

Análisis de la organización social del PDI del Campus de Terrassa de la UPC mediante redes de coautoría de artículos científicos.

GRADO EN INGENIERÍA MECÁNICA



UNIVERSITAT POLITÈCNICA DE CATALUNYA
BARCELONATECH

**Escola Superior d'Enginyeries Industrial,
Aeroespacial i Audiovisual de Terrassa**

FECHA DE ENTREGA: 30 de Junio de 2020

ESTUDIANTE: Carlos González Berengüí

DIRECTOR: Martí Rosas Casals

ÍNDICE

1.	INTRODUCCIÓN.....	5
1.1	OBJETO.....	5
1.2	ALCANCE	6
1.3	REQUERIMIENTOS	6
1.4	JUSTIFICACIÓN Y UTILIDAD.....	7
2.	DESARROLLO	8
2.1	ANTECEDENTES	8
2.2	PLANTEAMIENTO	8
2.2.1	Redes Complejas.....	9
2.2.2	Redes Sociales.....	16
2.2.3	Medidas de centralidad	18
2.2.4	Métricas globales	23
2.2.5	Tipos de redes	25
2.2.6	Redes colaborativas	35
2.2.7	Redes jerárquicas.....	36
2.3	HIPÓTESIS.....	38
2.3.1	Hipótesis de jerarquía	38
2.3.2	Hipótesis de colaboración.....	38
3.	DESARROLLO DE LA RED DE COAUTORIAS	39
3.1	CREACIÓN DE LA BASE DE DATOS EN FORMATO EXCEL	39
3.2	LIMPIEZA DE DATOS	40
3.2.1	Errores encontrados	41
3.2.2	Soluciones adoptadas	43
3.2.3	Creación de aristas.....	44
3.3	CREACIÓN DE LA RED MEDIANTE NODEXL.....	45
3.4	ANÁLISIS DE LAS DISTRIBUCIONES DE PROBABILIDAD	49
3.6	REDES EGO	51
3.6.1	Análisis de las redes ego de los 10 nodos con más grado	52
3.7	ANÁLISIS DE LAS TOPOLOGÍAS DE RED	55
3.7.1	Topología de las redes ego analizadas	56

4. RESULTADOS Y DISCUSIÓN	58
4.1 MÉTRICAS GENERALES	58
4.1.1 Componentes conectados	59
4.1.2 Vértices y aristas	60
4.1.3 Grado promedio	61
4.1.4 Modularidad	63
4.1.5 Densidad de la red	64
4.1.6 Centralidad del vector propio y PageRank	66
4.1.7 Distancia Geodésica	67
4.1.8 Intermediación	68
4.1.9 Cercanía y coeficiente de agrupamiento	69
4.2 ESTRUCTURA GLOBAL DE LA RED.....	72
4.3 ESTRUCTURAS LOCALES DE LA RED.....	79
5. CONCLUSIONES.....	85
5.1 COMPROBACIÓN DE HIPÓTESIS.....	86
5.2 COMENTARIOS ADICIONALES	88
6. BIBLIOGRAFÍA.....	89
7. ANEXOS.....	91
7.1 Topología de las 10 redes ego con más grado en 2018	91
7.1.1 "Ego network" de Vicenç Puig en el grupo G5.....	91
7.1.2 "Ego network" de Fatos Xhafa en el grupo G6.....	91
7.1.3 "Ego network" de TzankoTzanov en el grupo G1.....	92
7.1.4 "Ego network" de Kestutis Staliunas en el grupo G3.	93
7.1.5 "Ego network" de Cristina Masoller en el grupo G3.	93
7.1.6 "Ego network" de Jordi Morato en el grupo G1.....	94
7.1.7 "Ego network" de Jaume Pujol en el grupo G2.	95
7.1.8 "Ego network" de Joan Montanya en el grupo G8.....	95
7.1.9 "Ego network" de Oliva Asensio en el grupo G7.....	96
7.1.10 "Ego network" de Meritxell Vilaseca en el grupo G2.	97
7.2 Topología del resto de grupos de 2018.....	98
7.2.1 "Ego network" de Jasmina Casals Terre en el grupo G4.....	98
7.2.2 "Ego network" de Jaime Martí Herrero en el grupo G9	98
7.2.3 "Ego network" de Jordi Martorell en el grupo G10.....	99



7.2.4 "Ego network" de Lluís Gil en el grupo G11	100
7.2.5 "Ego network" de Miquel Casals en el grupo G12	100
7.2.6 "Ego network" de Anna Rius en el grupo G13	101
7.2.7 "Ego network" de Oriol Lordan en el grupo G14	102
7.2.8 "Ego network" de María Carmen Riva en el grupo G15	102
7.2.9 "Ego network" de Josep Pou en el grupo G16	103
7.2.10 "Ego network" de Meritxell Martí en el grupo G17	104
7.2.11 "Ego network" de Francisco Javier Oliver en el grupo G18	104
7.2.12 "Ego network" de Martí Rosas Casals en el grupo G19	105
7.2.17 "Ego network" de Ramón Quintanilla en el grupo G24	106

1. INTRODUCCIÓ

1.1 OBJETO

Este TFG tiene por objetivo el análisis de la organización social del personal docente e investigador (PDI) del Campus de Terrassa de la Universitat Politècnica de Catalunya (UPC) mediante la creación de redes de coautoría de artículos científicos, utilizando datos de dominio público para calificar los actores, en este caso el PDI del Campus de Terrassa, y sus relaciones de colaboración acumulada en función del número de coautoría de artículos científicos, y establecidas a lo largo del tiempo, en este caso desde 2014 hasta 2018. Cuando hablamos de organización social, y analizada mediante redes de coautoría de artículos científicos, nos referimos fundamentalmente al tipo de relación, de colaboración y/o jerárquica, que se puede establecer entre los y las diferentes autores.

Para ello, se ha procedido a la creación de una base de datos (en formato tabla de cálculo) ¹ limpia de errores. Se creará la lista de aristas "autor/a - autor/a" para cada año, de tal forma que podamos exportarla a un software especializado en el análisis de redes.² Mediante este software, analizaremos las principales medidas de centralidad de las redes de cada año y profundizaremos en este análisis mediante las diferentes funciones de distribución de probabilidad de cada medida de centralidad utilizada. Finalmente, se hará un estudio de las redes "ego" (i.e., primeros vecinos) de los autores con mayor importancia dentro de la red, y se relacionarán estos datos con la topología de la red, viendo si los resultados obtenidos guardan alguna relación con esta topología. Finalmente, con los datos obtenidos, se pretende concluir que tipo de estructura presenta esta red de coautoría de artículos científicos, y profundizar en el entendimiento de la organización del PDI del Campus de la UPC en Terrassa.

¹ Excel™

²NodeXL (<https://www.smrfoundation.org/nodexl/>), programa con el cual trabajaremos para la creación y análisis de la red de coautoría de artículos científicos.

1.2 ALCANCE

- Creación y limpieza de datos (in., *data cleaning*) de la base de datos y creación de la lista de aristas "autor/a - autor/a" por año, en formato exportable al software de análisis de redes complejas.
- Proceso de análisis de la red de coautoría de los artículos científicos mediante medidas de centralidad.
- Determinación de la organización social del PDI del Campus de Terrassa mediante la interpretación de los resultados de las medidas de centralidad anteriormente calculadas y la literatura existente referente al análisis cuantitativo de redes sociales.

1.3 REQUERIMIENTOS

Dada la especificidad del TFG, los únicos requerimientos son dos bases de datos de dominio público:

- Datos referentes al PDI vinculado al Campus de Terrassa, obtenidas a través de Futur UPC³ y CTT.⁴
- Datos referentes a la producción científica del PDI vinculado al Campus de Terrassa en los años 2014 - 2018, obtenidas a través del Servicio de Bibliotecas de la UPC, en particular de la Biblioteca del Campus de Terrassa.⁵

³<https://futur.upc.edu/>

⁴<https://rdi.upc.edu/ca>

⁵<https://bibliotecnica.upc.edu/bct>

1.4 JUSTIFICACIÓN Y UTILIDAD

En vista de la cantidad de artículos, trabajos, tesis y proyectos que se presentan cada año por parte de todo el profesorado docente e investigador, se quiere analizar la organización social que hay detrás de toda esta actividad, por lo que con los resultados obtenidos, ya sea una organización jerárquica o colaborativa, se pueda mejorar la eficiencia y la eficacia de esta producción de artículos en el campus de Terrassa así como su puesta en valor, la cual prima, esencialmente, la cantidad de artículos producidos a cualquier coste, y prescinde de las condiciones laborales y relacionales de y en los grupos en los que se realiza este trabajo de investigación. La ventaja de este proceso será el análisis puramente objetivo respecto a los datos públicos que se encuentran en FUTUR (portal de la producción científica de los investigadores de la UPC) y la posterior difusión y comprensión de los resultados obtenido.

