFT JOIN Relaciones ON con_rel_type = conexion
81 LEFT JOIN codigos_paises ON codigo_largo = ofi_country_codes
82 LEFT JOIN codigos_paises_2 ON codigo_largo_2 = ent_country_codes
83 GROUP BY Codigos_paises.codigo, Codigos_paises.pais, conexion_clasificada, Codigos_paises_2.codigo_2, Codigos_paises_2.pais_2)
84 UNION ALL
85 (SELECT Codigos_paises.codigo, Codigos_paises.pais, conexion_clasificada, Codigos_paises_2.codigo_2, Codigos_paises_2.pais_2, count(*)
86 FROM Officers INNER JOIN All_edges ON ofi_node_id = con_node_1
87 INNER JOIN Intermediaries ON int_node_id = con_node_2
88 LEFT JOIN Relaciones ON con_rel_type = conexion
89 LEFT JOIN codigos_paises ON codigo_largo = ofi_country_codes
90 LEFT JOIN codigos_paises_2 ON codigo_largo_2 = int_country_codes
91 GROUP BY Codigos_paises.codigo, Codigos_paises.pais, conexion_clasificada, Codigos_paises_2.codigo_2, Codigos_paises_2.pais_2)
92 UNION ALL
93 (SELECT Codigos_paises.codigo, Codigos_paises.pais, conexion_clasificada, Codigos_paises_2.codigo_2, Codigos_paises_2.pais_2, count(*)
94 FROM Officers INNER JOIN All_edges ON ofi_node_id = con_node_1
95 INNER JOIN Entities ON ent_node_id = con_node_2
96 LEFT JOIN Relaciones ON con_rel_type = conexion
97 LEFT JOIN codigos_paises ON codigo_largo = ofi_country_codes
98 LEFT JOIN codigos_paises_2 ON codigo_largo_2 = ent_country_codes
99 GROUP BY Codigos_paises.codigo, Codigos_paises.pais, conexion_clasificada, Codigos_paises_2.codigo_2, Codigos_paises_2.pais_2)
100 UNION ALL
101 (SELECT Codigos_paises.codigo, Codigos_paises.pais, conexion_clasificada, Codigos_paises_2.codigo_2, Codigos_paises_2.pais_2, count(*)
102 FROM Officers INNER JOIN All_edges ON ofi_node_id = con_node_1
103 INNER JOIN Intermediaries ON int_node_id = con_node_2
104 LEFT JOIN Relaciones ON con_rel_type = conexion
105 LEFT JOIN codigos_paises ON codigo_largo = ofi_country_codes
106 LEFT JOIN codigos_paises_2 ON codigo_largo_2 = int_country_codes
107 GROUP BY Codigos_paises.codigo, Codigos_paises.pais, conexion_clasificada, Codigos_paises_2.codigo_2, Codigos_paises_2.pais_2)
108 UNION ALL
109 (SELECT Codigos_paises.codigo, Codigos_paises.pais, conexion_clasificada, Codigos_paises_2.codigo_2, Codigos_paises_2.pais_2, count(*)
110 FROM Officers INNER JOIN All_edges ON ofi_node_id = con_node_1
111 INNER JOIN Entities ON ent_node_id = con_node_2
112 LEFT JOIN Relaciones ON con_rel_type = conexion
113 LEFT JOIN codigos_paises ON codigo_largo = ofi_country_codes
114 LEFT JOIN codigos_paises_2 ON codigo_largo_2 = ent_country_codes
115 GROUP BY Codigos_paises.codigo, Codigos_paises.pais, conexion_clasificada, Codigos_paises_2.codigo_2, Codigos_paises_2.pais_2)
116 UNION ALL
117 (SELECT Codigos_paises.codigo, Codigos_paises.pais, conexion_clasificada, Codigos_paises_2.codigo_2, Codigos_paises_2.pais_2, count(*)
118 FROM Officers INNER JOIN All_edges ON ofi_node_id = con_node_1
119 INNER JOIN Intermediaries ON int_node_id = con_node_2
120 LEFT JOIN Relaciones ON con_rel_type = conexion
121 LEFT JOIN codigos_paises ON codigo_largo = ofi_country_codes
122 LEFT JOIN codigos_paises_2 ON codigo_largo_2 = int_country_codes
123 GROUP BY Codigos_paises.codigo, Codigos_paises.pais, conexion_clasificada, Codigos_paises_2.codigo_2, Codigos_paises_2.pais_2)
124 UNION ALL
125 (SELECT Codigos_paises.codigo, Codigos_paises.pais, conexion_clasificada, Codigos_paises_2.codigo_2, Codigos_paises_2.pais_2, count(*)
126 FROM Officers INNER JOIN All_edges ON ofi_node_id = con_node_1
127 INNER JOIN Entities ON ent_node_id = con_node_2
128 LEFT JOIN Relaciones ON con_rel_type = conexion
129 LEFT JOIN codigos_paises ON codigo_largo = ofi_country_codes
130 LEFT JOIN codigos_paises_2 ON codigo_largo_2 = ent_country_codes
131 GROUP BY Codigos_paises.codigo, Codigos_paises.pais, conexion_clasificada, Codigos_paises_2.codigo_2, Codigos_paises_2.pais_2)
132 )as todosDatos
133 GROUP BY codigo, pais, conexion_clasificada, codigo_2, pais_2
134 ORDER BY codigo,apariciones DESC;

```

Figura 4-1: view_relaciones_paises

La vista consiste en la unión de sub-consultas que muestran la relación existente entre dos nodos que pueden ser entidades, encargados, intermediarios o direcciones. Estas sub-consultas se han realizado con todas las combinaciones posibles como se indica en el diagrama E-R Offshore Leaks. Las sub-consultas realizan un INNER JOIN con la tabla relación con el objetivo de simplificar el resultado final con solamente 19 relaciones. También es necesario hacer otro INNER JOIN con la tabla codigos_paises por cada nodo de la tabla de conexiones, para poder mostrar de formar individual los códigos que tienen varios países.

Por último, es necesario realizar una consulta a un nivel más alto que las anteriores, encargada de sumar todas las apariciones de tuplas semejantes, es decir, disponen del mismo código del país emisor, tipo de relación y código del país receptor.

Con la vista completada, solo falta realizar la siguiente consulta, para obtener la red de conexiones entre países sin normalizar.

```

78  -- Relaciones entre las diferentes paises
79  COPY (
80  SELECT * FROM view_relaciones_paises
81  WHERE codigo IS NOT NULL AND codigo !='XXX' AND codigo_2 IS NOT NULL AND codigo_2 !='XXX'
82  )TO '/tmp/consulta.csv' WITH DELIMITER ',' csv header;

```

Figura 4-2: consulta red de conexiones entre países sin normalizar

La red de conexiones entre países sin normalizar consiste en una tabla donde aparecen un país emisor, otro receptor, el tipo de relación y las veces que se repite esa combinación en la base de datos de los Papeles de Panamá. Esta tabla la encontramos a continuación:

código	país	conexión_clasificada	codigo_2	pais_2	apariciones
ESP	Spain	shareholder	VGB	British Virgin Islands	1602
ESP	Spain	shareholder	BRA	Brazil	3
HKG	Hong Kong	beneficiary of	COK	Cook Islands	70
HKG	Hong Kong	registered address	SGP	Singapore	99
...

Tabla 4-1 relación entre países sin normalizar

La tabla se lee de la siguiente manera: en la fila dos encontramos a España que posee 3 accionistas brasileños.

Para normalizar los datos, se ha realizado un programa en java que primero lee los datos de la tabla anterior, luego los normaliza e imprime finalmente una tabla con el país emisor, país receptor y el peso de esa conexión tras ser normalizado, consiguiendo la red de conexión entre los países.

Al normalizar los datos se ha decidido realizarlo de la siguiente manera:

Se ha normalizado por el atributo relación, cogiendo el valor en el que ese atributo en la red de conexiones sin normalizar contiene un mayor número de apariciones. Este valor se ha usado posteriormente de referencia para dar valores entre [0-4] dependiendo de lo próximo que se encuentre a este valor la relación de cada tupla de la red de conexiones entre países sin normalizar. Cuanto más cerca se encuentre del valor de 4, significa que esa relación tiene un peso elevado entre los dos países participantes. Se ha decidido que los valores estén comprendidos entre [0,4] porque de esta manera dependiendo del intervalo en el que se encuentre cada relación de par de países, podemos interpretar su grado de aparición en ese tipo de relación.

- [0,1] baja
- (1-3] media
- (3,4] alta

Se ha normalizado de esta manera porque hay una gran diferencia entre las apariciones de cada tipo de relación. De esta manera se consigue que cada relación tenga el mismo peso porque se normaliza respecto a su valor máximo, por ejemplo, si el valor máximo de la relación presidente es 40 y de intermediarios es 4000. Una tupla que tenga un valor de 10 en la relación presidente le corresponderá un 1 al ser normalizado y en el caso de la relación intermediario para conseguir un valor 1 debe de llegar a 1000 apariciones.

Este proceso se formaliza con la siguiente formula:

$$P = \frac{T(po, r, pd)}{Max[f(r)]} * 4$$

$T(po,r,pd)$ = tupla de la tabla relaciones entre países sin normalizar
 po = país origen
 r = relación
 pd = país destino
 $Max[f(r)]$ = valor máximo de la relación en la tabla

Figura 4-3: fórmula de normalización relaciones países

El valor máximo de cada conexión se consigue con la siguiente consulta:

```

84  |-- valor Max de cada relacion-pais
85  | COPY(
86  |   SELECT conexion_clasificada, MAX(apariciones) FROM
87  |   (
88  |     SELECT * FROM view_relaciones_paises
89  |     WHERE codigo IS NOT NULL AND codigo !='XXX' AND codigo_2 IS NOT NULL AND codigo_2 !='XXX'
90  |   ) AS red_paises
91  |   GROUP BY conexion_clasificada
92  |   ORDER BY MAX DESC
93  | ) TO '/tmp/consulta.csv' WITH DELIMITER ',' csv header;
  
```

Figura 4-4: consulta valor máximo de cada relación de la red de conexiones entre países

Se ha extendido la consulta de la red de conexiones entre países sin normalizar para seleccionar únicamente los valores de cada relación cuyo valor son los más grandes.

Con todo el proceso anterior se consigue el vector de relación de cada par de países.

