

Documentos de Trabajo

ISSN 2219-780X

3

SEPTIEMBRE 2018



Fiscalidad, Big Data y Análisis de Redes:

Un estudio introductorio de la
red mundial de convenios de
doble imposición

Santiago Díaz de Sarralde Miguez

Fiscalidad, Big Data y Análisis de Redes:

Un estudio introductorio de la red mundial de convenios de doble imposición

Santiago Díaz de Sarralde Miguez.
Director de Estudios e Investigaciones. CIAT

Fiscalidad, Big Data y Análisis de Redes:

Un estudio introductorio de la red mundial de convenios de doble imposición

© 2018 Centro Interamericano de Administraciones Tributarias (CIAT)

ISSN: 2219-780X

Propiedad Intelectual

Se autoriza la reproducción total o parcial de esta obra por cualquier medio o procedimiento, conocido o por conocer, siempre que se cite adecuadamente la fuente y los titulares del Copyright.

Contenido

1. Introducción	7
2. Introducción al Análisis de Redes	8
3. La red internacional de Convenios de Doble Imposición	15
4. Otros ámbitos de aplicación del Análisis de Redes en las Administraciones Tributarias	21
Referencias bibliográficas	24
Anexo 1. Clusters Grafo no direccional Convenios Doble Imposición	25

Tablas y Gráficos

Tablas

Tabla 1. Ejemplo gravamen internacional dividendos	8
Tabla 2. Ejemplo de análisis de red de convenios internacionales	12
Tabla 3. Análisis de la red internacional de convenios de doble imposición.	17
Tabla 4. Análisis de la red internacional de convenios de doble imposición con consideración de tasas.	19

Gráficos

Gráfico 1.	Ejemplo gravamen internacional dividendos	9
Gráfico 2.	Treaty Shopping	9
Gráfico 3.	Ejemplo red convenios internacional	10
Gráfico 4.	Ejemplo de presentación de red convenios internacional	11
Gráfico 5.	Ejemplo de transformación de la red de convenios internacionales	14
Gráfico 6.	La red internacional de convenios de doble imposición en "bruto"	15
Gráfico 7.	La red internacional de convenios de doble imposición con presentación organizada.	16
Gráfico 8.	Ejemplo de grafo de cluster.	18
Gráfico 9.	Ejemplo de detección de tramas potencialmente fraudulentas.	22
Gráfico 10.	La naturaleza multidisciplinar y las difusas fronteras entre las técnicas de análisis de datos.	23

Anexo 1

Clusters 1, 2 y 3	26
Clusters 4, 5 y 6	28

1. Introducción

Entre los, aproximadamente, 200 países y jurisdicciones existentes en el mundo pueden trazarse unos 40.000 posibles caminos por los que los beneficios se pueden trasladar entre dos países diferentes. En la actualidad hay más de 3.500 convenios bilaterales de doble imposición, además de acuerdos regionales (Unión Europea, CARICOM, West African Economic and Monetary Union, etc.) que alteran las condiciones de gravamen de dividendos, intereses o royalties, modificando más de 7.000 de esos posibles caminos, en general en un sentido favorable. Si, además, tenemos en cuenta que para el traslado de los beneficios entre jurisdicciones se pueden utilizar cientos de miles de caminos indirectos, aprovechando las diferencias en fiscalidad de forma justificable o como una modalidad de abuso de tratados (treaty shopping), la complejidad que ofrece el panorama de la fiscalidad internacional puede parecer inmanejable.

Sin embargo, en los últimos años la coincidencia del desarrollo de las capacidades de almacenamiento y procesamiento de datos, junto a la mayor disponibilidad de los mismos en muchos sectores, ha provocado un desarrollo exponencial de las técnicas matemáticas, estadísticas e informáticas para su análisis.

El paradigma que actualmente define la disponibilidad de bases de datos complejas y de dimensiones masivas se conoce como Big Data, mientras que, entre las nuevas técnicas para su análisis sobresale por la novedad de su enfoque el Análisis de Redes.

En este documento nuestro objetivo es ofrecer, en primer lugar, una introducción al uso del Análisis de Redes en el estudio de las bases de datos masivas que pueda ser de utilidad para las Administraciones Tributarias, utilizando como ejemplo ilustrativo el análisis de la red mundial de convenios de doble imposición.

El siguiente apartado ofrece una breve introducción al Análisis de Redes y justifica su potencial utilidad en el ámbito de la Administración Tributaria. A continuación, nos centraremos en su aplicación a la red mundial de convenios de doble imposición, describiendo sus características y analizando la posición de los diferentes países en dicha red para el caso particular de la transferencia internacional de dividendos.

Por último, recopilaremos otras experiencias y campos de aplicación de las nuevas técnicas de Análisis de Redes en el ámbito de la administración tributaria.

2. Introducción al Análisis de Redes¹

Un ejemplo sencillo: Una empresa X residente en el país C tiene una filial en el país A. Tanto en A como en C la retención en la fuente por dividendos distribuidos a no residentes es del 10%. Entre A y C no existe convenio de doble imposición, mientras que sí existe entre A y B (con una tasa reducida para dividendos pagados a no residentes del 5%) y entre B y C (con una tasa reducida del 1%). En todos los países la distribución interna de dividendos no está gravada.

¿Qué opciones tiene esta empresa X para repatriar sus dividendos procedentes de A? ¿Hay posibilidades de Treaty Shopping? ¿Cómo les podría interesar negociar a los distintos países su red de convenios o modificar su política de retenciones en la fuente?

Para su análisis los parámetros del problema pueden sintetizarse mediante una matriz y representarse mediante un grafo² en el que cada país estaría simbolizado por un nodo o vértice (nodes/vertex) y las relaciones entre estos mediante arcos o enlaces (edges):

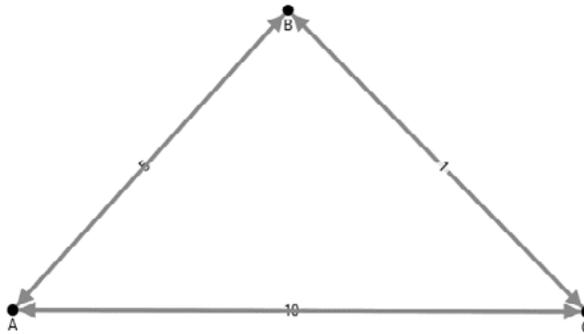
Tabla 1. Ejemplo gravamen internacional dividendos

		RESIDENCIA		
		A	B	C
FUENTE	A	0	5	10
	B	5	0	1
	C	10	1	0

Fuente: Elaboración propia

- 1 El origen del análisis matemático de las redes suele atribuirse a Euler y a su análisis del problema de los puentes de Königsberg (actualmente Kaliningrado), una ciudad dividida en cuatro partes por el río Pregel y que contaba con 7 puentes. Euler probó que era imposible trazar una ruta que permitiese recorrer los siete puentes cruzando cada uno de ellos sólo una vez.
- 2 Grafos representados mediante el complemento de Excel NodeXL.