2. DESARROLLO

2.1 ANTECEDENTES

La web FUTUR de la UPC se ha utilizado en los últimos años para generar el mapa de colaboraciones entre PDI del Campus de la UPC en Terrassa a partir de artículos científicos.⁶En él, podemos ver el mapa con los ámbitos de conocimiento y la relación que existe entre ellos basada en la producción científica del personal investigador de la UPC. Pero este mapa dice poco al respecto de la calidad de la relación entre los diferentes actores y no permite esclarecer nada de la organización social que hay detrás de esta producción científica

Este TFG pretende ahondar un poco más en este sentido mediante el análisis de redes complejas aplicada a artículos científicos, proceso de análisis ya utilizado desde hace años para establecer la calidad de los vínculos profesionales entre científicos(Bettencourt et al., 2009; Chen et al., 2016; H. Small, 1973; H. G. Small, 1978).

2.2 PLANTEAMIENTO

Antes de proceder al desarrollo del trabajo, se plantean aquí las bases teóricas necesarias para el entendimiento de todo el documento. Hablamos de las redes complejas, y de uno de sus casos particulares, las redes sociales, las cuales son el objeto de este TFG. Se explicarán los principales modelos existentes para redes complejas, el modelo Erdős-Rényi para las redes aleatorias, el modelo Watts-Strogatz para redes de mundo pequeño, y el modelo Barabási-Albert para redes sin escala. En cuanto a su estructura, se hablará de redes colaborativas y jerárquicas. Hablaremos también de las principales métricas asociadas al análisis de la teoría de grafos, las medidas de centralidad, y por último de las hipótesis planteadas al inicio de éste, las cuales discutiremos finalmente para ver si se han cumplido o no.

⁶<https://upcommons.upc.edu/handle/2117/177320>

2.2.1 Redes Complejas

2.2.1.1 *Introducción a las redes complejas*

Las redes complejas son redes modeladas como grafos, es decir, un conjunto de vértices o nodos unidos por unos enlaces o aristas, que permiten representar relaciones entre los elementos de un conjunto, ya sean fenómenos físicos, biológicos o sociales, con el fin de llegar a modelos predictivos de estos fenómenos. Las redes complejas, a diferencia de los grafos, describen las interacciones presentes en un sistema complejo, mientras que un grafo no tiene por qué estar asociado a un sistema complejo. Estas redes poseen ciertas propiedades estadísticas y topológicas no triviales que no encontramos en redes simples, como por ejemplo su estructura jerárquica o comunitaria.⁷

Las redes complejas son objeto de estudio de la ciencia de redes, donde se incluye la teoría de grafos de las matemáticas, la mecánica estadística de la física, la minería de datos y la visualización de la información de ciencias de la computación, la estadística inferencial de la estadística y la estructura social de la sociología. Vemos entonces, que el estudio de estas redes complejas implica el uso de varias disciplinas, dependiendo del campo en el que se usen. Algunos ejemplos de redes complejas son:

- las redes informáticas (Figura 1),
- redes biológicas,
- redes de telecomunicaciones,
- redes de transporte,
- redes sociales.

⁷ Esta parte introductoria se basa mayoritariamente en la obra de referencia (Newman, 2010).

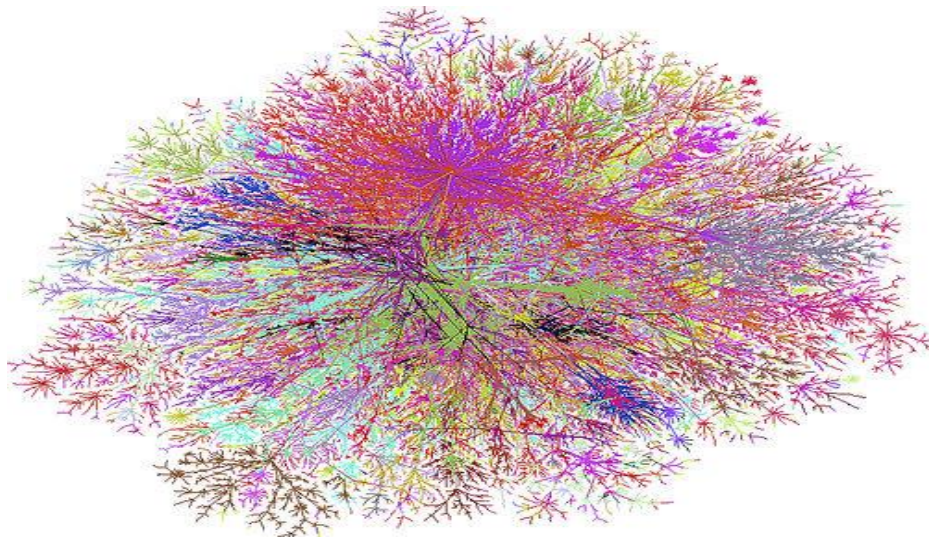


Figura 1. Representación gráfica de los ISP y nodos de Internet a la izquierda..

Para clarificar un poco más estos ejemplos, viendo la función que podrían desempeñar los nodos y las aristas en una red compleja, presentamos la siguiente tabla:

RED COMPLEJA	NODOS	ARISTAS
Internet	Routers	Líneas físicas
World Wide Web	Páginas web	Direcciones URL
Red de actores	Actores	Actuar en la misma película
Red de colaboraciones	Científicos	Coautores en una publicación
Red de interacción de proteínas	Proteínas	Interacción física
Red cerebral	Neuronas	Conexiones sinápticas
Red de contagios	Personas	Interacciones entre personas
Red eléctrica	Centrales eléctricas	Conexiones eléctricas
Red aeropuertos	Aeropuertos	Vuelos entre aeropuertos

Tabla 1. Ejemplos de redes complejas.

2.2.1.2 Contexto histórico

El artículo más antiguo sobre redes complejas fue el famoso problema de los siete puentes de Königsberg (actual Kaliningrado, capital del Oblast de Kaliningrado), propuesto por el matemático y físico suizo Leonhard Euler en 1736, el cual resolvió y puso las bases en cuanto a la teoría de grafos se refiere.

El problema en su formulación original decía: "Dado el mapa de Königsberg, con el río Pregel dividiendo el plano en cuatro regiones distintas, que están unidas a través de los siete puentes, ¿es posible dar un paseo comenzando desde cualquiera de estas regiones, pasando por todos los puentes, recorriendo sólo una vez cada uno, y regresando al mismo punto de partida?".

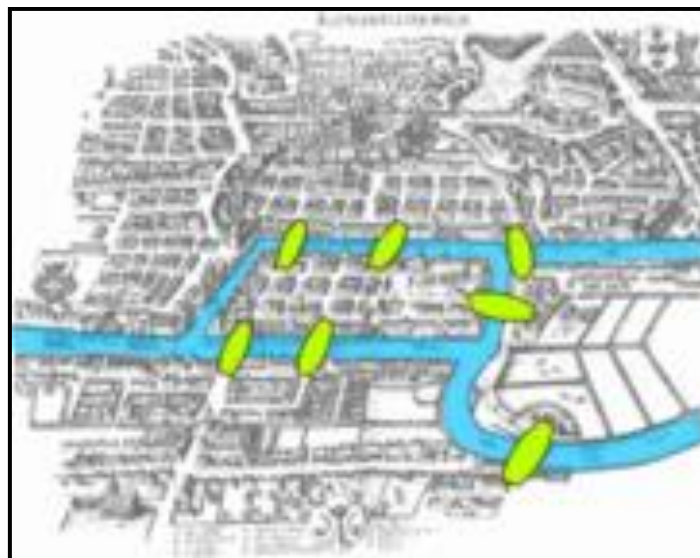


Figura 2. Los siete puentes de Königsberg.

Viendo esta representación de los puentes en la ciudad de Königsberg (Figura 2), Euler tuvo la idea de representar la ciudad como un grafo en el que las cuatro partes de la misma eran los vértices y los siete puentes eran las aristas (Figura 3), haciendo por primera vez uso de las propiedades estructurales de los grafos, quedando una figura como la que podemos ver a continuación:

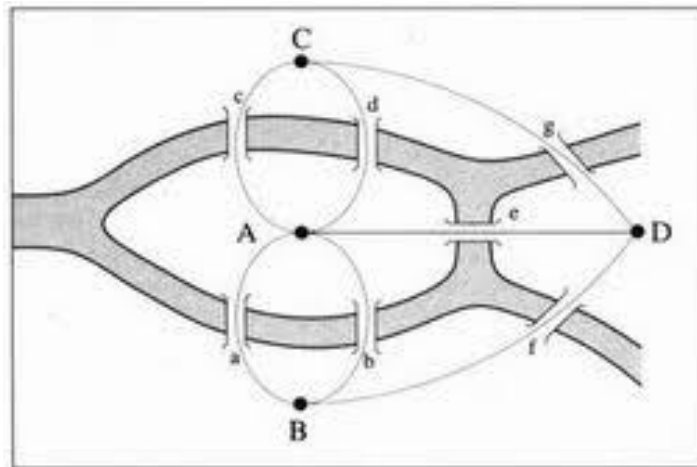


Figura 3. Grafo de los puentes de Königsberg.