Emisor	Receptor	R1	R2	R3	R4	R5	...	R19	Peso
VGB	HKG	0	0,54	1,06	0,02	0,02	...	0	2,95
ESP	VGB	0,001	0	0,32	0	0,02	...	0	0,36
IDN	SGP	0,01	0,01	0,25	0,03	0,08	...	0	3,56
ESP	GBR	0,005	0	0,01	0	0	...	0	0,06
...

Tabla 4-2 vector relación entre países

Los valores de la tabla anterior que comienzan con R, hacen referencia a cada uno de los tipos de relación de la base de datos de los Papeles de Panamá. En el glosario aparecen su significado.

Finalmente, para calcular el peso de la conexión entre dos países se ha sumado los valores de cada relación normalizada.

Se ha realizado un proceso similar a la red de conexiones entre países, con sus agrupaciones, con el objetivo de analizar las conexiones existentes entre regiones.

4.3 Patrón diccionario

El patrón diccionario consiste en identificar el patrón de las emisiones de los países según su número de apariciones de cada tipo de relación, con el fin de posteriormente buscar similitudes entre ellos. Se obtiene de la siguiente manera:

De la vista `view_relaciones_paises` se realiza una consulta donde únicamente se selecciona el país emisor, el tipo de relación y el número de apariciones de cada una de estas combinaciones. Para conseguir estos datos se ha realizado la siguiente consulta:

```
31 -- Numero de apariciones de cada relacion en cada pais (diccionario)
32 COPY (
33     SELECT codigo, conexion_clasificada, SUM(apariciones) FROM view_relaciones_paises
34     WHERE codigo IS NOT NULL AND codigo != 'XXX'
35     GROUP BY codigo, conexion_clasificada
36     ORDER BY codigo, SUM
37 ) TO '/tmp/consulta.csv' WITH DELIMITER ',' csv header;
```

Figura 4-5: consulta diccionario países sin normalizar

De esta manera se consigue una tabla con todos los códigos de los países, su relación y su número de apariciones.

Código	conexión_clasificada	sum
ARE	Intermediary of	6457
IDN	registered address	4916
ESP	shareholder	2576
WSM	intermediary of	3423
...

Tabla 4-3 apariciones diccionario países

La tabla anterior será tratada por un programa en java que leerá los datos, luego los normalizará y volcará en un fichero una tabla con el patrón de emisiones de cada país, además de las distancias relativas entre países respecto a sus patrones.

Al igual que la red de conexiones entre países, se ha realizado una normalización por el valor máximo de cada tipo de relación. A diferencia de la red de países, el diccionario no tiene un nodo receptor, por lo tanto, la obtención y normalización de datos son diferentes:

```
39 -- valor Max de cada relacion-pais (diccionario)
40 COPY (
41     SELECT conexion_clasificada, MAX(SUM) FROM
42     (
43         SELECT codigo, conexion_clasificada, SUM(apariciones) FROM view_relaciones_paises
44         WHERE codigo IS NOT NULL AND codigo != 'XXX'
45         GROUP BY codigo, conexion_clasificada
46     ) AS diccionario_pais
47     GROUP BY conexion_clasificada
48     ORDER BY MAX DESC
49 ) TO '/tmp/consulta.csv' WITH DELIMITER ',' csv header;
```

Figura 4-6: consulta valor máximo de cada relación del diccionario

Este proceso se formaliza con la siguiente formula:

$$D = \frac{T(po,r)}{Max[f(r)]} * 4$$

T(po,r) = tupla de la tabla apariciones diccionario países.

po = país origen.

r = relación.

Max[f(r)] = valor máximo de la relación en la tabla apariciones diccionario países.

Figura 4-7: fórmula de normalización diccionario países

De este modo se consigue el patrón diccionario de cada país, donde se muestra el resultado en la siguiente tabla:

Emisor	R1	R2	R3	R4	R5	R6	R7	R8	R9	R10	R11	...	R19
VGB	0,03	4	2,7	3,9	1,7	2,7	2,6	1,5	0,3	3,9	1,7	...	0,11
HKG	2	2	3	0	4	1,2	1	2,7	3,1	3,4	1,1	...	3,4
ESP	0	0	0,3	0	0	0	0	0	0,2	2,6	0	...	2,3
SGP	0,5	0,5	1,5	4	1,6	0,2	0,2	2,6	0,9	3,8	0	...	1,1
...

Tabla 4-4 patrón diccionario de los países

El cálculo de las distancias entre los patrones de dos países, se ha realizado de la siguiente manera: sumatorio del valor absoluto de la diferencia de cada tipo de relación. Este proceso se ha hecho entre todos los países.

4.4 Algoritmo Dijkstra

El algoritmo de Dijkstra, encuentra el camino mínimo entre dos nodos de una red. Para realizar este algoritmo se ha decidido hacer un programa en Python. El algoritmo consiste en leer de un fichero que contengan el siguiente formato:

(Todos los nodos que aparecen en la red)

Nodo1

Nodo2

Separador de nodos y conexiones

(Conexiones entre nodos)

Nodo1 Nodo2 distancia

De esta manera el programa leerá los datos y construirá el grafo de la red de conexiones de países al que se le aplicará el algoritmo de Dijkstra.

Para calcular la distancia existente entre dos nodos, se ha realizado de tres maneras diferentes, haciendo uso de la red de conexiones entre países.

La primera forma consiste en asignar como distancia entre países, el inverso al peso de la conexión entre ellos dos, de esta manera un peso mayor implica una menor distancia.

La segunda forma consiste en asignar como distancia 1 a cada peso de la conexión de países que sea mayor que 0.

La tercera forma consiste en aplicar un umbral a los pesos. Las conexiones de los países que superen el umbral se le asignará distancia 1 y en caso contrario 0. El uso del umbral se hace con el objetivo de eliminar las conexiones que se encuentren cercanas a 0.

En cualquiera de las tres formas anteriores, posteriormente se aplicará el algoritmo de Dijkstra y se conseguirá la matriz de distancia entre todos los nodos de la red.

Una vez se ha conseguido la matriz de distancias mínimas entre todos los pares de nodos, se ha realizado la media de distancia tanto por emisiones como recepciones de cada país. Las filas de la matriz representan las emisiones y las columnas se tratan de las recepciones.

4.5 Correlación de los datos

Se ha querido comprobar la correlación entre las apariciones de los países en los papeles de Panamá con atributos como la población, PIB y superficie. Estos datos se han conseguido gracias a la web del Banco mundial, donde se cogieron los datos del año 2015-2016.

El número de apariciones de los países se ha conseguido con la siguiente consulta:

```
210  |-- Consulta para ver el numero de apariciones de los paises en los papeles de panama
211  |COPY (
212  |    SELECT codigo, pais, SUM(count) as apariciones FROM (
213  |        (SELECT codigo, pais, count(*)
214  |        FROM Addresses INNER JOIN All_edges ON con_node_1 = addr_node_id
215  |        LEFT JOIN codigos_paises ON addr_country_codes = codigo_largo
216  |        GROUP BY codigo, pais)
217  |        UNION ALL
218  |        (SELECT codigo, pais, count(*)
222  |        UNION ALL
223  |        (SELECT codigo, pais, count(*)
227  |        UNION ALL
228  |        (SELECT codigo, pais, count(*)
232  |        UNION ALL
233  |        (SELECT codigo, pais, count(*)
237  |        UNION ALL
238  |        (SELECT codigo, pais, count(*)
242  |        UNION ALL
243  |        (SELECT codigo, pais, count(*)
247  |        UNION ALL
248  |        (SELECT codigo, pais, count(*)
252  |    )as apariciones_unidas
253  |    GROUP BY codigo, pais
254  |    ORDER BY apariciones DESC
255  |)TO '/tmp/consulta.csv' WITH DELIMITER ',' csv header;
```

Figura 4-8: consulta apariciones países

La consulta consiste en ir sumando las apariciones de cada país tanto si es el nodo emisor o receptor en cada tipo de entidad de la base de datos.

Para ver la correlación se ha realizado un programa en R que lee los datos anteriores y aplica la correlación de Spearman entre dos vectores.

Además, se ha realizado la correlación entre el patrón diccionario de los países para comprobar cuáles son los países que tienen patrones similares.

5 Integración, pruebas y resultados

5.1 Red de conexiones entre países

La red de países tiene un total de 11829 aristas y 208 nodos lo que lo hace un grafo disperso porque $E \leq V^2$. Un grafo disperso implica que existe un gran número de nodos que no se encuentran conectados. Esto hace indicar que la mayoría de la actividad se encuentra centrada en los paraísos fiscales.

Para ver las relaciones existentes entre los países, se ha decidido filtrar por diferentes umbrales según el peso de las conexiones, de este modo se consigue la lista de adyacencias de los que tienen mayores movimientos en los Papeles de Panamá. Para representar esta lista se ha pintado estas relaciones en un mapa mudo político del mundo. Los países que aparecen coloreados de amarillo significa que tienen también una conexión con ellos mismos.

En el siguiente mapa se observan las conexiones entre los países que superan un umbral de conexión 3. Los territorios que aparecen son los que tienen mayor actividad dentro de los Papeles de Panamá. La mayoría son paraísos fiscales como las Islas Vírgenes Británicas o Hong Kong, aunque también aparecen países con PIB muy altos como Estados Unidos, China, Reino Unido, India y Rusia.

Se puede observar que las Islas Vírgenes Británicas es uno de los grandes núcleos, el cual, dispone de una fuerte conexión con el resto de territorios que se encuentran en este umbral. Los territorios europeos están aislados del resto de países salvo conexiones entre ellos mismos. En estas conexiones europeas aparecen paraísos fiscales como son Guernesey y Jersey; Reino Unido y Rusia que tienen un PIB muy grande y por último Suiza donde se encuentran grandes patrimonios debido al secreto bancario que se ha estado ejerciendo en ese país.