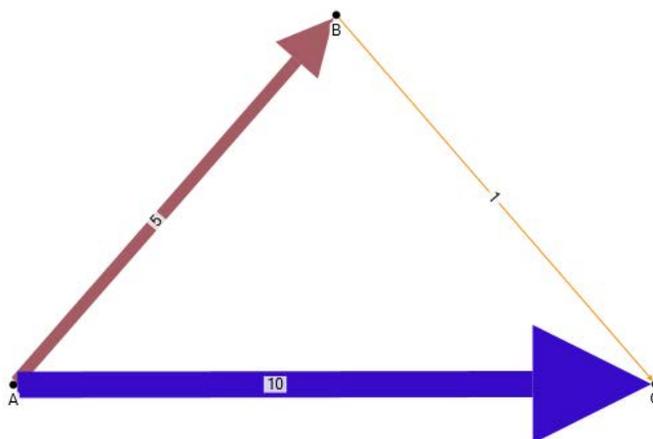
Gráfico 1. Ejemplo gravamen internacional dividendos



Fuente: Elaboración propia con NodeXL.

En este marco de análisis la determinación, por ejemplo, del camino más corto (entendido como el de menor tributación) desde A hasta C resulta obvio: podría resultar rentable utilizar una empresa intermediaria (conduit company) situada en B para aprovechar la menor tributación derivada de la red de convenios (rebajando de 10 a $5 + 1$ las tasas soportadas).

Gráfico 2. Treaty Shopping

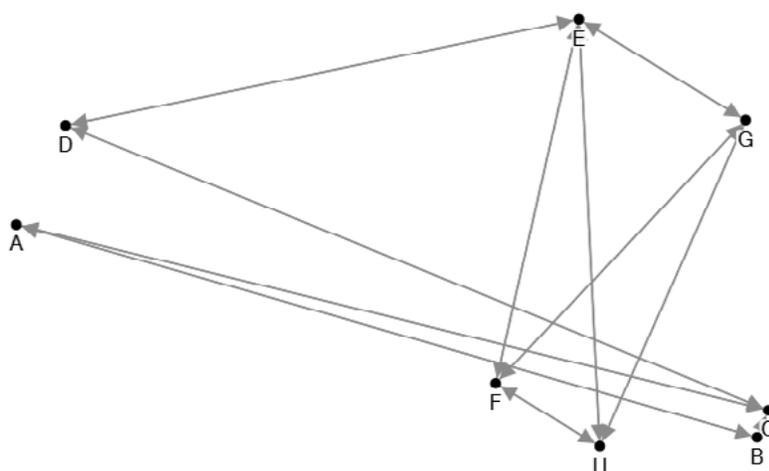


Fuente: Elaboración propia con NodeXL.

Desde el punto de vista de la Administración Tributaria del país A, la identificación de las estrategias posibles permite plantearse la renegociación de los convenios (por ejemplo, eliminando o aumentando la tasa reducida entre A y B o negociando un tratado con C que hiciese competitiva esa ruta) o la introducción de cláusulas anti-abuso.

La utilidad de esta aproximación sistemática al análisis de los datos se multiplica exponencialmente cuando aumenta el número de actores, al mismo tiempo que se amplían los indicadores métricos disponibles. Imaginemos una red ampliada de convenios de doble imposición³ como la representada en el gráfico 3.

Gráfico 3. Ejemplo red convenios internacional



Created with NodeXL (<http://nodexl.codeplex.com>)

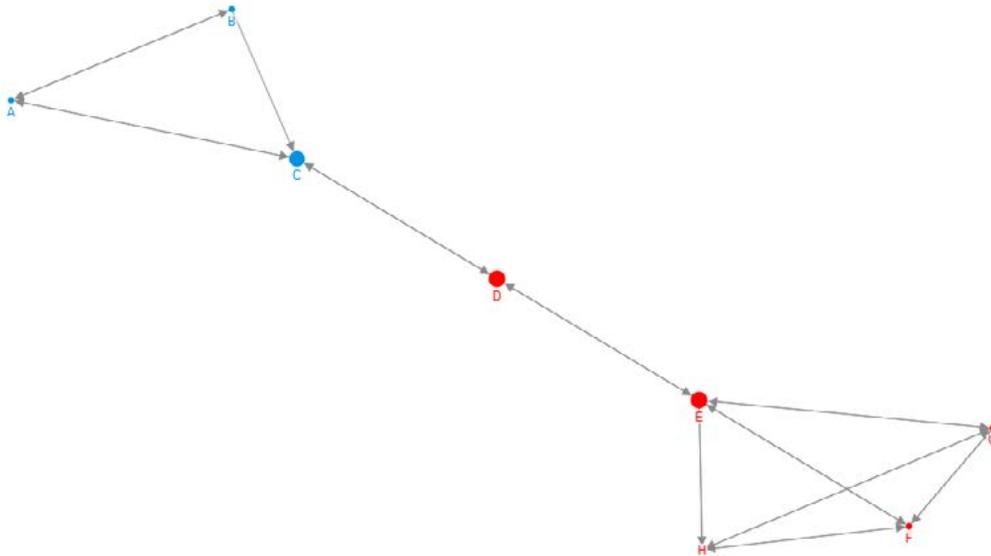
Fuente: Elaboración propia con NodeXL.

En este caso la estructura interna de la red y las posibles estrategias son más difíciles de identificar a simple vista. Afortunadamente el software disponible permite reorganizar la representación gráfica de acuerdo con múltiples pautas a través algoritmos (layout algorithms) que facilitan la comprensión de las relaciones. En este caso, la aplicación del método Harel-Koren Fast Multiscale Layout (basado en lo que se denominan “fuerzas” de atracción y repulsión, determinadas por las conexiones entre los nodos), la formación de grupos o clusters (representados por colores) y la diferenciación del tamaño de los nodos en función de su “centralidad”⁴ nos proporcionan una imagen mucho más clara de las relaciones (Gráfico 4).

3 Podemos suponer que el resto de conexiones aun siendo posible no fuese rentable al no disponer de convenio y verse sometido el traslado de dividendos a tasas de retención más elevadas. Adicionalmente, como caso particular para introducir variabilidad, la conexión desde B a C se considera unidireccional. El peso o ponderación de las nuevas aristas es de 1, excepto entre C y D (5).

4 En este caso la Centralidad de Intermediación cuya interpretación y cálculo se explican más adelante.

Gráfico 4. Ejemplo de presentación de red convenios internacional



Fuente: Elaboración propia con NodeXL

De esta forma, el papel que tienen las redes de convenios por grupos de países y, al mismo tiempo, la importancia relativa de cada uno de ellos en el conjunto de las conexiones resulta mucho más clara (las técnicas de representación permiten su mejora de forma extraordinariamente flexible, dando colores y dimensiones a los nodos y los enlaces en función de múltiples parámetros). Los países A, B y C conformarían un grupo, todos ellos con una importancia menor respecto al resto en términos de convenios en vigor, mientras que E, F, G y H forman otro cluster más integrado. De manera individual, C, D y E juegan un papel esencial de puente en la comunicación entre los diferentes países, siendo E entre ellos el que concentran mayor número de conexiones directas y estratégicas, mientras que el D, pese a su escaso número de relaciones, maximiza su relevancia gracias a su posición de intermediación.

Conforme se complique nuestro objeto de estudio es posible que la mera representación gráfica, pese a su utilidad como primera aproximación, no sea suficiente para analizar las conexiones con exactitud. Para ello es necesario recurrir a los diferentes indicadores o estadísticas que resumen diversas características del conjunto de la red, de los grupos de nodos y de cada uno de los vértices individualmente.