La solución de Euler a este problema mediante esta representación fue la que dio origen a la teoría de grafos. De lo que se dio cuenta es que para recorrer el grafo como proponía, en un recorrido continuo sin pasar dos veces por el mismo puente, los vértices intermedios debían tener un número par de aristas, es decir, una de entrada y otra de salida (siendo este vértice de grado par). Únicamente los nodos de inicio y salida podían tener un número impar de enlaces, pero además se tenía que cumplir que el punto inicial debía ser igual al final, por lo que tendrían también un número par de líneas.

Teniendo esto en cuenta, lo único que debía hacerse era calcular el grado de cada vértice, es decir, el número de aristas conectadas al vértice. Teniendo en el problema cuatro vértices, donde dos de ellos son intermedios, si tenemos más de dos vértices de grado impar el recorrido no es posible. Lo que vemos en la (Figura 3) es que todos los vértices son de grado impar, y por lo tanto el recorrido no era posible.

De este problema, aparte de asentar las bases de la teoría de grafos, se caracterizaron formalmente los caminos eulerianos en la teoría de grafos, un camino que pasa por cada arista una y solo una vez, y más tarde, Carl Hierholzer en 1873, caracterizó el ciclo euleriano, un camino cerrado que recorre cada arista exactamente una vez.⁸

⁸https://es.wikipedia.org/wiki/Carl_Hierholzer

Actualmente, los campos de la teoría de redes que se están desarrollando son los siguientes:

- las redes de pequeño mundo gracias a la investigación de los matemáticos Duncan Watts y Steven Strogatz en el año 1998 en la revista *Nature*, del cual surge el modelo Watts y Strogatz (Watts & Strogatz, 1998),
- y la aparición de redes complejas invariantes de escala, en los cuales se propone un nuevo paradigma para el estudio de las redes complejas, con el modelo de Barabási–Albert, un algoritmo empleado para generar redes aleatorias complejas libres de escala empleando una regla o mecanismo denominado conexión preferencial (Barabási & Albert, 1999).

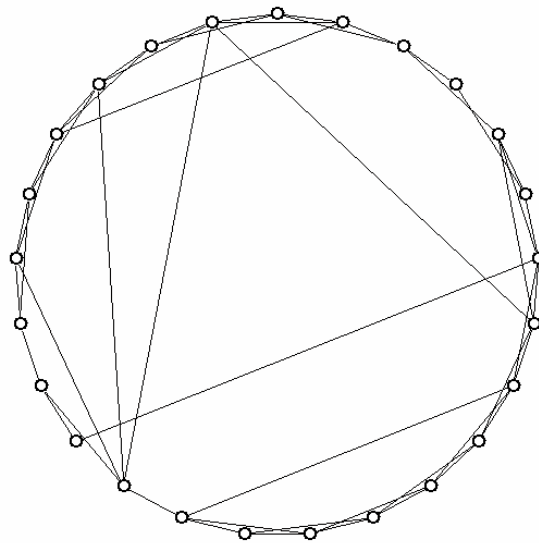


Figura 4. Red de mundo pequeño siguiendo el modelo Watts y Strogatz.

2.2.1.3 Marco Teórico

En cuanto al marco teórico de las redes complejas, debemos hablar de grafos y de la teoría de grafos, ya que es cómo podemos representar y analizar, respectivamente, estas redes formalmente. La teoría de redes se usa para hablar de grafos con un número elevado de vértices y aristas. Por ello al definir una red, lo haremos de igual manera que un grafo.

Definimos un grafo $G = (N, L)$ constituido por dos conjuntos $N = \{n_1, \dots, n_n\}$ cuyos elementos son los nodos o vértices del grafo y $L = \{l_1, \dots, l_n\}$ cuyos elementos son las uniones o aristas. Las aristas vienen definidas por los órdenes de los nodos que unen, es decir, la unión entre los vértices n_i y n_j se denota por $l_k = (i, j) = (n_i, n_j) = l_{ij}$.

En cuanto a vocabulario, si hay una unión o arista entre dos nodos, estos los nodos se llaman vecinos o adyacentes. Cuando dos aristas unen dos mismos vértices hablamos de aristas paralelas. Y cuando existen aristas que tienen en común un nodo o vértice hablamos de aristas adyacentes. Destacar también, que si a cada arista se le asigna un peso o valor numérico que mide la intensidad de la unión, nos referimos a grafos pesados. En caso contrario es un grafo no pesado.

Podemos decir que la representación más habitual de una red compleja es mediante su matriz de adyacencia A (Tabla 2). Esta matriz A es una matriz cuadrada de dimensión $N \times N$, donde las filas y las columnas hacen referencia a los vértices para almacenar en cada casilla la longitud entre cada par de vértices del grafo.

La celda A_{ij} almacena la longitud entre el vértice i y el vértice j . Por lo tanto, la celda A_{ij} de la red representara la existencia o no de enlace entre los vértices i y j . Si $A_{ij} = 0$, los nodos ij no estarán conectados, mientras que si es $A_{ij} = 1$, sí que lo estarán.

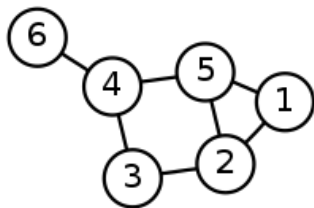


Figura 5. Ejemplo grafo no dirigido

A	1	2	3	4	5	6
1	0	1	0	0	1	0
2	1	0	1	0	1	0
3	0	1	0	1	0	0
4	0	0	1	0	1	1
5	1	1	0	1	0	0
6	0	0	0	1	0	0

Tabla 2. Matriz de adyacencia del grafo no dirigido

Al hablar de grafos, podemos diferenciar también entre dos tipos: grafos dirigidos y no dirigidos. En el caso de los grafos no dirigidos, el orden de los índices en las uniones es el mismo. Es decir, la existencia de un enlace entre el vértice i y el vértice j , implica que exista también entre el vértice j y el vértice i , por lo que decimos que $A_{ij} = A_{ji} = 1$, siendo la matriz A simétrica. En el caso de los grafos dirigidos, se tiene que el orden de los índices en las uniones es importante de forma que $l_{ij} \neq l_{ji}$. Cabe destacar que si la red no es pesada, esto es, si los enlaces no tienen una "cantidad" definida (como lo sería una distancia en una red espacial), todos los enlaces entre nodos se consideran de igual valor siendo $A_{ij} = 1$.

2.2.2 Redes Sociales

Como caso particular de las redes complejas, hablamos a continuación de las redes sociales. Como humanos, hemos nacido y vivimos en un mundo de conexiones. Las redes sociales se crean cuando interactuamos con el resto de personas, ya sea por nuestro lenguaje, por conflictos, por comercio o por colaboraciones. La teoría de redes sociales es un campo que lleva desarrollándose durante los últimos 80 años, pero ha sido gracias a las nuevas tecnologías computacionales y la evolución de las aplicaciones en internet que han ayudado notablemente a este desarrollo. Mediante el análisis de redes, podemos visualizar grupos de relaciones complejas mediante mapas o grafos, y calcular ciertas métricas como la forma, tamaño o densidad de la red, así como la posición de cada elemento dentro de ella, y cómo influye al resto del grupo, pudiendo ver si está centrada en la red o en la periferia de la estructura. De esta manera, el análisis de redes sociales ayuda a encontrar patrones que pueden ayudar al entendimiento de un grupo social, así como a predecir posibles futuros comportamientos.