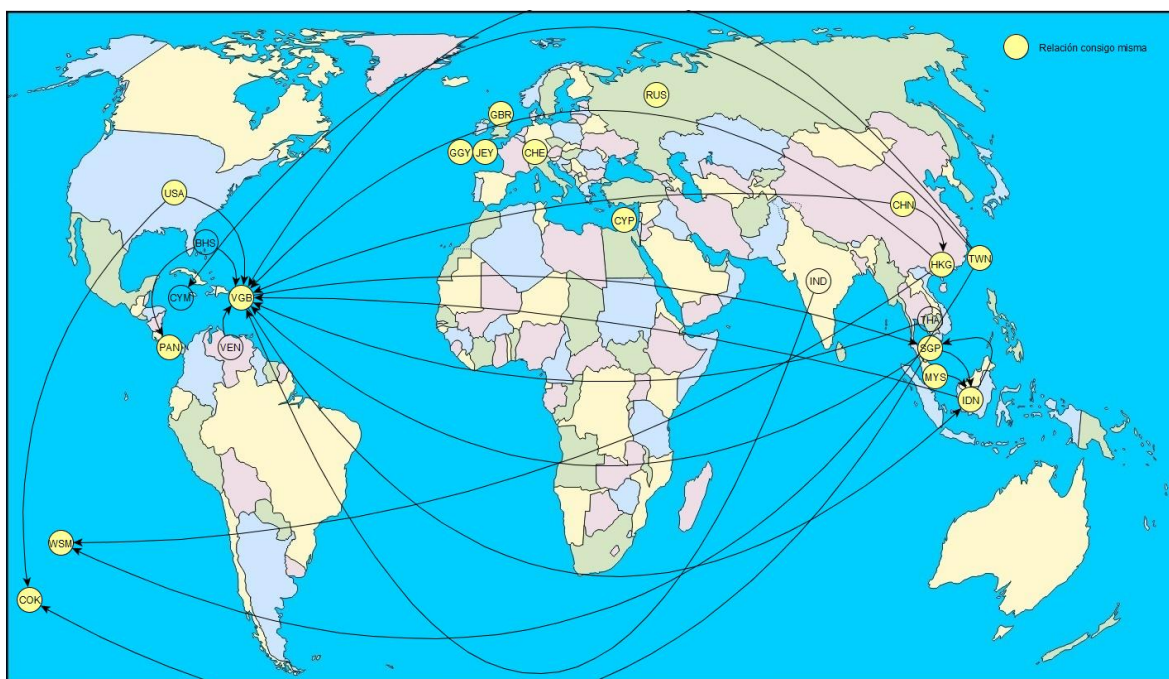


Figura 5-1: conexiones de países umbral 3

El siguiente mapa muestra las conexiones entre los países que se encuentran entre un umbral de [2-3). Estos territorios siguen mayoritariamente representados por paraísos fiscales o países con un gran PIB.

Cabe destacar que la actividad se encuentra concentrada en Indonesia y territorios de su alrededor.

Aparecen nuevos territorios europeos como Italia, Turquía y paraísos fiscales como Gibraltar, Isla de Man. También se encuentran las primeras conexiones con territorios fuera de Europa. Destaca que todos ellos se conectan únicamente con Indonesia.

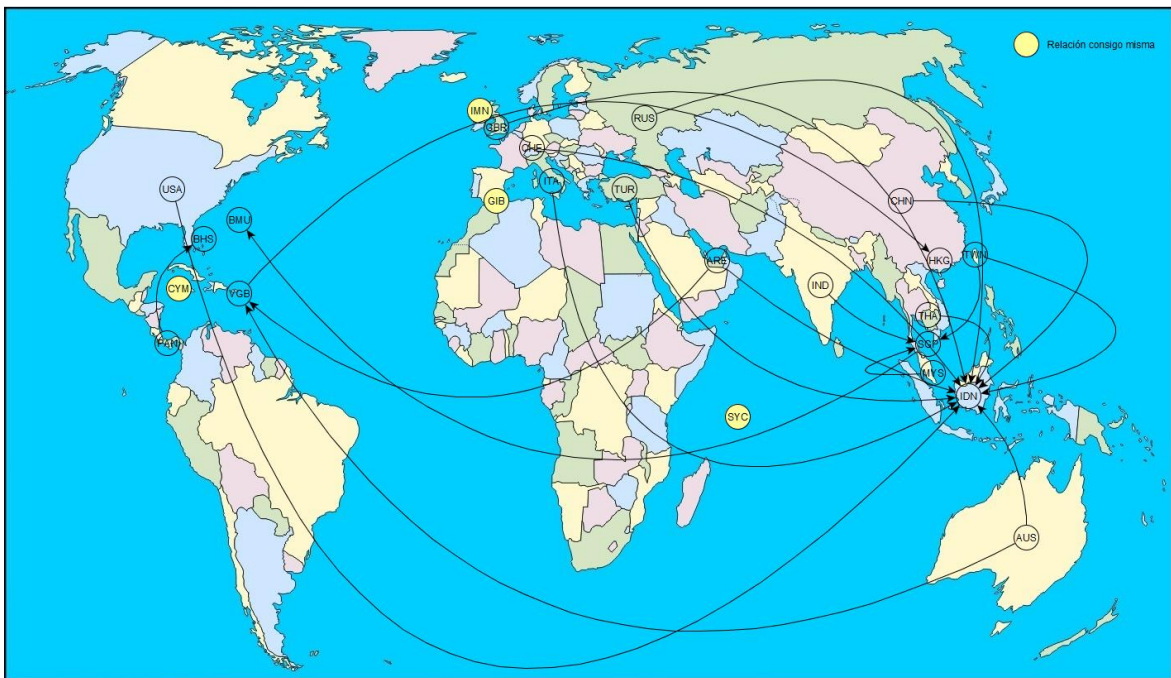


Figura 5-2: conexiones de países umbral [2-3)

El siguiente mapa muestra las conexiones entre los países que se encuentran entre un umbral de [1,5-2).

En este umbral aparece por primera vez España que se encuentra conectada a Indonesia, como ya hicieron el resto de países europeos previamente en el umbral [2-3). También aparecen más territorios europeos como Portugal, Alemania, Grecia y paraísos fiscales como Luxemburgo y Chipre.

Se observa como Indonesia y las Islas Vírgenes Británicas siguen siendo los países de recepción por excelencia.

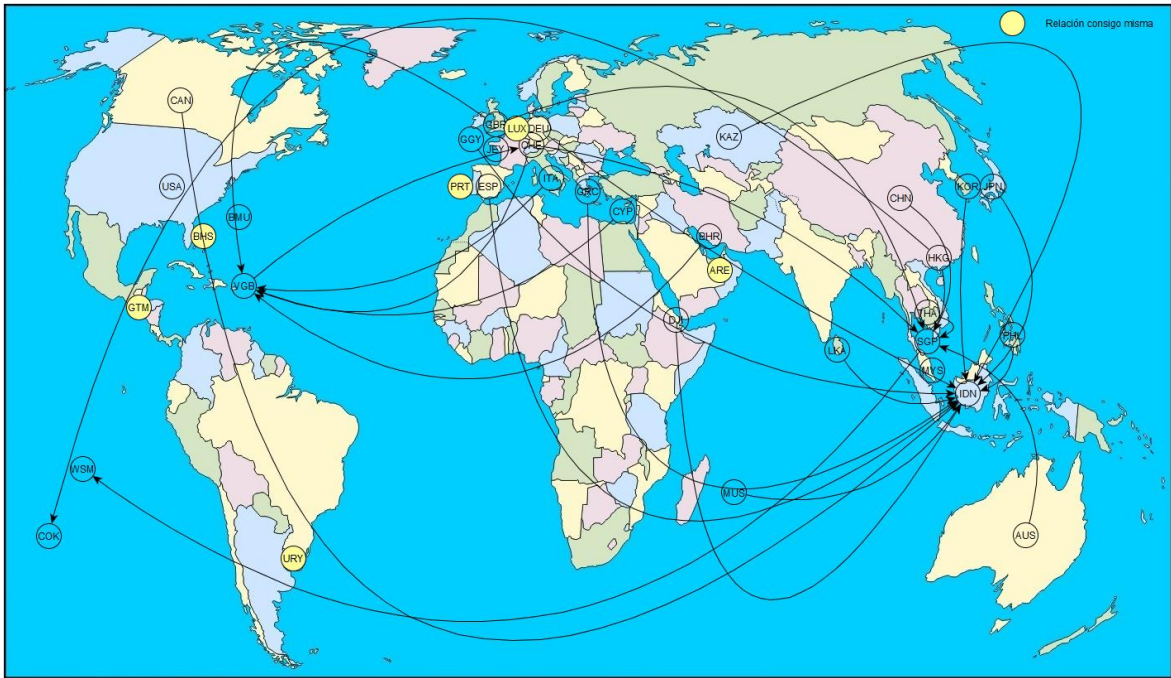


Figura 5-3: conexiones de países umbral [1,5-2)

El siguiente mapa muestra las conexiones entre los países que se encuentran entre un umbral de [1-1,5).

En el mapa se puede observar que el siguiente gran receptor se trata de Singapur, en el cual, los países europeos y España tienen una conexión directa con él.

A este nivel de umbral aparecen los primeros receptores europeos. Estos receptores son Jersey que recibe transacciones desde Reino Unido y Suiza que lo hace desde Panamá.