En el ejemplo anterior las características básicas globales de la red de convenios serían:

- Diámetro (Diameter): 5; el máximo número de conexiones necesarias para enlazar los nodos más lejanos.
- Densidad de Grafo (Graph density): 0,375; proporción del número de enlaces existentes (21) respecto al total de las posibles ($n*(n-1) = 8*7=56$).
- Longitud media de los caminos (Average path length): 2,321; la media de los caminos más cortos entre todos los países.

La interpretación de estas estadísticas varía según el contexto, siendo mucho más útiles si se tienen mediciones de redes similares o de la misma red en diferentes momentos del tiempo para poder comparar. En todo caso, nos muestran cómo de completa y “eficiente” es la red de contactos.

Las principales estadísticas individuales que podemos manejar aparecen resumidas en la tabla 2.

Tabla 2. Ejemplo de análisis de red de convenios internacionales

Países	Grado de entrada	Grado de salida	Excentricidad	Centralidad de Intermediación	Centralidad de Cercanía	Centralidad de Eigenvector	Page rank	Coefficiente de clustering
A	2	2	4	0,14	0,37	0,13	0,11	0,50
B	1	2	4	0,00	0,37	0,06	0,06	1,00
C	3	2	3	0,48	0,47	0,24	0,14	0,33
D	2	2	3	0,57	0,54	0,42	0,12	0,00
E	4	4	4	0,57	0,54	1,00	0,18	0,50
F	3	3	5	0,00	0,41	0,89	0,13	1,00
G	3	3	5	0,00	0,41	0,89	0,13	1,00
H	3	3	5	0,00	0,41	0,89	0,13	1,00

Fuente: elaboración propia con Gephi (<https://gephi.org/>).

De forma muy resumida, los grados de entrada y salida (in-degree, out-degree) muestran el número de relaciones direccionales⁵ que inciden en cada nodo, en nuestro caso los caminos de entrada o salida para los dividendos. La excentricidad (eccentricity) mide el número máximo de pasos que debe realizarse desde cada nodo para alcanzar el punto más alejado en la red. Son las medidas más simples y una primera aproximación (muy matizable, como veremos) a la importancia (mayor grado; menor excentricidad) y papel de cada nodo en la red⁶.

⁵ Por supuesto estas técnicas también pueden aplicarse a redes no dirigidas.

⁶ En la tabla se ofrecen los valores sin ponderar, esto es, sin tener en cuenta el valor de los enlaces (en nuestro caso, las tasas de los convenios).

Las medidas de Centralidad ofrecen diferentes perspectivas del papel de cada nodo (país) en las conexiones internas de la red (la circulación de los dividendos minimizando potencialmente la tributación internacional gracias al aprovechamiento de los convenios).

La Centralidad de Intermediación (*Betweenness centrality*) mide la frecuencia con la que cada nodo se encuentra en el camino más corto entre cualesquiera otros dos nodos (su valor en este caso está normalizado entre cero y uno, dividiendo por el número total de posibles caminos). La de Cercanía (*Closeness centrality*) es la inversa de la media de los caminos más cortos desde cada nodo a todos los demás⁷. Estas dos medidas tienen una relación inversa con las ponderaciones de cada enlace, en este caso con las tasas a pagar, con lo que a menores tasas establecidas en los convenios correspondería un mayor valor del índice de centralidad. Debido a la facilidad de interpretación de estos indicadores son los más utilizados, en particular el de Intermediación, en los análisis de las redes de tratados de doble imposición.

La Centralidad Eigenvector es una extensión de las medidas de centralidad en la que se pondera a los vecinos (nodos próximos) en función, a su vez, de su respectiva centralidad⁸. Utilizando esta información se puede obtener de forma iterativa el valor Pagerank, que mide la frecuencia o probabilidad de acabar en cada uno de los nodos teniendo en cuenta sus conexiones⁹. El método de "clasificación de páginas" de Google es el ejemplo más conocido de uso de este tipo de algoritmos, utilizando los datos de enlaces, búsquedas y frecuencia de visitas. Para el caso del análisis de las conexiones entre países a través de convenios (bilaterales por naturaleza) no resulta fácil interpretar estos indicadores sobre todo cuando se incluye el valor de las tasas, al ponderar en mayor medida las más elevadas, lo cual hace que su valor combine dos efectos contradictorios (mayor cuantos más convenios tengan los "vecinos", pero también cuanto más elevadas sean las tasas). Sin embargo, estas características pueden ser muy útiles en otros contextos, como el estudio de posibles redes de fraude contando con la información de los flujos comerciales y financieros.

Entre los indicadores mostrados en la tabla nos encontramos también con el Coeficiente de Clustering, el cual nos muestra para cada nodo la probabilidad de que los nodos próximos al mismo estén a su vez conectados entre ellos, con lo cual nos permite identificar los grupos más interconectados¹⁰ (en el gráfico identificados por colores). A nivel global, la red del ejemplo contaría con un Coeficiente de Clustering de 0,667 (estando sus valores entre 0 y 1 e indicando cuanto mayor sea una mayor frecuencia de posibles conexiones triangulares respecto a todas las posibles).

7 El tratamiento de los nodos inalcanzables representa una dificultad. En general este índice toma valor cero cuando se trata de un nodo aislado. La medida de centralidad armónica (*harmonic centrality*) es una variante que intenta simplificar este problema calculando la media de las inversas de los caminos más cortos.

8 Su valor está normalizado, siendo 1 para el nodo mejor conectado de acuerdo con estos parámetros.

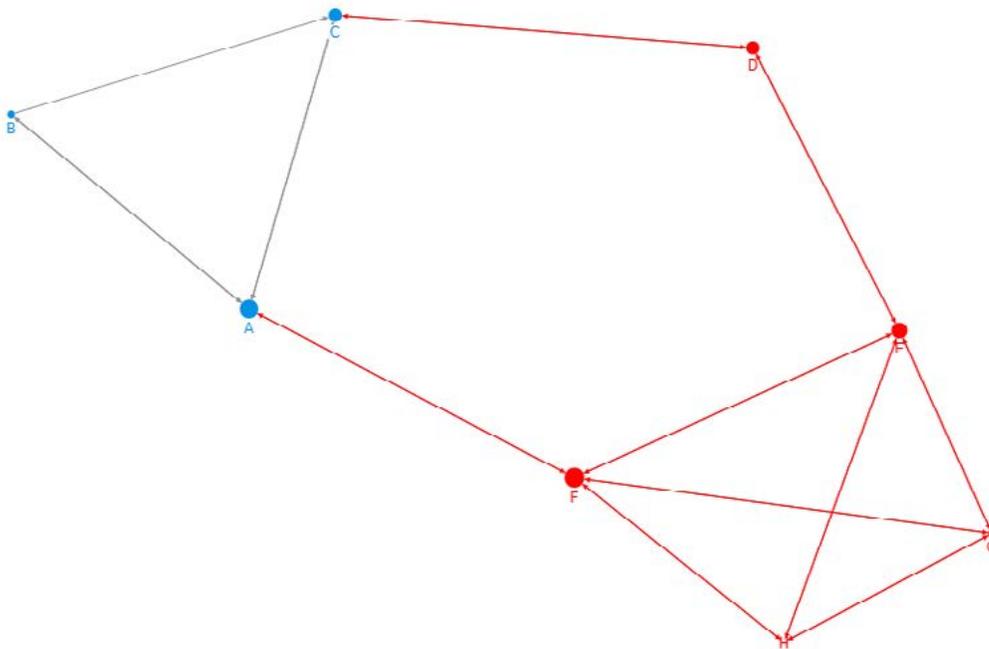
9 Los indicadores *HITS*, *Hub* y *Authority Scoring*, originalmente diseñados también para clasificar páginas web, valoran también la centralidad teniendo en cuenta el papel de cada nodo como receptor de enlaces procedentes de otros nodos relevantes (*Authority*) o como punto de enlace a otros nodos importantes (*Hub*).