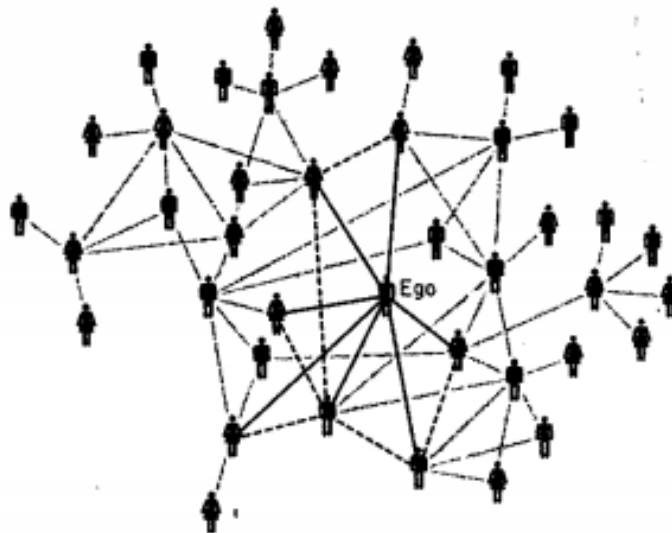
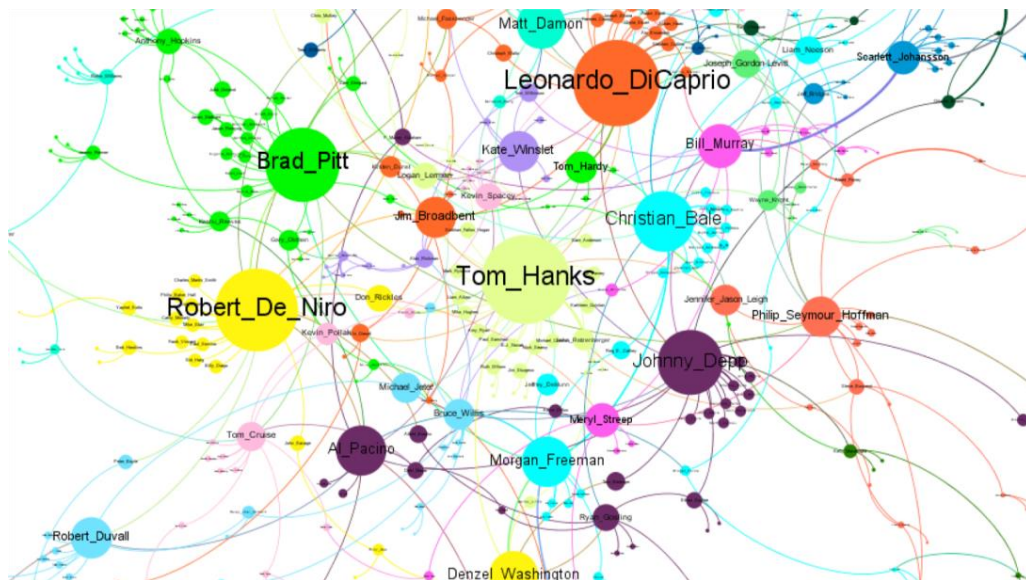


Figura 6. Ejemplo de una red social en la que se muestra cómo se tiende a formar triángulos.

El gran cambio en la manera de analizar toda esta información, implica centrar el foco del análisis de las redes sociales entre la gente, y no en la gente. Es decir, centrarnos en las relaciones y las conexiones que hacen que ese grupo de personas esté unido, y no en sus características intrínsecas o sus atributos como puede ser la edad o el género. Esta nueva visión de observar las redes frente a la tradicional, defiende que diferentes personas en circunstancias y posiciones sociales similares actúan normalmente de manera parecida. Por ello, el análisis de redes defiende que el éxito o el fracaso de muchas organizaciones es el resultado de una buena o mala organización en la estructura de sus relaciones (Girvan & Newman, 2001).



En la Figura 7 podemos ver un ejemplo de una red de actores, en la cual el tamaño de los nodos es proporcional a su centralidad de vector propio, o lo que es lo mismo, a sus influencias dentro de la red. Es decir, aquí se valora el número de conexiones que tiene, y si esas conexiones a su vez están bien conectadas o no. Esta, es una de las medidas de centralidad que pueden usarse para analizar las redes sociales. A continuación, definiremos las principales medidas de centralidad usadas en el análisis de redes.

⁹<http://ars-uns.blogspot.com/2018/03/redes-de-actores-en-la-base-de-datos.html>

2.2.3 Medidas de centralidad

La centralidad, en teoría de grafos, es un concepto introducido en 1950 por Alexander Bavelas, el cual se refiere a una posible medida de un vértice o nodo en dicho grafo, determinando así su importancia relativa dentro de la red (Bavelas, 1950). La centralidad mide según el criterio, la contribución o impacto de un nodo según su ubicación en la red. Es decir, hablamos de un atributo estructural, y no de un atributo intrínseco como podría ser la temperatura. Determinar la centralidad de un nodo, puede ayudarnos, por ejemplo, a conocer la importancia de una carretera dentro de la red urbana, porque colapsan los ecosistemas o sociedades, o el impacto de ciertas personas dentro de una red social.

Existen dos tipos de medidas de centralidad: las medidas radiales (in., *radial measures*) y mediales (in., *medial measures*). Las medidas radiales son las que toman como punto de referencia un nodo que inicia o termina recorridos por la red. Estas pueden clasificarse a su vez como medidas de volumen, midiendo el número de recorridos limitados a dicha longitud prefijada, o como medidas de longitud, midiendo la longitud de los recorridos necesarios para alcanzar un volumen prefijado. Por otro lado, las medidas mediales toman como referencia los recorridos que pasan a través de un nodo.

Existen cuatro medidas que son generalmente usadas en el análisis de redes, además de otros conceptos que también comentamos a continuación:

Centralidad de grado

La centralidad de grado (in., *degree centrality*) nos indica el número de enlaces o conexiones conectados al vértice, es decir, el número de aristas pertenecientes a ese nodo. Esta medida nos da una idea de la conectividad de un nodo. Formalmente, definimos la centralidad de grado, dado un grafo $G := (V, E)$, siendo V su conjunto de vértices, y E su conjunto de aristas, para cada nodo $v \in V$:

$$C_{DEG}(v) = \text{grado}(v) \quad (1)$$

Esta definición es en el caso de tener la matriz de adyacencia del grafo, donde cada posición a_{ij} asume el valor 1, si existe la arista (i,j) y el valor 0.

En caso de no existir, entonces la centralidad de grado de cada nodo j se puede definir como:

$$C_{DEG}(j) = \sum_i a_{ij} \quad (2)$$

Para grafos dirigidos, podemos distinguir dos medidas de grado diferentes, el grado de entrada y el de salida. El grado de entrada (in., *in-degree*) es el número de conexiones que apuntan a ese vértice. El grado de salida (in., *out-degree*) es el número de conexiones que se originan en el vértice y apuntan a otros vértices.

Es decir, en este tipo de grafos, las aristas tienen un sentido. Un ejemplo de grafo dirigido puede ser la red de aguas de una ciudad, ya que, en este caso, cada tubería solo admite que el agua se recorra en un único sentido.

En nuestro caso, trataremos con un grafo no dirigido, ya que al tratarse de colaboraciones entre el profesorado de investigación, no tiene sentido preguntarse el sentido de la colaboración, sino que más bien estudiaremos la estructura de estas colaboraciones.

Centralidad de cercanía

La cercanía (in., *closeness centrality*) es la más conocida y utilizada de las medidas radiales de longitud. Fue definida por el matemático Murray Beauchamp en 1965 y luego popularizada por Freeman en 1979 (Freeman et al., 1991). Se basa en el cálculo de la distancia media entre un vértice y el resto de los vértices de la red.

Es decir, calcula las rutas más cortas entre todos los nodos, y luego asigna una puntuación a cada nodo, basándose en la suma de las distancias más cortas.

Formalmente podemos definir la cercanía como:

$$C_{CL}(i) = \frac{1}{\sum_{j=1}^n (S)_{ij}} \quad (3)$$

Donde S es la matriz de distancias de la red, cuyos elementos (i, j) corresponden a la distancia más corta desde el nodo i hasta el j .

Cuanto menor sea el valor de cercanía, significa que ese nodo o vértice está directamente conectado o a una pequeña distancia del resto de nodos en la red. Es decir, que a menor valor en este parámetro, más importancia tendrá el nodo en la red, ya que su distancia media con respecto al resto de vértices es más pequeña y por lo tanto, se encuentra más cerca a todos ellos que el resto. Gracias a esta medida, en una red de flujo, se puede interpretar como la rapidez en la que una información se puede propagar desde un nodo al resto, o en una red social, por ejemplo, a través de cuantas personas podemos llegar a otra en esa red.

Centralidad de intermediación

La centralidad de intermediación (in., *betweenness centrality*) mide el número de veces que un vértice particular está situado en el camino más corto (o lo que es lo mismo entre la distancia geodésica) entre otros dos vértices de la red. Es decir, mide la frecuencia en la que un nodo sirve de puente dentro de la ruta más corta entre otros dos nodos de la red. Si este puente resulta unir muchos vértices de la red por el camino más corto, este parámetro aumentará, y nos indicará una mayor importancia de este nodo en la red, ya que si este puente o vértice desapareciera, la unión del resto de vértices debería hacerse por otros caminos más largos.