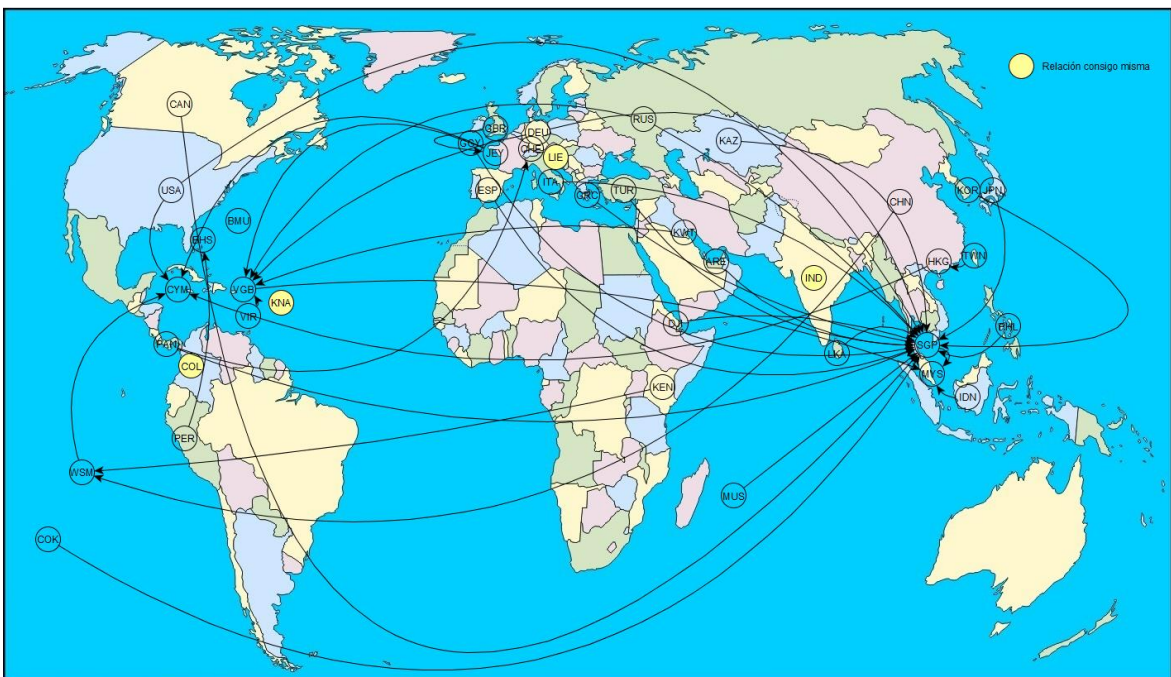


Figura 5-4: conexiones de países umbral [1-1,5)

A continuación, se muestra las conexiones entre territorios. Estas conexiones se han decidido representar de la siguiente forma, como se muestra en la leyenda:

Las flechas de color azul representan las conexiones más fuertes entre territorios, esta intensidad se va reduciendo a color rojo que se trataría de una conexión media fuerte, color verde conexión media y color rosa conexión débil.

Las agrupaciones de países que disponen de conexiones consigo mismas se encuentra en color amarillo si es una conexión fuerte y en naranja si es débil.

Se puede observar a semejanza de la red de conexiones entre países que el umbral más fuerte se encuentra concentrado en América Antillas donde hay paraísos fiscales como las islas vírgenes o Bahamas y Sudeste Asiático donde se encuentra Singapur e Indonesia. Desde el punto de vista de las agrupaciones se puede encontrar la gran actividad que tiene Oceanía con islas como Samoa o las islas Cook que respecto al mapa de los países no se apreciaba con tanta claridad.

Europa utiliza estas tres grandes regiones América Antillas, Sudeste Asiático y Oceanía para el envío en sus transacciones. También se encuentra una conexión entre Europa Germana con Europa Latina. En el comportamiento de Europa se puede observar que, a pesar de utilizar las tres grandes regiones para enviar, únicamente es receptora de América Antillas, esto hace indicar que el Sudeste Asiático y Oceanía hacen de intermediarios para sus transacciones.

El continente africano queda desconectado salvo por la conexión de África Subsahariana que la conecta con el resto del mundo gracias al Sudeste Asiático.

Observando el mapa se puede realizar la siguiente clasificación de las regiones: receptoras, emisoras y desconectados:

Receptoras: América Antillas, Sudeste Asiático, Oceanía y Asia Oriental.

Emisoras: América del Norte, Sur, Central, Europa Latina, Germana, Eslava, Oriente Medio, Asia del Sur, África Subsahariana.

Desconectadas: África Magreb.

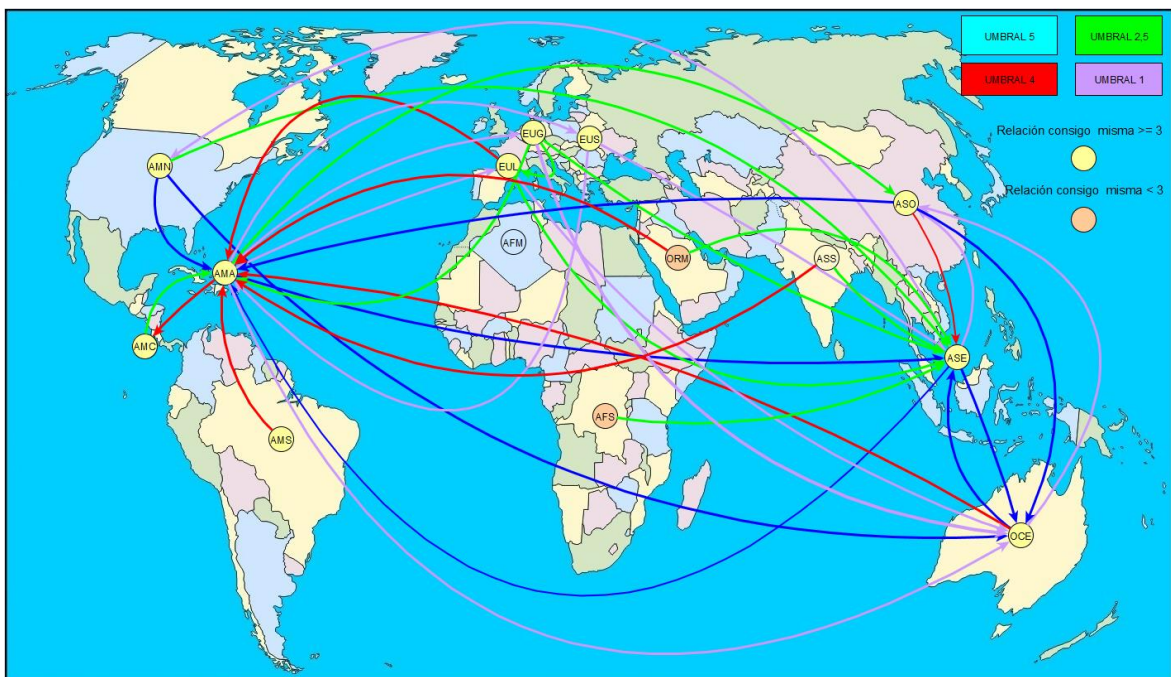


Figura 5-5: conexiones de agrupaciones

5.2 Resultados Dijkstra

En anexos se encuentran imágenes de las tablas que se mencionan a continuación.

Sin aplicar ningún tipo de filtro a la red, se observa que existen países desconectados del resto a la hora de emitir, sin embargo, ninguno se encuentra desconectado al tratarse del receptor. Los paraísos fiscales son los receptores con distancias menores.

Aplicando un filtro que elimine a los países con las distancias más grandes se observa que muchos de los países quedan desconectados, como se mostró anteriormente con la correlación del diccionario de los países.

Países con distancia más cercanas a España:

Emisiones con valores entre 0,5 y 2,5: Samoa, Islas Vírgenes Británicas, Singapur, Rusia, Panamá, Malasia, India, Indonesia, Hong Kong, Reino Unido, Islas Caimán, Islas Cook, Suiza, Bahamas, Chipre, Guernesey.

Las emisiones son principalmente dirigidas a paraísos fiscales.

Recepción con distancias entre 15 y 23: Emiratos árabes unidos, Australia, Bahrein, Bahamas, Canadá, Suiza, Chipre Yibuti, Alemania, Guernesey, Reino Unido, Francia, Grecia, Hong Kong, Indonesia, India, Italia, Jersey, Kazakstán, Kuwait, Japón, Sir Lanka, Luxemburgo, Islas Marshall, Panamá, Perú, Filipinas, Rusia, Seychelles, Tailandia, Turquía, Taiwán, Estado Unidos, Venezuela, Islas Vírgenes Británicas y de Estados Unidos.

Las recepciones se concentran en un mayor número de países distribuidos en diferentes regiones.

Se ha comprobado que los países que tienen distancias menores, suelen estar cercanos geográficamente.

5.3 Correlación de datos

En este apartado se comentarán las correlaciones del patrón diccionario de los países y también la correlación del número de apariciones en los Papeles de Panamá respecto a la superficie, población y PIB.

5.3.1 Correlación patrón diccionario

En el patrón diccionario de los países se puede comprobar que muchos de ellos tienen valores muy bajos en las relaciones. Esto se refleja al realizar las correlaciones entre estos países, dando un total de 84 países con correlaciones muy altas debido a su baja participación en los Papeles de Panamá. La mayoría de estos países se encuentran en el continente africano y Europa Eslava. También llama la atención la baja aparición de

algunos paraísos fiscales como son: Jordania, Liberia, Curasao, Barbados, Nauru, Norte de las islas Marianas, Maldivas, Jamaica y Granada.

Se ha querido comprobar si existe correlación entre los países con más participación en los Papeles de Panamá. Estos países se tratan de las Islas Vírgenes Británicas, Hong Kong, Indonesia y Singapur.

	VGB	HKG	IDN	SGP
VGB	1	-0,22	-0,02	0,24
HKG	-0,22	1	0,5	0,16
IDN	-0,02	0,5	1	0,54
SGP	0,24	0,16	0,54	1

Tabla 5-1 correlación entre países con mayores apariciones

Se puede observar en la tabla anterior, que apenas existe correlación entre estos países. Esto es debido a que cada uno de ellos se especializa en un grupo de relaciones concreto:

VGB: registered address, records & registers of, Nominee Shareholder of, related entity.

HKG: director, trustee of trust of, signatory, officer, president, others.

IDN: records & registers of, secretary, Nominee Shareholder of, president.

SGP: records & registers of.

A continuación, se listan los países con mayor correlación con España:

Rusia (0,74), Panamá (0,7), Italia (0,69), Francia (0,69), Kazajstán (0,68), Jersey (0,67), Australia (0,63), Guatemala (0,63), Mónaco (0,61), Yibuti (0,61), Grecia (0,6).

Se observa que los países con más correlación con España son principalmente otros territorios europeos. También tiene patrones similares a paraísos fiscales como Panamá y Jersey. Con el ejemplo de España se puede observar que países de una misma región es donde se encontrarán una mayor correlación en su patrón diccionario. Como se ha

comentado anteriormente esto también ocurre con los países de Europa Eslava y del continente africano. Estas similitudes también se repiten en el resto de regiones.

5.3.2 Apariciones en los Papeles de Panamá y su correlación

El primer proceso que se ha realizado antes de analizar la correlación entre las apariciones de los países ha sido analizar su histograma.