10 De forma intuitiva implica analizar si nuestros "amigos" son, a su vez, amigos entre ellos. A este efecto se le denomina "*small-world*", lo que en español podría traducirse con la expresión "el mundo es un pañuelo". Técnicamente se procede a identificar todos los posibles "triángulos" y medir las conexiones realmente existentes en la red.

Como puede apreciarse, el cálculo de los indicadores refuerza, precisa y cuantifica las características ya observadas en el análisis visual del grafo (Gráfico 4), lo cual será de máxima utilidad en entornos más complejos, al mismo tiempo que expande las posibilidades de análisis a otros entornos de interés para las administraciones tributarias.

Adicionalmente, el análisis puede adoptar una perspectiva dinámica y/o de comparación de potenciales transformaciones. Por ejemplo, podrían analizarse las consecuencias de la firma de un nuevo convenio entre A y F (con tasa 1). La red se transformaría tal y como aparece en el Gráfico 5, aumentando notablemente el papel de ambos países en la circulación de los dividendos, que desplazarían a D y E en las clasificaciones basadas en los valores de los indicadores de centralidad. Por supuesto, a la vez, la red en su conjunto aumentaría su conectividad, reduciendo su diámetro a 3, aumentando la densidad a 0,411 y rebajando la longitud media de las conexiones a 1,786.

Gráfico 5. Ejemplo de transformación de la red de convenios internacionales

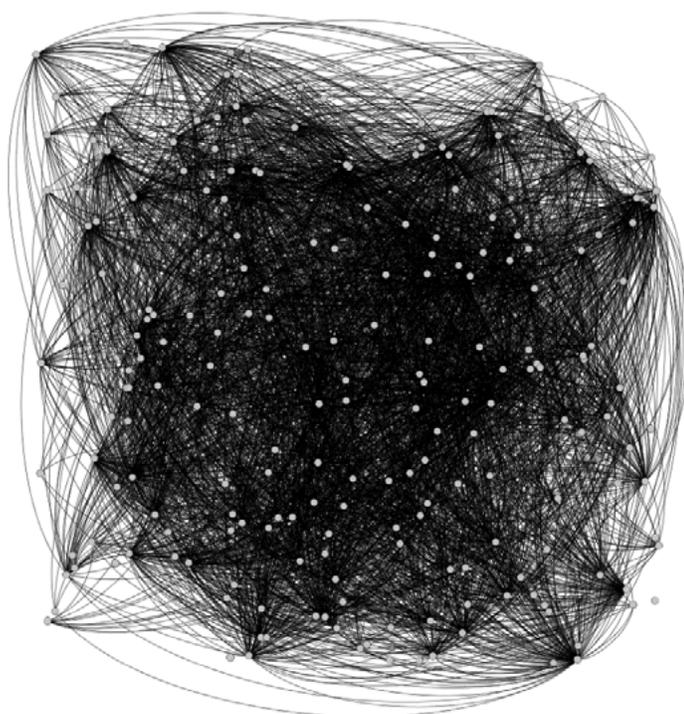


Fuente: Elaboración propia con NodeXL.

3. La red internacional de Convenios de Doble Imposición

Partiendo de la base de datos de convenios tributarios del IBFD¹¹ podemos configurar una red formada por 217 nodos (países/jurisdicciones) y 3.417 convenios (que suponen el doble de aristas al ser bilaterales). Su primera representación gráfica con disposición aleatoria (Gráfico 6) pone claramente de relieve su complejidad.

Gráfico 6. La red internacional de convenios de doble imposición en “bruto”

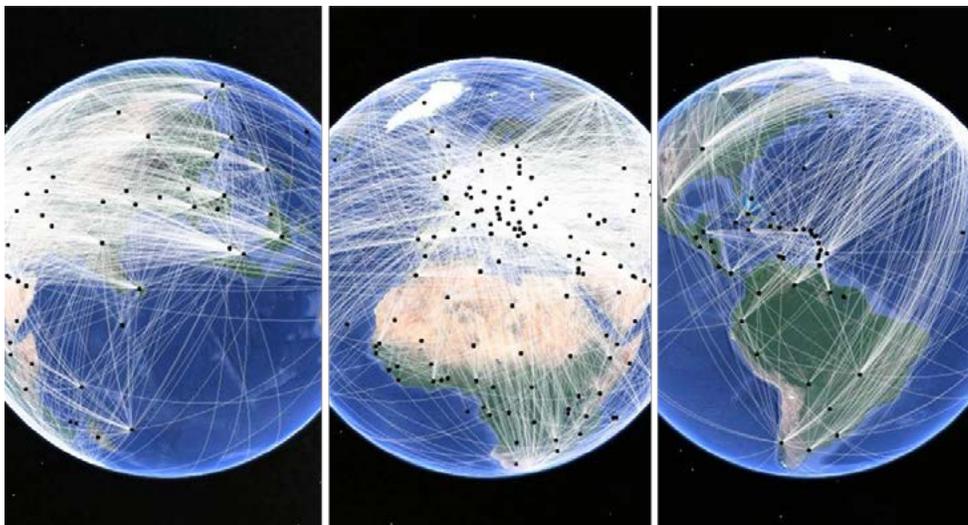
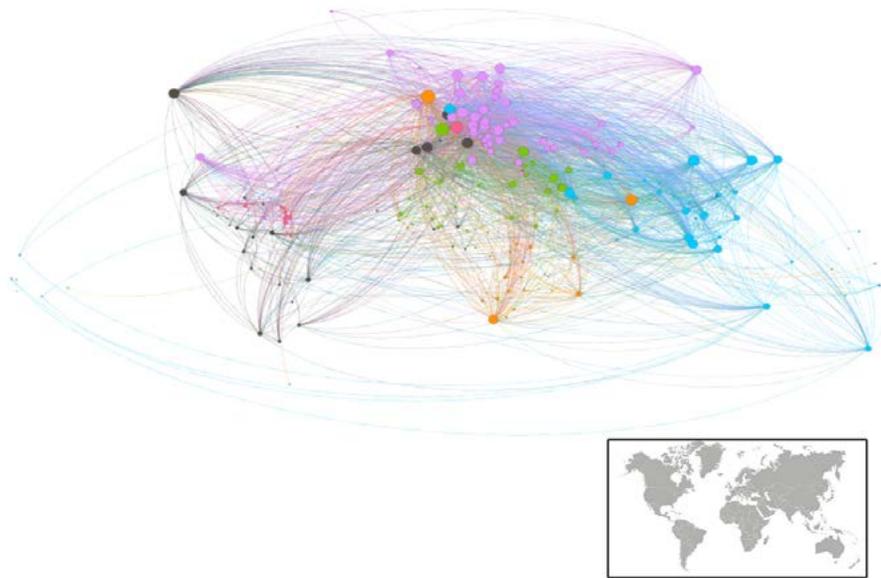


Fuente: Elaboración propia con Gephi.