Formalmente, podemos definir la intermediación de un nodo i en la red como:

$$C_{BTW}(i) = \sum_{j,k} \frac{n_{jik}}{n_{ji}} \quad (4)$$

Donde n_{ji} es el número de caminos más cortos desde el nodo j hasta el nodo k , n_{jik} es el número de caminos más cortos desde j hasta el nodo k , que pasan a través del nodo i .

Centralidad de vector propio

La centralidad del vector propio o (in., *eigenvector centrality*) fue propuesta por Phillip Bonacich en 1972 y corresponde al principal vector propio de la matriz de adyacencia del grafo analizado (Bonacich & Lu, 2012). Este parámetro mide la influencia o conectividad de un nodo dentro de la red. Es decir, un nodo que tenga pocas conexiones, puede poseer un valor alto de centralidad del vector propio, si estas pocas conexiones que tiene, a su vez están muy bien conectadas. Por lo tanto, nodos que tengan un valor alto en esta medida de centralidad, serán buenos candidatos por ejemplo para difundir información o propagar enfermedades en una red de contagios.

Una variante de esta medida de centralidad, es el PageRank. Este algoritmo fue creado por Google en 1999 y es utilizado en su motor de búsqueda (Mitchell, 2009). La diferencia es que aparte de medir la cantidad de enlaces que posee una web, también tiene en cuenta la dirección y el peso del enlace, por lo que los enlaces solo pueden transmitir influencia en una dirección y transmitir diferentes cantidades de influencia.

Coeficiente de agrupamiento

Aún sin ser una medida de centralidad, comentamos brevemente este parámetro ya que más tarde será usado para el análisis de nuestra red. El coeficiente de agrupamiento (in., *clustering coefficient*) es una medida del grado en el que los nodos de una red tienden a agruparse entre ellos, es decir, cuantifica qué tanto está de agrupado o interconectado un vértice con sus vecinos. Dado un grafo $G(N, L)$ no dirigido se define el coeficiente de agrupamiento de un nodo i de dicho grafo y se denota como C_i al cociente:

$$C_i = \frac{\text{número de pares de nodos vecinos de } i \text{ conectados por aristas}}{\text{número total de pares de nodos vecinos de } i \text{ que podrían existir}} \quad (5)$$

Un ejemplo de cálculo para este coeficiente lo podemos ver en la siguiente imagen:

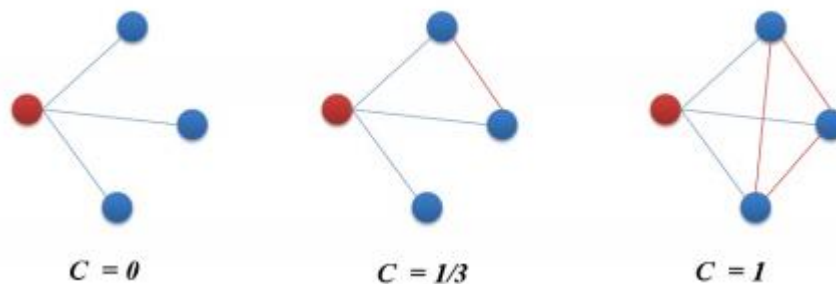


Figura 8. Ejemplo de cálculo del coeficiente de agrupamiento. Fuente¹⁰

Podemos ver como si el vértice de color rojo está agrupado como un clique, es decir, como un subgrafo en el que cada nodo está conectado a todos los demás nodos del subgrafo, el coeficiente de agrupamiento será máximo, de valor 1, mientras que, si ninguno de los vecinos del vértice rojo están conectados entre ellos, su valor será 0.

¹⁰http://www.ugr.es/~jtorres/Tema_2_redes_complejas.pdf

2.2.4 Métricas globales

En cuanto a las métricas globales de red que se deben conocer antes de proceder al análisis de la red de coautorías, destacamos las siguientes:

Longitud de camino

La longitud de camino (in., *path length*) (L_i) es el número medio de aristas que se deben cruzar para ir de un nodo i al resto de nodos de la red:

$$L_i = \sum_{j \in N} \left(\frac{1}{n-1} * \sum_{j \in N, j \neq i} d_{ij} \right) \quad (6)$$

Donde N es el número total de nodos de la red, n es el número de nodos implicados y d_{ij} es el shortestpathlength entre el nodo i y j . Hay que tener en cuenta para esta métrica la definición de camino (in., *path*). Un path es una secuencia de nodos unidos por aristas en la que nunca se pasa por un mismo nodo más de una vez.

Cuando hablamos de "characteristic path length o average path length" nos referimos a una medida global de la red, es decir, que solo hay un valor para la red entera. La longitud promedio de caminos o distancia promedio consiste en la media del path length de cada nodo de la red, es decir, el promedio de las distancias más pequeñas entre dos nodos cualesquiera de la red:

$$L = \frac{1}{N} \sum_{i \in N} L_i \quad (7)$$

Diàmetro

El diàmetro (D) de una red conectada se define como el máximo de las distancias más cortas entre cualquiera dos nodos de la red, es decir, al camino geodésico más largo entre dos vértices cualquiera:

$$D = \max d(n_i, n_j) \quad \text{para } i, j \neq i \quad (8)$$

Entendemos por distancia geodésica el mínimo número de nodos que hay que recorrer para unir dos nodos cualesquiera. Es decir, el camino más corto que existe entre dos vértices, entendiéndose que puede haber más de uno.

2.2.5 Tipos de redes

2.2.5.1 Redes aleatorias

Las redes aleatorias se basan en grafos generados por algún tipo de proceso aleatorio. El modelo más aplicado en la generación de grafos aleatorios, dentro de la teoría de grafos, fue propuesto por los matemáticos Paul Erdős y Alfréd Rényi, y es conocido habitualmente como el modelo ER por las iniciales de sus autores (Newman et al., 2002). La característica principal de estas redes radica en que un nuevo nodo posee independencia estadística con el resto de nodos de la red, es decir, que cada nodo que aparece se enlaza con igual probabilidad con el resto de nodos en la red.

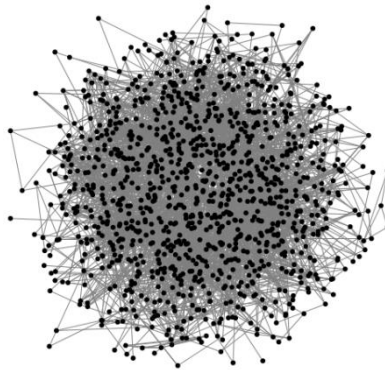


Figura 9. Red aleatoria. Elaboración propia en NodeXL.

Es decir, teniendo una red $G = (N, L)$ con N nodos y L aristas, diremos que G es aleatoria si las aristas que conectan dos vértices cualesquiera se generan aleatoriamente.

Podemos hablar de dos modelos de redes aleatorias:

1. Redes tipo $G = (N, L)$ con N nodos y L aristas, donde los nodos se conectan aleatoriamente entre sí, utilizando estas L aristas.
2. Redes tipo $G = (N, p)$ donde las aristas se generan siguiendo una distribución de probabilidad binomial que nos indica si dos vértices están conectados o no. Este es el caso del modelo Erdős-Rényi, y p nos indica la probabilidad de que dos nodos de G estén conectados. Cabe destacar, que las aplicaciones del modelo Erdős-Rényi son bastante limitadas, ya que son pocas las redes reales que se comportan como este modelo. La distribución de grado de este modelo es de tipo Poisson.

2.2.5.2 Redes de mundo pequeño

Las redes de pequeño mundo (in., *small world*) se basan en la idea de los seis grados de separación y son bastante frecuentes en nuestro entorno, tanto en redes sociales, físicas o biológicas. Esta idea defiende que cualquier persona en la Tierra está conectada a otra persona del planeta a través de una cadena de conocidos que no tiene más de cinco intermediarios. Esta idea fue introducida por primera vez en 1930 por el escritor húngaro Frigyes Karinthy en su cuento "Chains". Posteriormente probada y reconocida por el psicólogo Stanley Milgram en 1963 mediante el "experimento de pequeño mundo". Finalmente fue impulsada en los últimos años por Watts y Strogatz, del cual nace el modelo de redes de mundo pequeño Watts y Strogatz, introducido en el año 1998 en su artículo en la revista Nature (Watts & Strogatz, 1998).

El experimento de Milgram es esencial en el entendimiento del fundamento de las redes de pequeño mundo. Consistía en escoger de manera aleatoria a dos personas de los Estados Unidos, típicamente una en Boston y otra en Omaha, ya que representaban una gran distancia tanto geográficamente como socialmente. A estos individuos se les enviaba información detallando el propósito del estudio además de información básica acerca del destinatario. Se añadía también una lista en la cual todos los participantes debían apuntar sus nombres, conjuntamente con unas tarjetas de respuesta pre-dirigidas a Harvard, que aseguraban un rastreo del progreso durante la cadena al destinatario, especialmente para aquellas cadenas que quedaban interrumpidas y donde el destinatario no llegaba a recibir la carta. La idea era que el primer individuo entregara una carta al segundo, mediante una cadena de conocidos intermediarios, pensando en quién de sus conocidos o "vecinos" tenía más posibilidades de conocer directamente al destinatario final. En el caso de que el primer individuo conociese al destinatario final, se le pedía que le enviara la carta directamente.