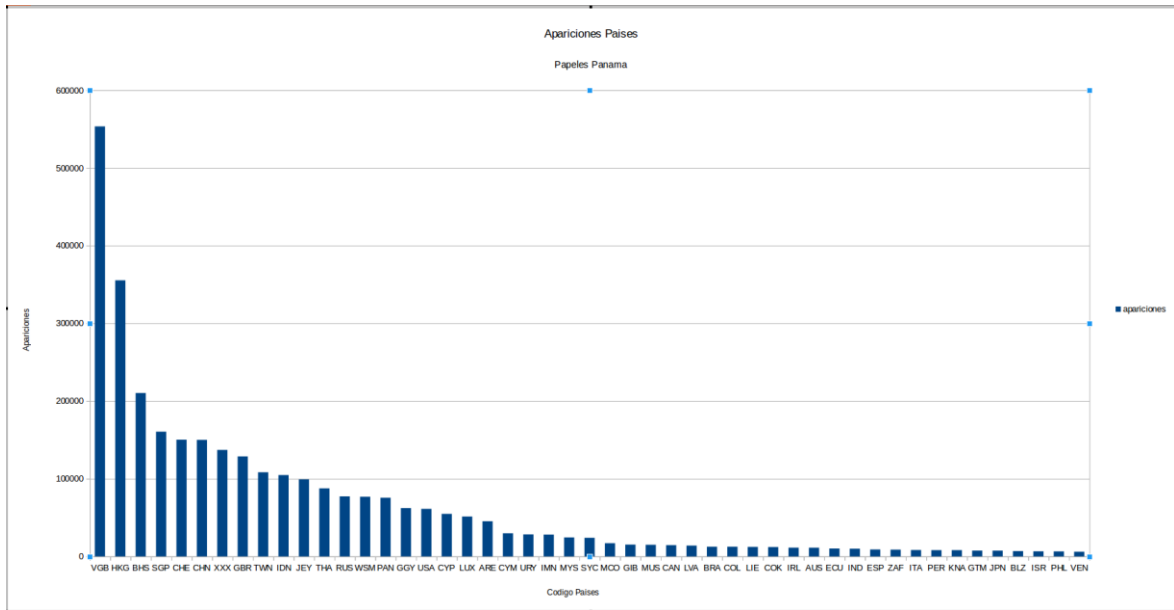


Figura 5-6: histograma apariciones países

Se puede comprobar que la mayor actividad se encuentra centrada en los paraísos fiscales, como, las Islas Vírgenes Británicas, Hong Kong, Bahamas, Singapur, etc. También se observa que hay una gran diferencia entre las apariciones de la primera mitad del histograma y la segunda. Por lo tanto, se ha querido analizar si existen dos comportamientos diferentes.

Para ello se ha aplicado escala logarítmica en el eje “y”, lo que transforma una función exponencial en una lineal. De esta manera es más fácil identificar si existe un ajuste exponencial.

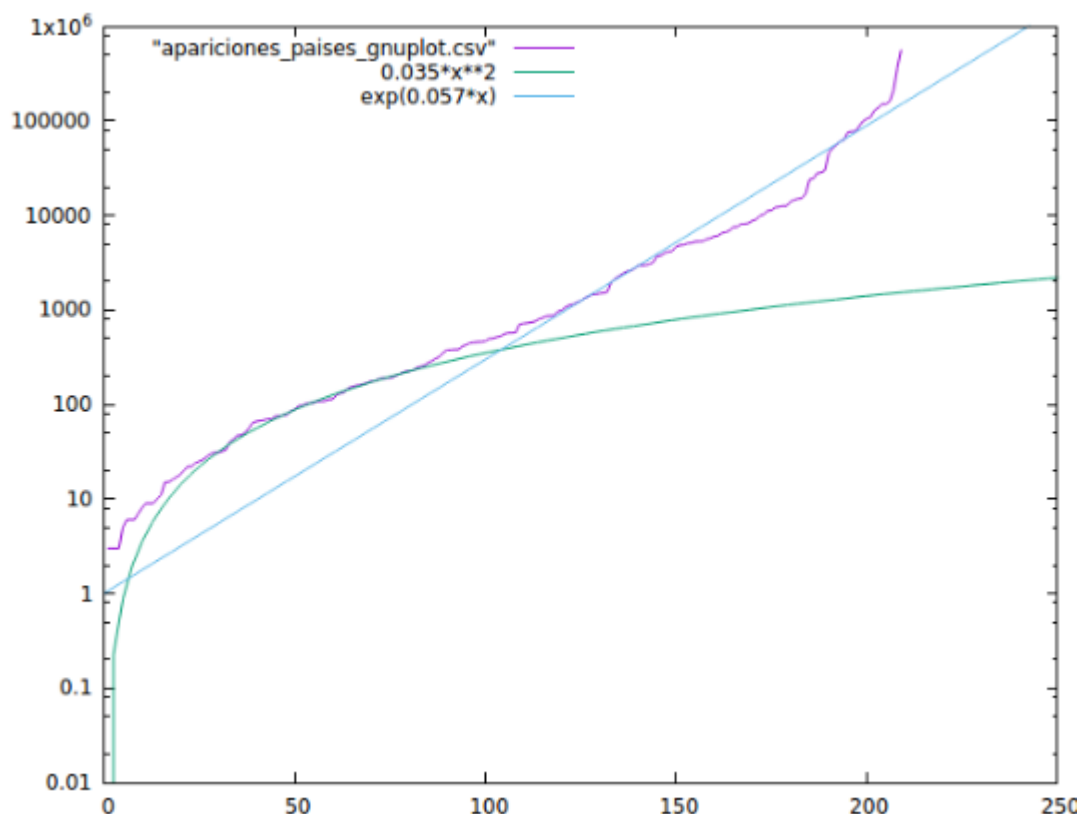


Figura 5-7: ajuste apariciones países

Alrededor de la primera mitad se ajusta a una función potencial, mientras que la segunda mitad es ajustable a una función exponencial. Esto se debe a que la primera parte se encuentran los países que poseen un menor número transacciones en los Papeles de Panamá y además están desconectados de la red de países. La segunda mitad contiene a los grandes paraísos fiscales, países con un alto PIB, estos países tienen un alto grado de transacciones en dichos papeles.

A continuación, se muestran los países clasificados según los continentes por orden de mayor aparición en los Papeles de Panamá:

- Europa: Suiza, Reino Unido, Letonia, Irlanda, España, Italia, Alemania, Portugal, Francia, Grecia.
- América: Uruguay, Brasil, Colombia, Ecuador, Perú, Guatemala, Venezuela
- Asia: Taiwán, Indonesia, Tailandia, Malasia, India, Japón, Israel, Filipinas, Turquía,
- África: Sur África, Kenia
- Oceanía: Australia

Para finalizar, se muestra la correlación de las apariciones con datos como población, superficie y PIB.

	Población	Superficie (km ²)	PIB (millones \$)	Apariciones
Población	1	0,85	0,8	0,12
Superficie (km ²)	0,85	1	0,65	0,02
PIB (millones \$)	0,8	0,65	1	0,4
Apariciones	0,12	0,02	0,4	1

Tabla 5-2 correlación apariciones

Los resultados no demuestran que exista una correlación entre estos datos. Pero se ha comprobado que un número alto de apariciones en los papeles con una población y superficie pequeña generalmente se encuentra clasificado como paraíso fiscal.

5.4 Datos de España

España se trata del decimotercer país con más apariciones dentro de la zona Europa, por detrás de Suiza, Reino Unido, Jersey, Rusia, Guernesey, Chipre, Luxemburgo, Isla de Man, Mónaco, Gibraltar, Letonia, Liechtenstein, Irlanda. De los cuales muchos de ellos son paraísos fiscales. Además, la podemos encontrar en el puesto 38 del mundo con un total de 9010 apariciones en los Papeles de Panamá lo que conlleva a un total del 0,27% del total.

Los intermediarios españoles con más actividad dentro de los papeles de Panamá se tratan de:

- MANAGEMENT CONSULTING ENTERPRISES INC., 258 apariciones
- INTERNATIONAL COMPANY LAWYERS LTD., 178 apariciones
- (GENEVA CLIENT) BUFETE G. PRETUS, 98 apariciones
- ESTUDIO LEGAL ROCA & ASOCIADOS, 95 apariciones
- SHORTT, THOMAS - BARON GALMOY 65 apariciones

Podemos encontrar un total de 1169 entidades españolas.

Estas entidades se encuentran principalmente representadas en las siguientes 6 jurisdicciones

- Panamá 540 entidades
- Niue 186 entidades
- Islas Vírgenes Británicas 185 entidades
- Bahamas 163 entidades
- Seychelles 38 entidades
- Samoa 38 entidades

Solo encontramos 3 proveedores diferentes en las entidades españolas, Mossack Fonseca con 1165 entidades, Commonwealth Trust Limited con 3 entidades y por último con una única entidad Portcullis Trustnet.

Las entidades con más apariciones se tratan de las siguientes:

- GESTER FINANCE CORP. 22 apariciones
- Nautilus Limited 17 apariciones
- KAWAMURA SECURITIES INC. 12 apariciones
- PERFUMES INTERNACIONALES CO., LTD. 12 apariciones
- GLASBOWN FIRST HOLDINGS INC. 12 apariciones

Los encargados con más apariciones se tratan de las siguientes:

- Sharecorp Limited 2695 apariciones
- THE BEARER 360 apariciones
- Pedro Aquiles GRIMALDI Rey 27 apariciones
- PAOLA TORNINI MUÑIZ 26 apariciones
- MR. FADEL MOHAMED IRAKI 22 apariciones

5.5 Recopilación de datos

Para finalizar se ha creado un archivo Excel con diferentes atributos de los países. Estos atributos consisten en información general obtenida del Banco Mundial, apariciones en los Papeles de Panamá y la relación entre estos dos últimos datos, forma de gobierno, el patrón diccionario de cada país, un atributo que indica si es considerado paraíso fiscal según un artículo del periódico ABC en el que se listan los paraísos fiscales según la OCDE al que se le han añadido los paraísos fiscales que aparecen en la lista de la agencia tributaria del gobierno de España. Por último, se ha realizado un programa que clasifica las apariciones en los papeles según sean locales si se realiza dentro de los países de su respectiva región o global en caso contrario y nos indica el porcentaje de cada una de estas dos.

Analizando las divisiones de población, tamaño del territorio y PIB, con las apariciones en los Papeles de Panamá, se puede observar que los paraísos fiscales tienen comportamientos similares. El comportamiento que siguen es que son los que tienen mayores valores en estos atributos. Además, se observa que existen países que siguen esos mismos patrones como Suiza, Taiwán, Yibuti, Samoa Americana.