11 <https://www.ibfd.org/IBFD-Products/Tax-Treaties-Database>

Sin embargo, las herramientas del análisis de redes y grafos nos permiten clarificar el panorama. El Gráfico 7 muestra esta misma información ordenada espacialmente (a través de los datos de longitud y latitud de cada país), dimensionando los nodos en función de su índice de Centralidad por Intermediación y clasificando por grupos (clusters) los países. A efectos divulgativos la información puede incluso transponerse en tres dimensiones a Google Earth e incluso presentarse de forma dinámica.

Gráfico 7. La red internacional de convenios de doble imposición con presentación organizada.



Fuente: Elaboración propia con Gephi y Google Earth.

A continuación, los indicadores cuantitativos nos permitirían un análisis más técnico. Partiendo de la información proporcionada por la mera existencia de convenios (sin tener en cuenta las tasas para cada una de las modalidades de rentas pasivas -dividendos; intereses; royalties-), obtendríamos un grafo no direccional sin ponderaciones. La Tabla 3 muestra los veinte primeros países en términos de número de convenios y según el valor del índice de Intermediación. En ambos casos nos encontramos con una mayoría de países europeos y con los mismos tres países encabezando la lista (Reino Unido, Francia y Suiza), sin embargo, las posiciones se alteran dependiendo del criterio de ordenación. Así, por ejemplo, Portugal sobresale en términos de intermediación o puente (4º) pese a su menor número de convenios (76). De igual forma, España, Australia o Sudáfrica mejoran su posición al tener en cuenta su potencial papel de enlace, mientras que otros como China descienden apreciablemente en la clasificación.

Tabla 3. Análisis de la red internacional de convenios de doble imposición.

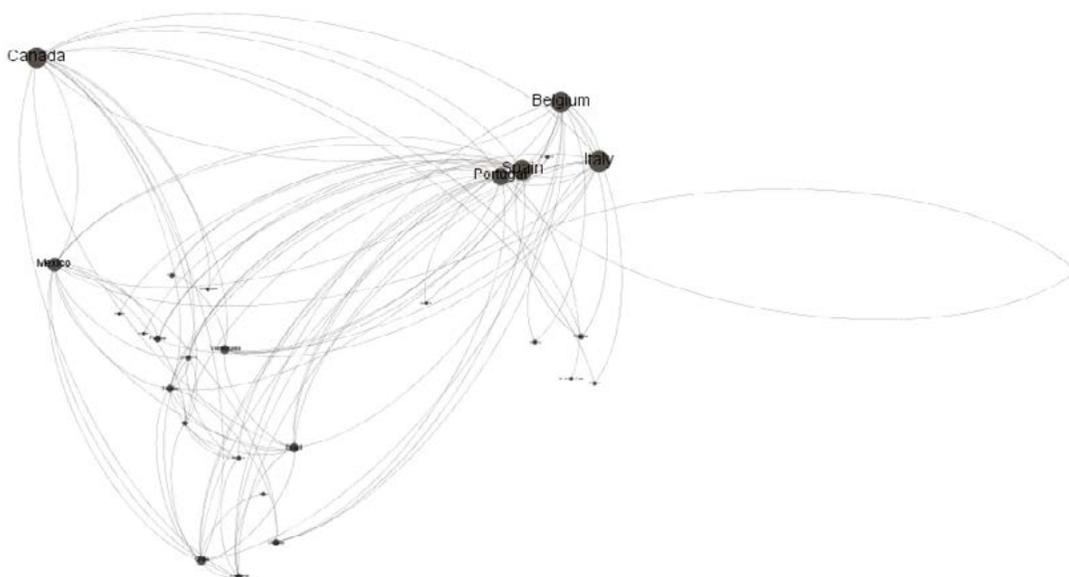
Países	Núm. Conv.	Países	Intermediación	Ejemplo Cluster	
United Kingdom	130	France	0,096537	Portugal	Nigeria
France	121	United Kingdom	0,090146	Spain	Paraguay
Switzerland	105	Switzerland	0,046105	Canada	Gabon
China (People's Rep.)	103	Portugal	0,030836	Belgium	Cape Verde
India	100	Spain	0,024188	Italy	Costa Rica
Italy	99	Norway	0,022729	Chile	Dominican Republic
Germany	96	India	0,022503	Mexico	S.Tome and Principe
Norway	95	Australia	0,020902	Brazil	El Salvador
Canada	94	South Africa	0,020666	Venezuela	
Sweden	94	Germany	0,020575	Ecuador	
Spain	93	United Arab Emirates	0,020328	Bolivia	
Netherlands	93	China (People's Rep.)	0,019311	Argentina	
Belgium	93	Canada	0,018979	Macau	
Korea (Rep.)	93	Sweden	0,014817	Colombia	
Austria	89	Netherlands	0,014239	Uruguay	
Romania	89	Belgium	0,01372	Peru	
Czech Republic	89	Singapore	0,012885	Andorra	
United Arab Emirates	87	Italy	0,012734	Panama	
Poland	87	Korea (Rep.)	0,011767	Cuba	
Finland	85	Denmark	0,011685	Ghana	

Fuente: Elaboración propia con Gephi.

Los clusters -formados, recordemos, en función de la similitud de su red de convenios y su (relativa) lejanía respecto a las redes de otros países e identificados por colores en el Gráfico 7- muestran una pauta de agrupación geográfica notable con algunas

excepciones consecuencia de las áreas específicas de influencia de ciertos países¹². Las últimas columnas de la Tabla muestran los países englobados en uno de los clusters, donde 21 de los 28 países son miembros de CIAT, con una mayoría de países americanos, pero incluyendo también, por ejemplo, a los también miembros Portugal, España, Italia o Nigeria.

Gráfico 8. Ejemplo de grafo de cluster.



Fuente: Elaboración propia con Gephi.

Esta primera aproximación contribuiría al presentar los datos de forma ordenada a identificar, por ejemplo, los posibles caminos para aprovechar los convenios de cara a minimizar la tributación a escala internacional. Sin embargo, sin tener en cuenta, al menos, las tasas estipuladas en cada uno de los tratados, las tasas de retención existentes en caso de no existir convenio y los acuerdos multilaterales regionales de reducción del gravamen la visión seguiría siendo incompleta.

La incorporación de esta información no resulta sencilla dado que la legislación y los tratados estipulan en la mayoría de los casos diversas tasas posibles en función de las circunstancias (grado de participación empresarial, la existencia de gravamen previo, la fuente de los royalties, el tipo de endeudamiento que genera los intereses, etc.). Para nuestro ejemplo simplificado, nos ceñimos al gravamen de los dividendos, suponiendo que es de aplicación la tasa más favorable posible (incluyendo los acuerdos regionales de la Unión Europea, CARICOM o la West African Economic and Monetary Union) y creamos una matriz de países fuente y residencia identificando las rutas más favorables en cada

¹² Por ejemplo, Suiza aparece encuadrado en un grupo junto a la mayoría de los países caribeños. Francia se engloba junto a gran parte de los países árabes y del norte de África. Reino Unido junto al sur de África y la India. El Anexo 1 detalla los principales *clusters* y su representación gráfica.

caso por comparación con las tasas generales de retención para eliminar los convenios no relevantes (las cuales prevalecen en caso de ser las menores posibles). Adicionalmente, en todos los casos añadimos un coste mínimo adicional (0,001%) para que el incremento de pasos en la repatriación de dividendos se vea al menos ligeramente penalizada reflejando los costes adicionales de tener presencia en diversas jurisdicciones¹³.

La Tabla 4 recoge la clasificación obtenida en función de la Centralidad de Intermediación, que, como hemos visto, refleja la posibilidad de que cada jurisdicción se encuentre en el camino más corto -de menor tributación- entre cualesquiera otros dos países.