Se llegaron a enviar 296 paquetes. De éstos, sólo 64 casos llegaron al final de la cadena. Y de estos 64 casos, la longitud promedio de la cadena de conexiones fue de entre 5'5 y 6 personas. La conclusión a la que se llegó fue que se podía conectar a dos personas cualesquiera con tan solo seis saltos de media, fenómeno al que se denominó "seis grados de separación" y del cual sale su expresión coloquial: "el mundo es un pañuelo".



Figura 10. Stanley Milgram. Fuente ¹¹

Basándonos en la idea de los seis grados de separación, podemos definir a las redes de mundo pequeño como un tipo de grafo en el que la mayoría de nodos no son vecinos entre sí, y sin embargo, podemos llegar desde un nodo cualquiera del grafo a otro, mediante una cadena de nodos intermediarios relativamente bajo.

Esto significa que las redes de pequeño mundo presentan una estructura intermedia entre las redes ordenadas y las redes aleatorias, como podemos ver en la siguiente figura, caracterizadas por longitudes de camino de características similares a las de las redes aleatorias, pero coeficientes de agrupamiento similares al de las redes de tipo malla:

¹¹<https://www.buscabiografias.com/biografia/verDetalle/9921/Stanley%20Milgram>

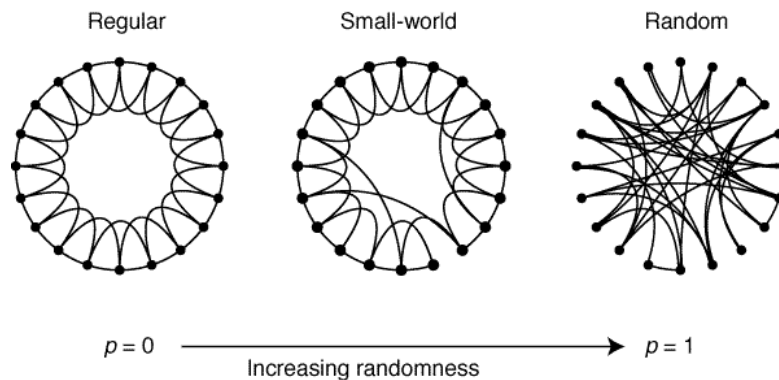


Figura 11. Dinàmiques de una red de mundo pequeño. Fuente¹²

Esta idea fue descubierta por los matemáticos Duncan Watts y Steven Strogatz en el año 1998. Siguiendo con la misma idea de encontrar las redes de pequeño mundo entre las ordenadas y las aleatorias, tenemos el siguiente gráfico dividido en estas tres etapas, desde las redes ordenadas a la izquierda, hasta las aleatorias a la derecha, donde se representan el coeficiente de agrupación (in., *clustering coefficient*) (C) y la longitud de camino característica (in., *path length*) (L).

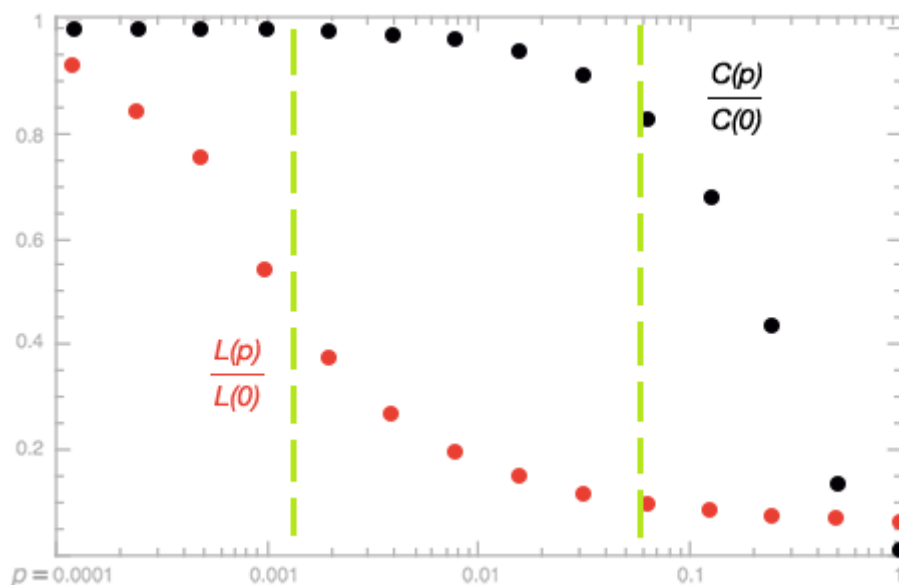


Figura 12. Gráfico log-lineal de distancia media y coeficiente de agrupamiento en función de la probabilidad. Fuente: Adaptación del trabajo de Watts y Strogatz "Collective dynamics of 'small-world' networks" en Nature 1998.

¹²<https://blogs.kent.ac.uk/strongcorrelations/2018/06/28/the-small-world-effect/>

En la zona intermedia, encontramos el comportamiento de estos parámetros cuando se trata de redes de pequeño mundo. Como vemos, el coeficiente de agrupación se mantiene con unos valores altos, mientras que la longitud de camino característica, entendida como la distancia mínima promedio entre nodos elegidos al azar, es pequeña.

Partiendo de estos resultados, Watts y Strogatz definieron su modelo a partir de dos conceptos:

- El primero, una distancia mínima promedio entre nodos pequeña. Es decir, que dentro de la red, cualquier nodo se comunique a otro mediante un número de nodos intermedio relativamente pequeño, como indica el fenómeno pequeño mundo.
- El segundo, un coeficiente de agrupamiento elevado. El coeficiente de agrupamiento es una medida del grado en que los nodos en un gráfico tienden a agruparse en una red. Este valor crecerá cuantas más conexiones haya entre los vecinos de un nodo dado. Es decir, un valor alto en este parámetro, indica que aunque dos nodos no estén directamente conectados entre sí, existe una gran probabilidad de que conecten mediante la intervención de otros nodos.

En cuanto a la distribución de grado, las redes regidas por el modelo Watts-Strogatz de mundo pequeño (global y local) siguen una distribución de Poisson al igual que las redes aleatorias, y está normalmente distribuida alrededor del grado promedio $\langle k \rangle$.

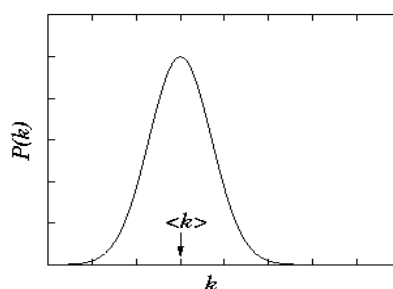


Figura 13. Distribución de grado alrededor del grado promedio $\langle k \rangle$.

Esta distribución de grado tipo Poisson, que da lugar a una distribución de forma exponencial de forma $P(k) \sim e^{-k}$, y en forma de campana de Gauss, significa que la mayoría de nodos dentro de la campana tienen el mismo número de enlaces, mientras que en la base de la campana se refleja la inexistencia de nodos altamente conectados.

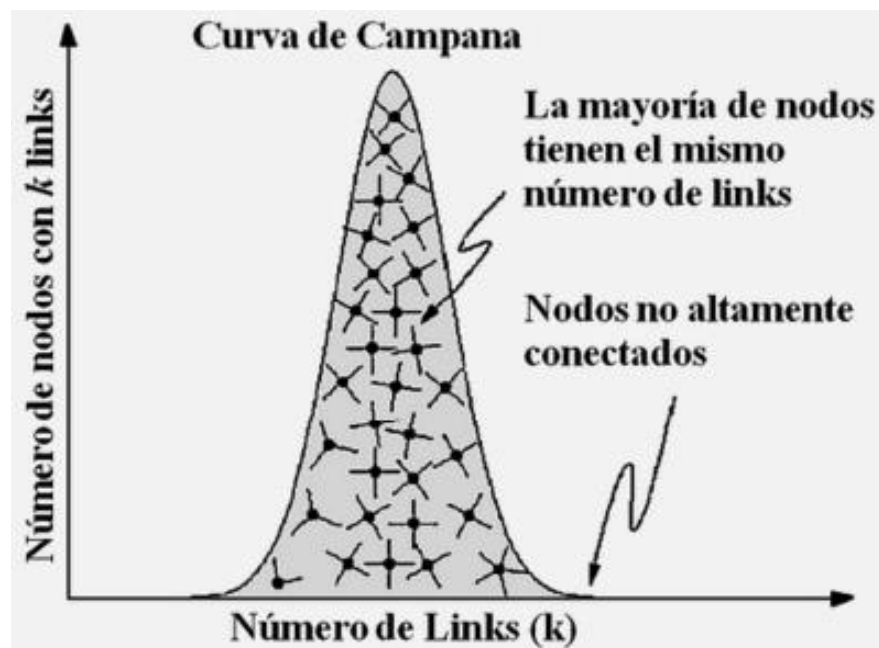


Figura 14. Distribución Poisson. Fuente: Barabási, 2002:71

Las propiedades de este tipo de redes presentan alta conectividad en vértices cercanos y una distancia pequeña entre vértices, **lo que permite una alta eficiencia en el flujo de información dentro de éstas**. Por ello, las redes que cumplen $0 < p < 1$ tienen efectos de mundo pequeño.