6 Conclusiones y trabajo futuro

6.1 Conclusiones

El trabajo realizado con los datos de los Papeles de Panamá ha consistido en realizar una red que conecta a los países según los tipos de relaciones que tienen entre ellos. De esta manera se ha podido observar que se pueden realizar tres grandes separaciones según sus actuaciones, estas son: paraísos fiscales, países que interactúan con ellos y los países que se encuentran desconectados de esta red. También se ha comprobado como Indonesia, Suiza que no aparecen en la lista como paraísos fiscales tiene una actuación similar a ellos.

Posteriormente se ha analizado las correlaciones del patrón diccionario de los países y las distancias entre ellos aplicando el algoritmo de Dijkstra, donde se ha comprobado que los países cercanos geográficamente seguían comportamientos similares.

Para finalizar a pesar de que no existe correlación entre las apariciones de los países en los papeles de Panamá con los datos como la superficie, población y PIB, se ha observado que un número alto de apariciones en los papeles, con una población y superficie pequeña generalmente se encuentra clasificado como paraíso fiscal.

6.2 Trabajo futuro

Con la colección de datos de los países que se ha obtenido a lo largo del desarrollo del TFG, se podría plantear identificar qué países son paraísos fiscales y cuales tienen patrones similares a ellos. Para realizar esta operación se aplicarían algoritmos de aprendizaje automático.

Con los datos obtenidos, también es posible plantear con el grupo de investigación en transparencia internacional, liderado por el profesor de la facultad de ciencias económicas, catedrático de contabilidad, Jesús Lizcano la siguiente propuesta. Proponer medidas o normativas resultantes de los resultados estadísticos obtenidos, con la finalidad de publicar y facilitar el control de transferencias económicas, y disminuir la opacidad de divisas, así como la evasión fiscal. Eso implica mejorar la implementación de la base de datos, y extenderlo a países en concreto y sus divisiones internas, como: comunidades, provincias, distritos, ciudades, estados, etc.

Referencias

- [1] <https://offshoreleaks.icij.org/> septiembre 2016
- [2] <http://datos.bancomundial.org/> octubre 2016
- [3] <http://es.gizmodo.com/que-son-los-papeles-de-panama-y-por-que-es-importante-l-1768860761>
- [4] <https://www.uaeh.edu.mx/scige/boletin/tlahuelilpan/n3/e5.html>
- [5] http://www.abc.es/internacional/abci-lista-paraisos-fiscales-segun-ocde-201604090244_noticia.html
- [6] <https://www5.aeat.es/ES13/S/IAFRIAFC11F?REFERENCIA=0112604>
- [7] https://en.wikipedia.org/wiki/List_of_countries_by_system_of_government abril 2017
- [8] Apuntes, diseño y análisis de algoritmos, José Dorronsoro, Escuela Politécnica Superior Universidad Autónoma de Madrid “Design and Analysis of Algorithms”
- [9] <https://es.wikipedia.org/wiki/Grafo>
- [10] https://es.wikipedia.org/wiki/Coeficiente_de_correlaci%C3%B3n_de_Spearman
- [11] <http://www.sciencedirect.com/science/article/pii/S0186104216300407>
- [12] <http://www.sciencedirect.com/science/article/pii/S018519181630023X>

Glosario

Patrón diccionario	El patrón diccionario consiste en las relaciones de cada país cuando se trata del nodo emisor de la red.
R1	Relación 1: intermediary of
R2	Relación 2: registered address
R3	Relación 3: shareholder
R4	Relación 4: records & registers of
R5	Relación 5: director
R6	Relación 6: same address as
R7	Relación 7: same name as
R8	Relación 8: secretary
R9	Relación 9: beneficiary of
R10	Relación 10: Nominee Shareholder of
R11	Relación 11: same name and registration date as
R12	Relación 12: trustee of trust of
R13	Relación 13: others
R14	Relación 14: signatory
R15	Relación 15: settlor
R16	Relación 16: related entity
R17	Relación 17: protector of
R18	Relación 18: officer
R19	Relación 19: president

Anexos

Colección de datos de los países

Muestra de la colección de datos que se ha conseguido a lo largo del TFG.

codigo	pais	codigo_agrupacion	agrupacion	poblacion	superficie (km²)	pib (millones us dolares)	apariciones	forma de gobierno
VGB	British Virgin Islands	AMA	America Antillas	30117	151	839	553443	Constitutional monarchy Parliamentary
HKG	Hong Kong	ASO	Asia Oriental	7305700	2755	309928	355388	Republic Single party
BHS	Bahamas	AMA	America Antillas	389019	13943	8884	210323	Constitutional monarchy Parliamentary
SGP	Singapore	ASE	Asia Sudeste	5535002	719	292739	160596	Republic Parliamentary
CHE	Switzerland	EUL	Europa Latina	8286976	41284	664737	150271	Republic Parliamentary
CHN	China	ASO	Asia Oriental	1371220000	9596961	10866443	149995	Republic Single party
GBR	United Kingdom	EUG	Europa Germana	65138232	331652	2848755	128711	Constitutional monarchy Parliamentary
TWN	Taiwan	ASO	Asia Oriental	23492000	36193	474000	108326	Republic Presidential
IDN	Indonesia	ASE	Asia Sudeste	257563815	1910931	861933	104737	Republic Presidential
JEY	Jersey	EUL	Europa Latina	100080	116	5100	99218	Constitutional monarchy Parliamentary
THA	Thailand	ASE	Asia Sudeste	67959359	513120	395281	87597	others
RUS	Russia	EUS	Europa Eslava	144096812	17098242	1326015	77277	Republic Presidential
WSM	Samoa	OCE	Oceania	193228	2842	761	76848	Republic Parliamentary
PAN	Panama	AMC	America Central	3929141	75417	52132	75543	Republic Presidential
GGY	Guernsey	EUL	Europa Latina	62948	78	2742	62191	Constitutional monarchy Parliamentary
USA	United States	AMN	America del Norte	321418820	9371174	17946996	61210	Republic Presidential
CYP	Cyprus	EUS	Europa Eslava	1165300	9251	19319	54796	Republic Presidential
LUX	Luxembourg	EUG	Europa Germana	569676	2586	57793	51357	Constitutional monarchy Parliamentary
ARE	United Arab Emirates	ORM	Oriente Medio	9156963	83600	370292	45298	Absolute monarchy
CYM	Cayman Islands	AMA	America Antillas	59967	264	1012	29619	Constitutional monarchy Parliamentary
URY	Uruguay	AMS	America del Sur	3431555	176215	53442	28338	Republic Presidential
IMN	Isle of Man	EUG	Europa Germana	87780	572	4076	28142	Constitutional monarchy Parliamentary
MYS	Malaysia	ASE	Asia Sudeste	30331007	330803	296217	24461	Constitutional monarchy Parliamentary
SYC	Seychelles	AFS	Africa Subsahariana	92900	452	1437	23979	Republic Presidential

apariciones/poblacion	apariciones/area	apariciones/pib (millones us dolares)	pib (millones us dolares)/poblacion	global	local	Paraiso fiscal
18,37643	3665,18543	659,64601	0,02786	0,85	0,15	TRUE
0,04865	128,99746	1,14668	0,04242	0,93	0,08	TRUE
0,54204	15,08449	23,67436	0,0229	0,91	0,09	TRUE
0,02901	223,36022	0,5486	0,05289	0,87	0,13	TRUE
0,01813	3,63993	0,22606	0,08022	0,85	0,15	FALSE
0,00011	0,01563	0,0138	0,00793	0,89	0,11	FALSE
0,00198	0,38809	0,04518	0,04373	0,89	0,11	FALSE
0,00461	2,99301	0,22854	0,02018	0,91	0,09	FALSE
0,00041	0,05481	0,12151	0,00335	0,86	0,14	FALSE
0,99139	855,32759	19,45451	0,05096	0,87	0,13	TRUE
0,00129	0,17071	0,22161	0,00582	0,88	0,12	FALSE
0,00054	0,00452	0,05828	0,0092	0,88	0,12	FALSE
0,39771	27,04011	100,98292	0,00394	0,96	0,04	TRUE
0,01923	1,00167	1,44907	0,01327	0,94	0,06	TRUE
0,98797	797,32051	22,68089	0,04356	0,85	0,15	TRUE
0,00019	0,00653	0,00341	0,05584	0,96	0,04	FALSE
0,04702	5,92325	2,83638	0,01658	0,77	0,23	TRUE
0,09015	19,85963	0,88864	0,10145	0,85	0,15	TRUE
0,00495	0,54184	0,12233	0,04044	0,9	0,1	TRUE
0,49726	112,95076	29,46542	0,01688	0,91	0,09	TRUE
0,00826	0,16081	0,53026	0,01557	0,72	0,28	FALSE
0,3206	49,1993	6,90432	0,04643	0,9	0,1	TRUE
0,00081	0,07394	0,08258	0,00977	0,84	0,16	FALSE
0,25812	53,05088	16,68685	0,01547	0,97	0,03	TRUE
0,45164	8520,5	2,8051	0,16101	0,8	0,2	TRUE
0,47264	2537,83333	14,28424	0,03309	0,82	0,18	TRUE
0,01189	7,35637	1,30382	0,00912	0,87	0,13	TRUE
0,00041	0,00146	0,00939	0,04325	0,97	0,03	FALSE
0,00706	0,2165	0,517	0,01367	0,68	0,32	FALSE
0,00006	0,00147	0,00705	0,00854	0,87	0,13	FALSE
0,00026	0,01096	0,04285	0,00606	0,79	0,21	FALSE
0,32674	76,64375	2,54103	0,12859	0,85	0,15	TRUE
1,1133	51,41949	66,31148	0,01679	0,83	0,17	TRUE
0,00243	0,16055	0,0474	0,05129	0,86	0,14	FALSE
0,00047	0,00145	0,0084	0,05633	0,91	0,09	FALSE

Muestra del patrón diccionario

intermediary of registered address	shareholder	records & registers of director	same address as	same name as secretary	beneficiary of Nominee Shareholder	same name and registration date as trustee of trust of others	signatory settlor	related entity	protector of officer	president
0.03	4	2.72	3.99	1.76	2.75	2.66	1.5	0.35	3.8	1.14
2.04	2	2.92	0	4	1.17	1.08	2.65	3.08	3.37	4
4	0.11	0.38	0	0.16	0.51	0.04	0.14	0	0	0
0.45	0.46	1.5	4	1.64	0.23	0.2	2.63	0.86	3.84	0.04
0.79	0.32	0.7	0	0.11	0.39	0.29	1.32	0.77	2.58	0.79
0.12	1.64	4	0	2.07	1.58	1.55	0.67	4	2.84	0.02
1.14	0.69	1.38	0	0.4	0.33	0.33	0.17	1.13	2.83	0.12
0.33	1.01	2.87	0	2.86	0.25	0.24	3.48	0.73	3.41	0
0.05	0.23	0.96	3.99	1.05	0.09	0.09	4	0.73	4	0
0.34	0.35	1.52	0	0.07	4	4	0.07	0.23	0.01	0.05
0.04	0.1	0.46	3.99	0.14	0.04	0.04	0.03	0.18	3.43	0.02
0.31	0.66	0.91	0	0.52	0.22	0.06	1.59	1.59	2.67	0.18
0.08	0.19	1.67	0	0.36	0.08	0.07	3.26	0.03	0.13	0.07
0.32	0.37	1.06	0	0.52	1.02	1.01	0.01	0.32	0.04	3
0.16	0.18	0.92	0	0.05	3.47	3.48	0.11	0.14	2.59	0.01
0.26	0.57	0.76	0	0.32	0.15	0.14	0.2	3.12	2.75	0.23
0.4	0.31	0.52	0	0.62	0.82	0.8	0.81	0.37	0	0.07
0.35	0.59	0.2	0	0	0.09	0.09	0	0.06	0	0.1
0.15	0.21	0.64	0	0.63	0.21	0.21	0.19	0.46	2.67	0.04
0.01	0.19	0.23	0	0.15	0.07	0.06	0.83	0.06	0.03	0.34
0.12	0.1	0.19	0	0	0.11	0.11	0	0.42	0	0.05
0.12	0.1	0.38	0	0.01	0.56	0.56	0.04	0.17	0	0.64
0.03	0.18	0.54	0	0.82	0.11	0.11	0.82	0.3	3.41	0
0.03	0.12	0.64	0	0.31	0.98	0.97	0.09	0.05	0	0.23
0.07	0.07	0.17	0	0.62	0.07	0.07	0.06	0.24	0.13	0
0.04	0.05	0.18	0	0.02	0.71	0.71	0	0.15	0	0
0.03	0.09	0.35	0	0.02	0.16	0.16	0.02	0.04	2.6	0
0.08	0.12	0.3	0	0.1	0.1	0.1	0.05	0.28	2.6	0.01
0.11	0.08	0.02	0	0.02	0	0	0	0.05	0	0.03
0.03	0.1	0.2	0	0.01	0.08	0.08	0	0.47	0	0.01
0.04	0.06	0.1	0	0	0.15	0.15	0	0.48	0	0.07
0.05	0.05	0.16	0	0.02	0.29	0.29	0	0.17	0	0
0	0.04	0.16	0	0.29	0.01	0.01	1.54	0.09	0.19	0

Tabla resultados Dijkstra sin filtrar

Muestra de los resultados del algoritmo Dijkstra sin aplicar ningún tipo de filtro en las distancias.