Tabla 4. Análisis de la red internacional de convenios de doble imposición con consideración de tasas.

(1-20)	(21-40)	(41-60)	(61-80)	(81-100)	(101-120)	(121-140)	(141-160)	(161-180)	(181-200)
United Kingdom	South Africa	Monaco	Mauritania	Ukraine	Macedonia (FYR)	Montenegro	Ethiopia	Syria	Grenada
France	Moldova	Romania	Ghana	Senegal	Azerbaijan	Montserrat	Kenya	Georgia	Guinea
India	Portugal	Czech Republic	Liechtenstein	Croatia	Zambia	Turkmenistan	Panama	Andorra	Iraq
Niger	Iran	San Marino	Yemen	Fiji	Oman	Argentina	Antigua and Barbuda	Anguilla	Kiribati
Singapore	Belgium	Chile	Mexico	Israel	Kazakhstan	Ecuador	Mozambique	Bermuda	Liberia
Vietnam	Netherlands	Luxembourg	Pakistan	Trinidad and Tobago	Peru	Albania	Togo	Bhutan	Malawi
Mauritius	Norway	Turkey	Macau	Saudi Arabia	Uzbekistan	Bosnia and Herzegovina	Laos	British Virgin Islands	New Caledonia
Malaysia	Italy	Denmark	Ireland	Guernsey	Algeria	St. Kitts and Nevis	Nigeria	Cambodia	Samoa
Hungary	Jordan	Morocco	Nepal	Isle of Man	Lesotho	Kyrgyzstan	Rwanda	Cameroon	Sierra Leone
Canada	New Zealand	Austria	Egypt	Jamaica	Barbados	Tajikistan	Gambia	Chad	Solomon Islands
Germany	Belize	Kosovo	Myanmar	Ivory Coast	Seychelles	Mongolia	Paraguay	Congo (Rep.)	St. Martin
Spain	St. Lucia	Tunisia	Slovak Republic	Slovenia	Philippines	Uruguay	Congo (Dem, Rep.)	Costa Rica	St. Pierre and Miquelon
Kuwait	St. Vincent and the Grenadines	Russia	Sudan	Serbia	Iceland	Bangladesh	Namibia	Dominican Republic	Suriname
Qatar	Korea (Rep.)	Korea (Dem, People's Rep.)	Belarus	Central African Republic	Lebanon	Bolivia	Swaziland	El Salvador	São Tomé and Príncipe
Brazil	China (People's Rep.)	Brunei	Thailand	Madagascar	Aruba	Guyana	Uganda	Equatorial Guinea	Tanzania

¹³ Este ajuste sería necesario en todo caso para que los enlaces en los que el gravamen es cero pudiesen tenerse en cuenta, ya que si no la mayoría de los programas estándar para el análisis de redes no reconocen la existencia de esas rutas.

(1-20)	(21-40)	(41-60)	(61-80)	(81-100)	(101-120)	(121-140)	(141-160)	(161-180)	(181-200)
Malta	Sweden	Poland	Papua New Guinea	Greece	Curaçao	Benin	Dominica	Falkland Islands	Turkish Cyprus
Latvia	Hong Kong	Finland	Bulgaria	Netherlands Antilles	St, Maarten	GuineaBissau	Palestine	Faroe Islands	Tuvalu
Cyprus	Taiwan	Jersey	United States	Sri Lanka	Armenia	Colombia	TimorLeste	French Polynesia	French Guiana
Switzerland	Botswana	Libya	Japan	Cape Verde	Venezuela	Burkina Faso	United Arab Emirates	Gabon	Guadeloupe
Estonia	Indonesia	Cuba	Australia	Lithuania	Mali	Zimbabwe	Bahrain	Greenland	Martinique

Fuente: Elaboración propia con Gephi.

Por ejemplo, el Reino Unido tendría un valor de 0,045 en su índice de intermediación, lo que implicaría estar presente en el 4,5% de los caminos más cortos (de menor tributación) -sin tener en cuenta las tasas de los convenios y de retención generales, Tabla 3, aparecía en el 9,6% de los mismos-, seguido de Francia (3,4%) y la India (2,2%).

Entre los veinte primeros, siete de ellos (Reino Unido, Francia, Suiza, India, Alemania, Canadá y España) ya aparecían en la Tabla 3, por su elevado número de convenios. El otro factor relevante para alcanzar este papel de puente es tener una tasa cero de retención general -sin necesidad de convenio- para la repatriación de dividendos (entre los veinte primeros las únicas excepciones son Francia, Canadá, Alemania, España y Suiza).

El análisis podría seguir refinándose teniendo en cuenta el régimen de gravamen de los dividendos recibidos en cada país, así como su régimen de corrección de la doble imposición¹⁴, las listas de paraísos fiscales¹⁵, el gravamen de resto de rentas pasivas -intereses, royalties-¹⁶ o las condiciones de seguridad y estabilidad de cada país, entre otros factores. En todo caso, esta primera aproximación esperamos que sirva de ejemplo de la potencialidad de la técnica de Análisis de Redes en el ámbito de la Administración Tributaria.

¹⁴ Barrios et al (2009) ofrece un marco teórico para su tratamiento. Petkova et al (2018), Van 't Riet y Lejour (2014) aplican, con variantes, este enfoque.

¹⁵ Van 't Riet y Lejour (2014).

¹⁶ Nakamoto e Ikeida (2018).

4. Otros ámbitos de aplicación del Análisis de Redes en las Administraciones Tributarias

En su publicación *Advanced Analytics for Better Tax Administration: Putting Data to Work*, la OCDE (2016) identifica la aplicación del Análisis de Redes en diversos países (Irlanda, Malasia, Países bajos, Singapur¹⁷, Nueva Zelanda) para la prevención del fraude carrusel en el IVA, así como para la detección de otros tipos de fraude que implican a redes de agentes económicos, si bien advierte que no es todavía la práctica más habitual. De forma más concreta, el artículo seleccionado para el número 44 (2018) de la Revista de Administración Tributaria CIAT/IEF/AEAT por Ignacio González García de la Agencia Tributaria de España (*Analytics y Big Data. La Nueva Frontera*) analiza las características técnicas básicas de esta aproximación en comparación con las técnicas estadísticas más tradicionales y ofrece ejemplos concretos de casos de uso desarrollados en la AEAT para la detección de redes familiares extendidas, cálculo de la riqueza societaria de los contribuyentes teniendo en cuenta las participaciones indirectas y cruzadas, detección de tramas de corrupción y blanqueo o de estructuras de fraude, en lo que constituye una lectura muy recomendable. De igual forma la DIAN en Colombia ha estado utilizando estas técnicas para la detección de patrones de fraude y lavado de dinero en el sector de la comercialización del oro.