2.2.5.3 Redes de libre escala

Las redes de libre escala se basan en la idea de que algunos nodos (llamados hubs) están altamente conectados, es decir, tienen un gran número de enlaces a otros nodos, mientras que el resto de nodos de la red poseen un grado de conexión bastante bajo. Como ejemplos de este tipo de redes tenemos las redes de distribución eléctrica o los aeropuertos, en las sociales las redes de amistad o de crimen organizado, y en biología las redes de neuronas en los organismos dotados de sistema nervioso.

Barabási junto a otros investigadores, descubrieron este fenómeno al hacer un mapa de la "Word Wide Web" en 1999 (Barabási & Albert, 1999). Al hacerlo, esperaban un modelo tipo Erdős-Rényi, sin embargo, se encontraron con que solo unas pocas páginas estaban altamente conectadas, y que el resto (más del 80%) estaban como máximo enlazadas a otras cuatro. Nació aquí el término "hub" y se dejó de tratar todas las redes con el modelo habitual de Erdős-Rényi.

El modelo de Barabási-Albert (o modelo BA) es un algoritmo empleado para generar redes aleatorias complejas libres de escala empleando una regla o mecanismo denominado conexión preferencial. Albert-László Barabási y Réka Albert, fueron los re-descubridores del proceso de conexión preferencial en 1999 y lo aplicaron a la distribución del número de enlaces que tiene cada sitio de Internet, aunque sus orígenes se pueden atribuir a Yule en 1925 o a Herbert Simon en el año 1955.

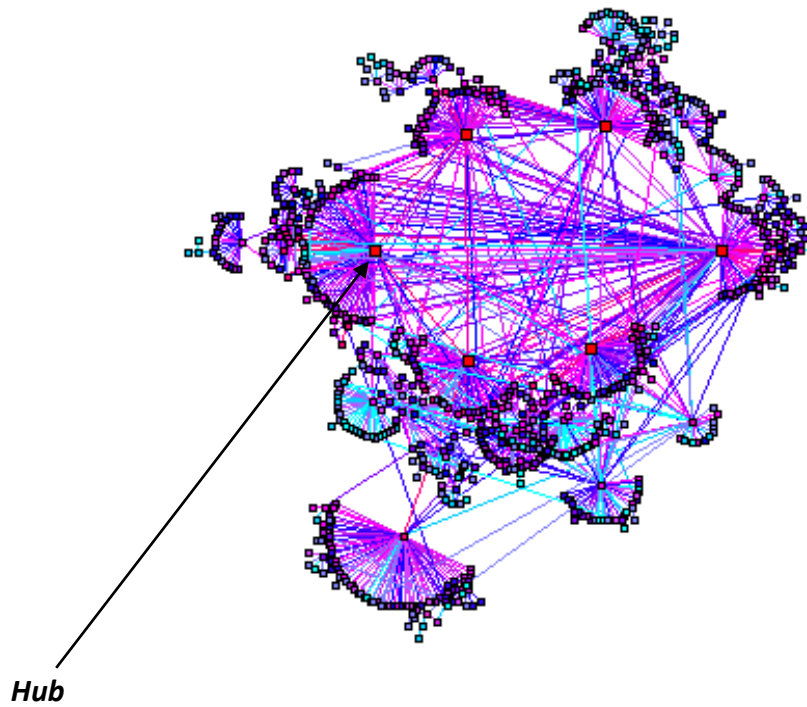


Figura 15. Mapa de la World Wide Web por Barabási y Albert. Fuente: Nature, 401 130 (1999)

Este tipo de redes poseen una distribución de grado que sigue la ley de potencias de forma $P(k) \sim k^{-\gamma}$, como podemos ver a continuación:

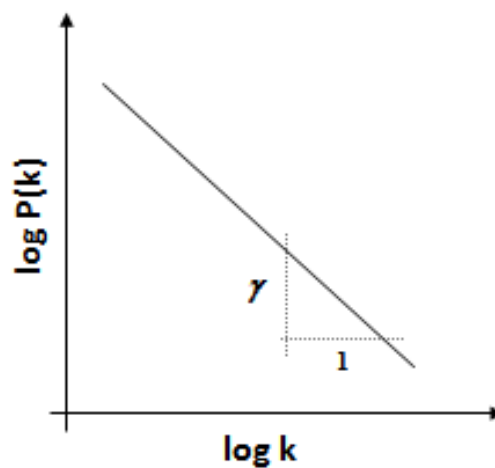


Figura 16. Distribución potencial. Elaboración propia.

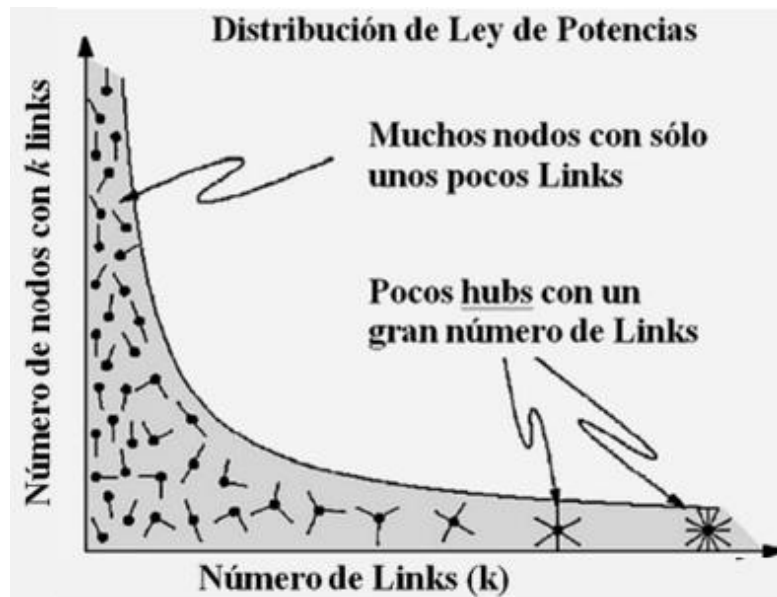


Figura 17. Distribución siguiendo la Ley de potencia. Fuente: Barabási, 2002:71

Esto implica como podemos observar, que al igual que Barabási y sus investigadores vieron al mapear la World Wide Web, la mayoría de nodos de la red están poco enlazados (a la izquierda del gráfico) y unos pocos nodos (in., *hubs*) están altamente conectados.

De estos resultados se sacan algunas propiedades de las redes de libre escala:

1. Si se desconectan el 80% de los nodos con menos conexiones, el resto de la red seguiría conectada.
2. Si se elimina entre el 5 y el 10% de los hubs, el sistema podría desmoronarse.
3. Jerarquía de grupos o "clusters", permitiendo la aparición de picos.
4. Cualquier nodo de la red está conectado a otro cualquiera con pocos grados de separación.
5. Existencia de nodos cuyo valor de conectividad supera varias veces el número promedio.
6. El exponente $-\gamma$ en la ley de potencias, $k^{-\gamma}$ suele estar entre 2 y 3.



Es importante destacar que en las redes de escala libre, el número de nodos no es fijo a lo largo del tiempo, aparecen y desaparecen nuevos nodos continuamente, aunque se sigue una tendencia de crecimiento. En este proceso de evolución de la red, existe un problema y es que los nodos con alta conectividad tienden a hacerse más ricos en conexiones a medida que se evoluciona, es decir, un nuevo nodo tendrá más posibilidades de unirse a uno que ya tiene un número elevado de conexiones. Esto lo denominó el propio Barbási como: "Rich get richer" (el rico se hace más rico).