	ZWE	ZMB	ZAF	YEM	WSM	VUT	VNM	VIR	VGB	VEN	VCT	UZB	USA	URY	UKR	UGA	TZA	TWN	TUR
1	emisiones/recepciones																		
2	ZWE		93.233inf	16.419	248.22	69.356	116.971	16.285	37.814	1009.105inf	18.891	30.727	50.9inf	18.891	30.727	485.423	18.664	32.857	
3	ZMB	348.05	248.798inf		172.054	403.785	224.921	272.536	171.85	194.368	1164.764inf	174.456	157.281	206.465inf	640.968	174.229	158.434		
4	ZAF	179.452inf		0inf	2.09	235.187	56.153	103.938	3.252	25.77	996.51inf	5.688	18.683	37.867inf	472.39	5.631	19.836		
5	YEM	1472.465inf	1373.283	0	1296.469	1528.27	1349.406	1397.021	1296.335	1318.853	2289.155inf	1298.941	1311.766	1330.95inf	1765.473	1298.714	1312.907		
6	WSM	177.362inf	78.11inf	0	233.097	54.063	101.848	1.162	23.68	994.42inf	3.598	16.593	35.777inf	470.3	3.541	17.746			
7	VUT	185.742inf	86.32inf	9.844	0	62.443	110.058	9.372	31.89	1002.63inf	11.978	24.803	43.987inf	478.51	11.751	25.956			
8	VNM	206.32inf	106.898inf	30.422	261.885	0	130.636	29.95	52.468	1023.208inf	32.556	45.381	64.565inf	499.088	32.329	46.534			
9	VIR	177.7inf	77.678inf	1.202	232.665	53.801	0	0.73	23.248	993.988inf	3.336	16.161	35.345inf	469.868	3.109	17.314			
10	VGB	176.949inf	76.849inf	0.472	231.935	53.071	100.686	0	22.518	993.250inf	2.606	15.431	34.515inf	469.138	2.379	16.584			
11	VEN	176.611inf	77.189inf	0.713	232.176	53.312	100.927	0.241	0	993.499inf	2.847	15.672	34.856inf	469.379	2.62	16.825			
12	VCT	418.069inf	318.817inf	240.707	473.804	294.77	342.555	241.869	264.387	0inf	244.305	257.3	276.484inf	711.007	244.248	258.453			
13	UZB	241.279inf	141.857inf	65.381	296.844	117.98	165.595	64.909	87.427	1058.167	0	67.515	80.34	99.524inf	534.047	67.288	81.493		
14	USA	176.666inf	77.244inf	0.681	232.231	53.367	100.982	0.296	22.814	993.385inf	0	15.727	34.911inf	469.434	2.675	16.88			
15	URY	193.939inf	94.517inf	18.041	249.504	70.64	118.255	17.569	37.76	1010.827inf	19.951	0	52.184inf	486.707	19.948	34.153			
16	UKR	188.095inf	88.913inf	12.099	243.9	65.036	112.651	11.965	34.483	1004.816inf	14.571	27.396	0inf	481.103	14.344	28.537			
17	UGA	1459.131inf	1359.949inf	1283.135	1514.936	1335.957	1303.697	1293.001	1305.519	2275.923inf	1205.492	1298.432	1317.616	0	1752.139	1205.38	1299.573		
18	TZA	348.079inf	248.897inf	172.083	403.884	224.905	272.535	171.849	194.467	1164.771inf	174.44	157.38	206.564inf	0	174.328	188.521			
19	TWN	176.523inf	77.101inf	0.121	232.088	53.224	100.839	0.153	22.671	993.311inf	2.759	15.584	34.768inf	469.291	0	16.737			
20	TUR	176.67inf	77.488inf	0.674	232.475	53.611	101.226	0.54	23.058	993.369inf	3.146	15.971	35.155inf	469.678	2.919	0			
21	TTO	337.34inf	238.162inf	161.348	393.149	214.17	261.9	161.214	183.732	1154.036inf	163.705	176.645	195.829inf	630.352	163.593	177.786			
22	TON	691.981inf	592.559inf	516.083	747.546	568.682	616.297	515.611	538.129	1508.869inf	518.217	531.042	550.226inf	984.749	517.99	532.195			
23	TOM	202.274inf	102.852inf	26.376	257.839	78.975	126.59	25.904	48.422	1019.162inf	28.51	41.335	60.519inf	495.042	28.283	42.488			
24	TKM	372.32inf	273.139inf	196.325	428.126	249.262	296.877	196.191	218.709	1189.042inf	198.797	211.622	230.806inf	665.329	196.57	212.763			
25	THA	893.386inf	594.134inf	517.39	749.121	570.257	617.872	517.386	539.704	1510.11inf	519.792	532.617	551.301inf	996.324	519.665	533.77			
26	THA	176.535inf	77.146inf	0.539	232.133	53.259	100.894	0.188	22.716	993.312inf	2.804	15.629	34.813inf	469.336	2.577	16.792			
27	TGO	691.744inf	592.322inf	515.846	747.309	566.445	616.06	515.374	537.892	1508.632inf	517.98	530.805	549.989inf	984.512	517.753	531.958			
28	TCD	408.549inf	310.17inf	233.347	465.157	286.293	333.908	233.222	255.74	1226.48inf	235.828	248.653	267.837inf	702.36	235.601	249.806			
29	TCA	197.979inf	98.556inf	22.08	253.543	74.679	122.294	21.608	44.126	1014.866inf	24.214	37.039	56.223inf	490.746	23.987	38.192			
30	SYC	185.019inf	85.837inf	9.023	240.824	61.845	109.575	8.889	31.407	1001.711inf	11.38	24.32	43.504inf	478.027	11.268	25.461			
31	SYC	178.636inf	79.384inf	2.64	234.371	55.507	103.122	2.436	24.964	995.35inf	5.042	17.867	37.051inf	471.574	4.815	19.02			
32	SMM	691.414inf	592.162inf	515.418	747.149	568.295	615.9	515.214	537.732	1508.128inf	517.82	530.645	549.829inf	984.352	517.593	531.798			
33	SWZ	390.435inf	291.253inf	214.439	446.24	267.376	314.991	214.305	236.823	1207.125inf	215.911	229.736	249.92inf	683.443	216.684	230.877			
34	SWE	204.42inf	105.253inf	28.744	260.24	81.121	128.991	28.305	50.823	1021.538inf	29.686	43.736	62.92inf	497.443	30.669	44.876			
35	SVN	308inf	208.578inf	132.102	363.565	184.701	232.316	131.63	154.148	1124.888inf	134.236	147.061	166.245inf	600.768	134.009	148.214			
36	SVK	239.496inf	140.074inf	63.598	295.061	116.197	163.812	63.126	85.644	1056.384inf	65.732	78.557	97.741inf	532.264	65.505	79.71			

Tabla resultados Dijkstra sin filtrar simplificada

Muestra de los resultados del algoritmo Dijkstra en el que los pesos mayores que 0 se le ha asignado valor 1.