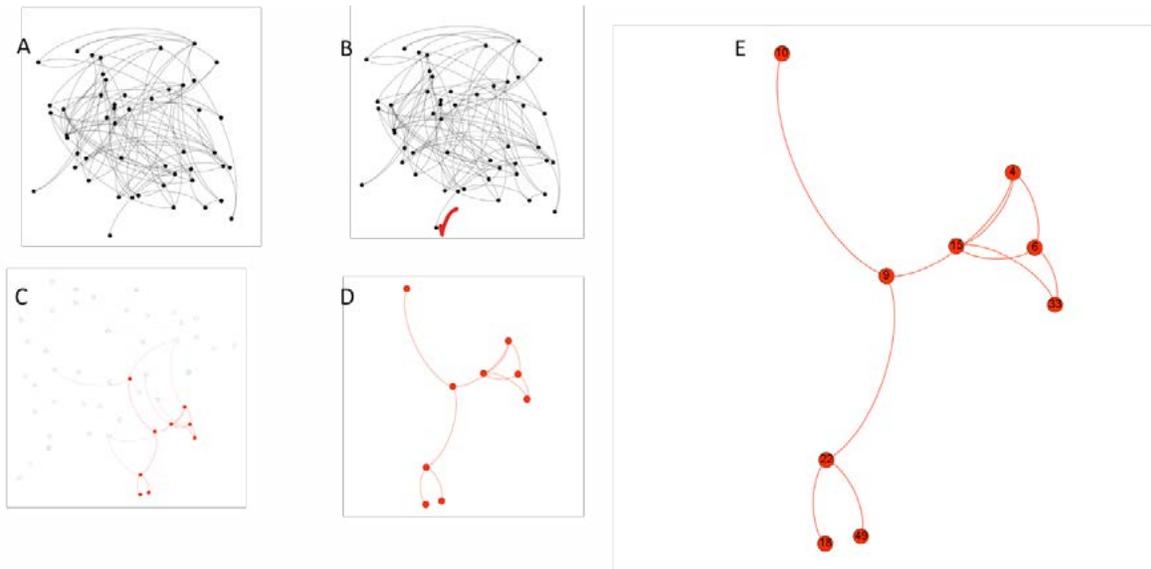
Dada la magnitud de la información que manejan las Administraciones Tributarias¹⁸ el ámbito de aplicación del Análisis de Redes es amplio. La flexibilidad que ofrece a la hora de organizar, visibilizar y categorizar la información lo convierte en una herramienta de trabajo óptimo, incluso para afrontar problemas tradicionales. Por ejemplo, si contamos con la información de transacciones económicas en una red (compras y ventas) y tenemos la sospecha de que una de las empresas tiene las características de una posible emisora de facturas falsas o de inicio de una red de fraude carrusel, el Análisis de Redes nos permite identificar sencilla y rápidamente las principales empresas conectadas para su posterior seguimiento y análisis. El Gráfico 8 muestra un ejemplo basado en una red de 50 nodos y 80 aristas generada aleatoriamente (A), en la que se identifica la empresa sospechosa (B), se efectúa una selección de las empresas más conectadas (C y D) con la misma (utilizando la identificación de clusters por modularidad, esto es de comunidades dentro de la red que se caracterizan por su alta densidad interna -muchas conexiones-, y su escasa conexión con el resto¹⁹) y se aíslan de las demás para su análisis mediante la aplicación de filtros (9 nodos, un 18% del total, y 10 aristas, el 12,35%). Adicionalmente se podrían utilizar los índices Pagerank o HITS para determinar la importancia y el papel de cada nodo en la red.

17 OCDE (2017), *The Changing Tax Compliance Environment and the Role of Audit*, destaca el caso de Singapur por el uso del Análisis de Redes para la elaboración de modelos de riesgo y selección de auditorías.

18 La introducción de la factura electrónica y de los sistemas de recopilación de la información individualizada e inmediata de compras y ventas aumenta exponencialmente las posibilidades de análisis.

19 Técnicamente se utiliza el método propuesto por Blondel et al (2008), altamente eficiente en la gestión de grandes bases de datos y regulable en su exigencia de cohesión interna de cada comunidad.

Gráfico 9. Ejemplo de detección de tramas potencialmente fraudulentas.



Fuente: Elaboración propia con Gephi.

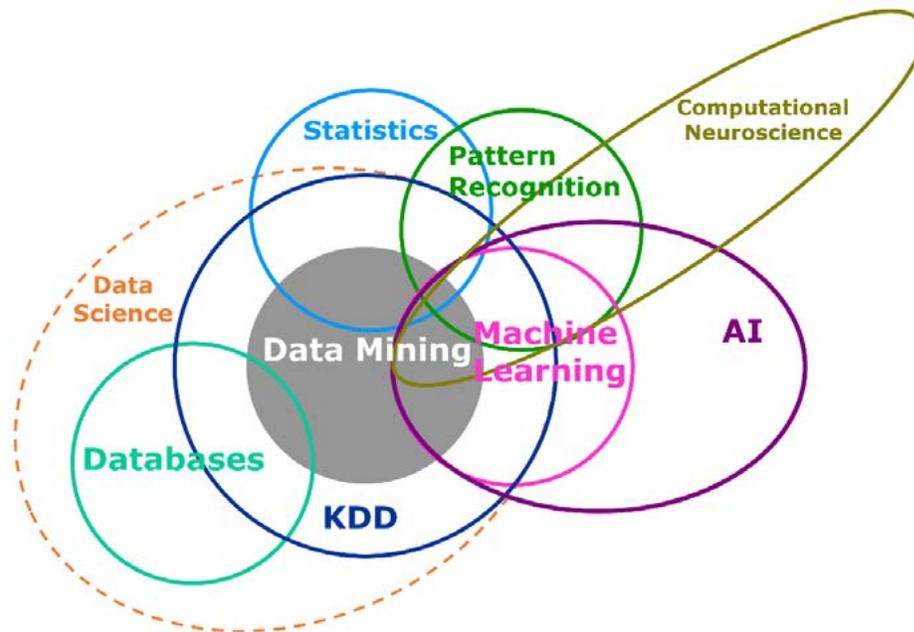
Las posibilidades de profundización en el análisis son numerosas: atribuyendo características adicionales a través de etiquetas (labels) para su filtrado y selección -antigüedad, inspecciones previas, localización geográfica, etc.-, diferenciando entre distintos tipos de nodos (gráfos multipartitos -multipartite graph-) -como, por ejemplo, personas físicas, jurídicas, holdings, etc.-, o analizando los cambios en el tiempo de las relaciones (grafos dinámicos a través de time-stamps). Siempre que los fraudes a combatir impliquen la creación de relaciones y redes²⁰ estas herramientas serán útiles para organizar y analizar la información²¹.

Por último, cabe resaltar que se trata de una técnica aséptica, podríamos decir, en el sentido de que no presupone una modelización teórica del fraude, sino que depende del conocimiento experto para orientar su uso e interpretación. El analista de sistemas tiene que trabajar codo con codo con el auditor fiscal para organizar la información, establecer los criterios de búsqueda y seleccionar los indicadores más adecuados. Esta supervisión puede complementarse con otras técnicas de análisis de datos, agrupadas en diversos campos y denominaciones que se solapan entre sí: machine learning; inteligencia artificial; data mining; Knowledge Discovery in Databases (KDD); etc.

²⁰ Algo que parece ser habitual en muchas situaciones reales. En inglés utilizan la expresión “*Birds of a feather flock together*”, trasladable al refrán español “Dime con quién andas y te diré quién eres”.

²¹ El libro de Baesens et al (2015) “*Fraud Analytics Using Descriptive, Predictive, and Social Network Techniques*” proporciona una introducción muy completa al uso de estas técnicas desde el punto de vista estadístico, incluyendo referencias a otras aplicaciones específicas de detección del fraude, por ejemplo en el ámbito de la Seguridad Social -“*GOTCHA! Network-Based Fraud Detection for Social Security Fraud*”, publicado de forma separada en Van Vlasselaer et al (2017)-. SAS ha desarrollado una aplicación denominada Social Network Analysis para el análisis y detección del fraude en los seguros mediante la aplicación de estas técnicas (https://www.sas.com/content/dam/SAS/en_us/doc/productbrief/sas-social-network-analysis-103857.pdf).

Gráfico 10. La naturaleza multidisciplinaria y las difusas fronteras entre las técnicas de análisis de datos.



Fuente: Hall et al (2014)

Estas técnicas permitirán amplificar la orientación experta del fiscalizador, rastreando las señales existentes en las bases de datos complejas respecto al comportamiento fraudulento a través de técnicas estadísticas y métodos recursivos para la optimización propulsados por la enorme capacidad y velocidad de cálculo de los equipos informáticos actuales (un ejemplo clásico es el uso de las denominadas redes neuronales para la detección de patrones de fraude, del mismo modo que se utilizan para el reconocimiento de imagen o voz)²².

²² En último extremo y debido al auge de la inteligencia artificial algunos plantean que podría llegarse incluso a la detección de comportamiento fraudulentos sin la supervisión experta ("*unsupervised machine learning*"), evitando así el sesgo y las limitaciones impuestas por los modelos que se basan en las pautas de fraude ya previamente conocidas. En mi opinión, sin embargo, no resulta útil enredarse en argumentaciones radicales sobre este tipo de planteamientos como enfrentamiento entre la nueva "inteligencia artificial" y la clásica "inteligencia" tributaria de los expertos y auditores. Siempre habrá que determinar externamente los objetivos y parámetros de búsqueda, validar los resultados obtenidos en función de la legislación y presentar las pruebas ante las instancias judiciales. Objetivos, legislación y jueces, todavía, muy humanos.

Referencias bibliográficas

Baesens, B., Van Vlasselaer, V., & Verbeke, W. (2015). Fraud Analytics Using Descriptive, Predictive, and Social Network Techniques.

Barrios, S., Huizinga, H., Laeven, L. and Nicodè, G. (2009) "International Taxation and Multinational Firm Location Decisions", Economic Papers 356 | January 2009 European Commission

Blondel, V., Guillaume, J.L, Lambiotte, R. and Lefebvre, E. (2008) "Fast unfolding of communities in large networks", Journal of Statistical Mechanics: Theory and Experiment. IOP and SISSA journal

González García, I., (2018) "Analytics y Big Data. La Nueva Frontera", Revista de Administración Tributaria CIAT/IEF/AEAT, número 44, 2018.

Hall, P., Dean, J., Kaynar Kabul, I. Silva. J (2014), "An Overview of Machine Learning with SAS® Enterprise Miner™". Paper SAS313-2014. SAS Institute Inc.

Nakamoto, T., Ikeda, Y. (2018) "Identification of Conduit Countries and Community Structures in the Withholding Tax Networks" <https://arxiv.org/abs/1806.00799>

OECD (2016), Advanced Analytics for Better Tax Administration: Putting Data to Work, OECD Publishing, Paris. <http://dx.doi.org/10.1787/9789264256453-en>

OECD (2017), The Changing Tax Compliance Environment and the Role of Audit, OECD Publishing, Paris. <http://dx.doi.org/10.1787/9789264282186-en>

Petkova, K., Stasio, A., Zagler, M. (2018) "On the relevance of double tax treaties", WU International Taxation Research Paper Series, No. 2018 - 05

Van 't Riet, M., Lejour, A (2014) "Ranking the Stars. Network Analysis of Bilateral Tax Treaties" CPB Discussion Paper | 290.

Van Vlasselaer, V., Eliassi-Rad, T., Akoglu, L., Snoeck, M., Baesens, B., (2017), "GOTCHA! Network-Based Fraud Detection for Social Security Fraud", Management Science Volume 63, Issue 9, September 2017

Wujek, B., Hall, P. and Güneş, F. (2016), "Best Practices for Machine Learning Applications", Paper SAS2360-2016. SAS Institute Inc.

Anexo 1. Clusters Grafo no direccional Convenios Doble Imposición

Cluster 1: 26,27% de los países			Cluster 2: 16,59%		Cluster 3: 15,67%	
Albania	Greece	Mongolia	Algeria	Mali	Australia	Netherlands
Armenia	Greenland	Montenegro	Bahrain	Mauritania	Bangladesh	New Zealand
Aruba	Guernsey	Netherlands Antilles	Benin	Morocco	Bhutan	Oman
Austria	Hungary	Norway	Burkina Faso	New Caledonia	Brunei	Pakistan
Azerbaijan	Iceland	Poland	Cameroon	Niger	Cambodia	Papua New Guinea
Belarus	Ireland	Romania	Central African Republic	Palestine	China (People's Rep,)	Philippines
Bermuda	Isle of Man	Russia	Congo (Rep,)	Qatar	Comoros Islands	Samoa
Bosnia and Herzegovina	Israel	San Marino	Egypt	Saudi Arabia	Cook Islands	Singapore
British Virgin Islands	Jersey	Serbia	Ethiopia	Senegal	Fiji	Sri Lanka
Bulgaria	Kazakhstan	Slovak Republic	France	St, Pierre and Miquelon	Hong Kong	Suriname
Cayman Islands	Kosovo	Slovenia	French Polynesia	Sudan	Indonesia	Taiwan
Croatia	Kyrgyzstan	Sweden	Guinea	Syria	Japan	Thailand
Cyprus	Latvia	Tajikistan	Guinea-Bissau	Togo	Kiribati	United Arab Emirates
Czech Republic	Liberia	Turkmenistan	Iran	Tunisia	Korea (Rep,)	Vietnam
Denmark	Liechtenstein	Ukraine	Iraq	Turkey	Laos	
Estonia	Lithuania	United States	Jordan	Yemen	Malaysia	
Faroe Islands	Luxembourg	Uzbekistan	Kuwait		Maldives	
Finland	Macedonia (FYR)		Lebanon		Marshall Islands	
Georgia	Malta		Libya		Myanmar	
Germany	Moldova		Madagascar		Nepal	

Cluster 4: 12,9%		Cluster 5: 11,52%		Cluster 6: 6%
Andorra	Gabon	Botswana	Seychelles	Anguilla
Argentina	Ghana	Congo (Dem, Rep,)	Sierra Leone	Antigua and Barbuda
Belgium	Italy	Falkland Islands	Solomon Islands	Barbados
Bolivia	Macau	Gambia	South Africa	Belize
Brazil	Mexico	India	Swaziland	Dominica
Canada	Nigeria	Kenya	Tanzania	Grenada
Cape Verde	Panama	Lesotho	Tuvalu	Guyana
Chile	Paraguay	Malawi	Uganda	Jamaica
Colombia	Peru	Mauritius	United Kingdom	St, Kitts and Nevis
Costa Rica	Portugal	Monaco	Zambia	St, Lucia
Cuba	Sao Tomao and Principe	Montserrat	Zimbabwe	St, Vincent and the Grenadines
Dominican Republic	Spain	Mozambique		Switzerland
Ecuador	Uruguay	Namibia		Trinidad and Tobago
El Salvador	Venezuela	Rwanda		

Clusters 4, 5 y 6





Serie **Documentos de Trabajo**



CIAT

Secretaría Ejecutiva del CIAT

Apartado: 0834-02129, Panamá, República de Panamá

Teléfono: (507) 3072428

Fax: (507) 2644926

Correo electrónico: ciat@ciat.org

Sitio Web: www.ciat.org