emisiones/r/ZWE	ZMB	ZAF	YEM	WSM	VUT	VNM	VIR	VGB	VEN	VCT	UZB	USA	URY	UKR
ZWE	0	208	1	208	2	2	2	2	1	2	2	208	2	2
ZMB	2	0	2	208	2	2	2	2	2	2	2	208	2	2
ZAF	2	208	0	208	1	2	2	2	1	2	2	208	1	1
YEM	3	208	2	0	2	2	2	2	2	2	3	208	2	2
WSM	2	208	1	208	0	2	2	2	1	1	2	208	1	1
VUT	2	208	2	208	1	0	2	2	1	2	2	208	1	2
VNM	2	208	2	208	1	2	0	2	1	2	2	208	1	2
VIR	2	208	1	208	1	1	2	0	1	2	2	208	1	2
VGB	2	208	1	208	1	1	1	1	0	1	2	208	1	1
VEN	2	208	2	208	2	2	2	2	1	0	2	208	1	2
VCT	2	208	2	208	1	2	2	2	1	2	0	208	1	2
UZB	2	208	2	208	2	2	2	2	1	2	2	0	2	2
USA	2	208	1	208	1	2	1	1	1	1	2	208	0	1
URY	2	208	2	208	2	2	2	2	1	1	2	208	1	0
UKR	2	208	2	208	1	2	2	2	1	2	2	208	1	2
UGA	2	208	2	208	2	3	2	2	2	2	3	208	2	2
TZA	2	208	2	208	1	2	2	2	1	2	2	208	2	2
TWN	2	208	1	208	1	1	1	1	1	2	2	208	1	2
TUR	2	208	2	208	1	2	1	1	1	2	2	208	1	2
TUN	2	208	2	208	2	2	2	2	2	2	2	208	2	2

Tabla resultados Dijkstra con filtro

Muestra de los resultados del algoritmo Dijkstra aplicando un filtro para eliminar las conexiones de los países con distancias mayores. Los países que se encontraban dentro del filtro se les ha puesto como distancia 1 y en caso contrario 208, dando lugar a la lista de adyacencias de cada país.

Emisiones/Recepciones	ZWE	ZMB	ZAF	YEM	WSM	VUT	VNM	VIR	VGB	VEN	VCT	UZB	USA	URY	UKR	UGA	TZA	TWN	TUR
ZWE	0	208	1	208	2	208	2	3	2	3	208	208	2	3	3	208	208	2	2
ZMB	3	0	3	208	2	208	2	3	2	3	208	208	2	3	2	208	208	2	2
ZAF	2	208	0	208	1	208	2	2	1	3	208	208	1	2	2	208	208	2	2
YEM	208	208	208	0	208	208	208	208	208	208	208	208	208	208	208	208	208	208	208
WSM	3	208	2	208	0	208	2	2	1	2	208	208	2	2	2	208	208	1	2
VUT	3	208	2	208	1	0	2	2	1	3	208	208	2	2	2	208	208	2	2
VNM	3	208	2	208	2	208	0	2	1	3	208	208	2	2	2	208	208	2	2
VIR	3	208	2	208	1	208	2	0	1	2	208	208	1	2	2	208	208	2	2
VGB	2	208	1	208	1	208	1	1	0	2	208	208	1	1	1	208	208	1	1
VEN	3	208	2	208	2	208	2	2	1	0	208	208	1	1	2	208	208	2	2
VCT	208	208	208	208	208	208	208	208	208	208	0	208	208	208	208	208	208	208	208
UZB	3	208	2	208	2	208	2	2	1	3	208	0	2	2	2	208	208	2	2
USA	2	208	2	208	1	208	1	2	1	2	208	208	0	1	2	208	208	1	1
URY	3	208	2	208	2	208	2	2	1	1	208	208	2	0	2	208	208	2	2
UKR	3	208	2	208	1	208	2	2	1	3	208	208	2	2	0	208	208	2	2
UGA	208	208	208	208	208	208	208	208	208	208	208	208	208	208	208	0	208	208	208
TZA	3	208	3	208	2	208	2	3	2	3	208	208	2	3	3	208	0	2	2
TWN	2	208	2	208	1	208	1	2	1	3	208	208	1	2	2	208	208	0	1
TUR	2	208	2	208	2	208	1	2	1	2	208	208	1	2	2	208	208	1	0
TUN	3	208	3	208	2	208	2	3	2	3	208	208	2	3	3	208	208	2	2
TTO	208	208	208	208	208	208	208	208	208	208	208	208	208	208	208	208	208	208	208
TON	4	208	3	208	2	208	3	3	2	4	208	208	2	3	3	208	208	2	3
TKM	3	208	3	208	3	208	2	3	2	3	208	208	2	3	2	208	208	2	2
TJK	208	208	208	208	208	208	208	208	208	208	208	208	208	208	208	208	208	208	208
THA	2	208	2	208	1	208	1	2	1	2	208	208	1	2	2	208	208	1	1
TGO	208	208	208	208	208	208	208	208	208	208	208	208	208	208	208	208	208	208	208
TCO	208	208	208	208	208	208	208	208	208	208	208	208	208	208	208	208	208	208	208
TCA	3	208	2	208	2	208	2	2	1	3	208	208	2	2	2	208	208	2	2
SYR	3	208	3	208	2	208	2	3	2	3	208	208	2	3	2	208	208	2	2
SYC	3	208	2	208	1	208	2	2	1	2	208	208	2	1	1	208	208	2	2
SXM	208	208	208	208	208	208	208	208	208	208	208	208	208	208	208	208	208	208	208
SWZ	208	208	208	208	208	208	208	208	208	208	208	208	208	208	208	208	208	208	208

Correlación del patrón diccionario

Muestra de los resultados de la correlación del patrón diccionario de los países.

	ZWE	ZMB	ZAF	YEM	WSM	VUT	VNM	VIR	VGB	VEN	VCT	UZB	USA	URY	UKR	UGA	TZA	TWN	TUR
ZWE	1	NA	0,73	NA	0,01	0,45	0,84	0,23	0,27	0,53	0,44	0,75	-0,11	0,83	0,76	NA	0,44	0,29	0,69
ZMB	NA	1	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
ZAF	0,73	NA	1	NA	0,3	0,69	0,73	0,21	0	0,23	0,39	0,59	0,05	0,55	0,63	NA	0,39	0,61	0,6
YEM	NA	NA	NA	1	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
WSM	0,01	NA	0,3	NA	1	0,47	0,2	0,47	0,08	0,08	0,35	0,1	-0,01	-0,03	0,1	NA	0,35	0,63	0,29
VUT	0,45	NA	0,69	NA	0,47	1	0,68	0,33	0,29	0,26	0,37	0,29	-0,17	0,29	0,52	NA	0,37	0,48	0,46
VNM	0,84	NA	0,73	NA	0,2	0,68	1	0,44	0,27	0,59	0,42	0,68	-0,11	0,66	0,87	NA	0,42	0,43	0,77
VIR	0,23	NA	0,21	NA	0,47	0,33	0,44	1	0,16	0,15	0,31	0,25	0,26	0	0,33	NA	0,31	0,46	0,56
VGB	0,27	NA	0	NA	0,08	0,29	0,27	0,16	1	0	0,17	0,17	-0,53	0,06	0,02	NA	0,17	0,02	0,4
VEN	0,53	NA	0,23	NA	0,08	0,26	0,59	0,15	0	1	0,31	0,53	0,08	0,57	0,56	NA	0,31	0,21	0,46
VCT	0,44	NA	0,39	NA	0,35	0,37	0,42	0,31	0,17	0,31	1	0,51	0,04	0,4	0,37	NA	1	0,3	0,34
UZB	0,75	NA	0,59	NA	0,1	0,29	0,68	0,25	0,17	0,53	0,51	1	0,17	0,65	0,66	NA	0,51	0,36	0,62
USA	-0,11	NA	0,05	NA	-0,01	-0,17	-0,11	0,26	-0,53	0,08	0,04	0,17	1	-0,17	0,02	NA	0,04	-0,02	-0,01
URY	0,83	NA	0,55	NA	-0,03	0,29	0,66	0	0,06	0,57	0,4	0,65	-0,17	1	0,78	NA	0,4	0,21	0,5
UKR	0,76	NA	0,63	NA	0,1	0,52	0,87	0,33	0,02	0,56	0,37	0,66	0,02	0,78	1	NA	0,37	0,29	0,61
UGA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	1	NA	NA	NA
TZA	0,44	NA	0,39	NA	0,35	0,37	0,42	0,31	0,17	0,31	1	0,51	0,04	0,4	0,37	NA	1	0,3	0,34
TWN	0,29	NA	0,61	NA	0,63	0,48	0,43	0,46	0,02	0,21	0,3	0,36	-0,02	0,21	0,29	NA	0,3	1	0,7
TUR	0,69	NA	0,6	NA	0,29	0,46	0,77	0,56	0,4	0,46	0,34	0,62	-0,01	0,5	0,61	NA	0,34	0,7	1
TUN	0,5	NA	0,35	NA	-0,26	-0,16	0,42	-0,04	-0,22	0,36	-0,06	0,61	0,22	0,45	0,41	NA	-0,06	0,13	0,38
TTO	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
TON	0,5	NA	0,35	NA	-0,26	-0,16	0,42	-0,04	-0,22	0,36	-0,06	0,61	0,22	0,45	0,41	NA	-0,06	0,13	0,38
TKM	-0,14	NA	-0,29	NA	-0,37	-0,16	-0,16	-0,27	-0,13	0,4	-0,06	-0,1	0,32	-0,17	-0,23	NA	-0,06	-0,3	-0,19
TJK	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
THA	0,19	NA	0,24	NA	-0,05	0,16	0,29	0,51	0,17	-0,16	0,26	0,35	0,27	0,04	0,22	NA	0,26	0,22	0,41
TGO	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
TCD	0,44	NA	0,39	NA	0,35	0,37	0,42	0,31	0,17	0,31	1	0,51	0,04	0,4	0,37	NA	1	0,3	0,34
TCA	0,42	NA	0,34	NA	0,46	0,61	0,61	0,59	0,24	0,39	0,4	0,24	-0,18	0,24	0,53	NA	0,4	0,52	0,53
SYR	0,45	NA	0,64	NA	0,2	0,4	0,38	-0,03	-0,14	0,22	0,51	0,58	0,21	0,37	0,3	NA	0,51	0,4	0,29
SYC	0,67	NA	0,47	NA	0,34	0,56	0,75	0,34	0,22	0,67	0,32	0,35	-0,33	0,71	0,73	NA	0,32	0,36	0,57
SXM	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
SWZ	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
SWE	0,43	NA	0,53	NA	0,14	0,47	0,56	0,14	-0,06	0,44	0,4	0,68	0,25	0,41	0,65	NA	0,4	0,33	0,39