2.2.6 Redes colaborativas

Las redes colaborativas (in., *collaborative networks*) fueron definidas por Peter Gloor como un equipo de personas autónomas, con una visión colectiva, que colaboran en la consecución de un objetivo común al grupo a través del intercambio de ideas, información y trabajo, y cuyas interacciones son compatibles con las redes informáticas¹³. Se puede visualizar esta organización como un solo cerebro, una entidad única integrada por millones de células generadoras de ideas, conocido como “El Cerebro Global”. Tienen como base, la transparencia interna en las organizaciones y la comunicación directa entre todos los miembros de la red. La disciplina de las redes colaborativas se centra en la estructura, el comportamiento y la dinámica evolutiva de las redes de entidades autónomas que colaboran para lograr de una mejor manera objetivos comunes o compatibles.

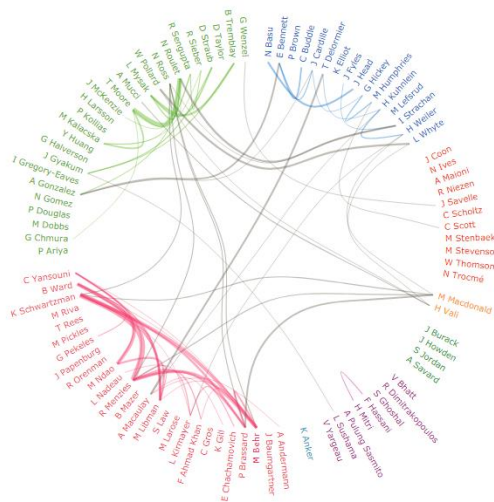


Figura 18. Red colaborativa de la Universidad McGill North.

Como elementos fundamentales, en este tipo de redes se permite a los usuarios buscar expertos o contenido. Son dirigidas por los mismos usuarios y se permite que los datos empresariales se integren en el sistema de manera dinámica. Se monitorizan los proyectos y su éxito de éstos mediante herramientas y paneles de control. Y por último el gobierno debe tener un acceso controlado a todo el contenido de la red.

¹³https://en.wikipedia.org/wiki/Collaborative_innovation_network

2.2.7 Redes jerárquicas

Las redes jerárquicas se basan en algoritmos iterativos para crear redes que pueden reproducir las propiedades únicas de la topología libre de escala y el alto agrupamiento de los nodos al mismo tiempo. En general, se derivan de forma iterativa al replicar el clúster inicial de la red de acuerdo con una determinada regla.

Son varias las redes en la vida real que presentan esta estructura, como redes metabólicas, la red de interacción de proteínas, la World Wide Web o algunas redes sociales. El modelo de red jerárquica, perteneciente a los modelos sin escala, tiene como propiedad principal el hecho de tener proporcionalmente más centros entre los nodos que por generación aleatoria. En los modelos jerárquicos, se espera que los nodos con más enlaces tengan un coeficiente de agrupación más bajo, a diferencia de otros modelos similares como el Barabási-Albert, o el modelo Watts-Strogatz, que predecirían un coeficiente de agrupamiento constante como función del grado del nodo. En los modelos jerárquicos, además, no existe una relación entre el tamaño de la red y su coeficiente de agrupación promedio.

La distribución de grado del modelo de red jerárquica sigue la ley de potencia (power law), lo que significa que un nodo seleccionado al azar en la red tiene k bordes con una probabilidad:

$$P(k) \sim ck^{-\alpha} \quad (9)$$

donde c es una constante y α es el exponente de grado, que se puede calcular como:

$$\alpha = 1 + \frac{\ln M}{\ln(M-1)} \quad (10)$$

donde M es el factor de replicación del modelo o grado de expansión.



El coeficiente de agrupación se puede expresar en función del grado de la siguiente manera:

$$C(k) \sim k^{-\beta} \quad \text{donde } \beta = 1 \quad (11)$$

2.3 HIPÓTESIS

Planteamos aquí algunas de las posibles hipótesis antes de empezar con el análisis de la red, según sea una red jerárquica o colaborativa, y que finalmente discutiremos una vez tengamos los resultados del análisis.

2.3.1 Hipótesis de jerarquía

Si fuera una red jerárquica, para cada componente conexa, y a medida que pasa el tiempo:

- Existen aristas con un peso mayor de lo que se podría esperar. Esto sería posible, por ejemplo, con una colaboración más intensa entre un jefe (de departamento o grupo de investigación) y un subordinado (i.e., alumno de doctorado, postdoc, etc.).
- Destaca un nodo (i.e., jefe de departamento o PDI con antigüedad notable) que:
 - Aumenta su grado de manera no lineal (estructura local en forma de estrella) y en número absoluto.
 - Aumenta su "clustering coefficient" de manera no lineal.
- El nodo (subordinado) repite el mismo patrón que el nodo (jefe de departamento) pero con un cierto retraso "delay".

2.3.2 Hipótesis de colaboración

- A medida que pasa el tiempo, el número de componentes conexas disminuye para cada año acumulado.
- Algunos nodos aumentan su "betweenness".
- El grado y el "clustering coefficient" de los nodos colaborativos aumentan de manera lineal con el tiempo, pero con valores absolutos que pueden ser órdenes de magnitud por debajo del caso jerárquico.
- El valor de cercanía (i.e., *closeness*) es muy bajo, con tendencia a disminuir.

3. DESARROLLO DE LA RED DE COAUTORIAS

3.1 CREACIÓN DE LA BASE DE DATOS EN FORMATO EXCEL

Primeramente, obtuvimos el listado en formato .txt de los artículos indexados desde 2014 hasta 2018.

```
articles indexats CT 2014-2018.bib.txt
@article{Raush2017,
author = {Raush, Gustavo and Gamez-Montero, Pedro J and Castilla, Roberto and Codina, Esteve},
doi = {10.1016/j.flowmeasinst.2016.10.014},
issn = {0955-5986},
journal = {FLOW MEASUREMENT AND INSTRUMENTATION},|
keywords = {IAFARG,LABSON,LMIT-CT},
mendeley-tags = {IAFARG,LABSON,LMIT-CT},
month = {jun},
pages = {13--22},
title = {{Experimental study on the impulsion port of a trochoidal wheeled pump}},
volume = {55},
year = {2017}
}
@article{Riba2018,
author = {Riba, Jordi-Roger and Moreno-Eguilaz, Manuel and Bogarra, Santiago and Garcia, Antoni},
doi = {10.3390/electronics7120393},
issn = {2079-9292},
journal = {ELECTRONICS},
keywords = {MCIA,PERC-UPC,QSE},
mendeley-tags = {MCIA,PERC-UPC,QSE},
number = {12},
title = {{Parameter Identification of DC-DC Converters under Steady-State and Transient Conditions Based on White-Box Models}},
volume = {7},
year = {2018}
}
@article{Rubido2018,
author = {Rubido, Nicolas and Masoller, Cristina},
doi = {10.1140/epjst/e2018-800070-1},
issn = {1951-6355},
journal = {EUROPEAN PHYSICAL JOURNAL-SPECIAL TOPICS},
keywords = {DONLL},
mendeley-tags = {DONLL},
month = {nov},
number = {10-11},
pages = {1243--1250},
title = {{Impact of lag information on network inference}},
volume = {227},
year = {2018}
}
```

Figura 19. Listado de artículos indexados.

Una vez conseguida esta base de datos en formato .txt, se necesitaba transfórmalo a formato .xlsx para poder trabajar con Excel, y así posteriormente hacer el análisis de la estructura social.

El objetivo era tener cada nombre en una columna distinta. Para ello, desde el editor de texto, se sustituyó los “and” y el símbolo “=” por un asterisco “*”.

Ahora desde Excel, importamos esta base de datos, con la condición que los asteriscos sean la referencia para separar por columnas. Una vez importada, podemos empezar con la limpieza de datos.

3.2 LIMPIEZA DE DATOS

Ya en Excel, procedemos a hacer los siguientes cambios:

- Eliminar todo input exceptuando “autor” y “year”.
- Añadir filtros a cada columna.
- Mediante la función “ESPACIOS” de Excel, duplicar cada columna de nombres, eliminando los espacios existentes antes de cada nombre.

Aquí podemos ver las columnas en amarillo duplicadas con la función “ESPACIOS”

	A	B	C	D	E	F
1	author	.T.b	c	ct	d	d1
2	author	Raush, Gustavo	Gamez-Montero, Pedro J	Gamez-Montero, Pedro J	Castilla, Roberto	Castilla, Roberto
4	year	2017)				
5	author	Riba, Jordi-Roger	Moreno-Eguilaz, Manuel	Moreno-Eguilaz, Manuel	Bogarra, Santiago	Bogarra, Santiago
7	year	2018)				
8	author	Rubido, Nicolas	Masoller, Cristina	Masoller, Cristina		
10	year	2018)				
11	author	Sune, Albert	Gibb, Jenny	Gibb, Jenny		
13	year	2015)				
14	author	Quintanilla, R				
16	year	2017)				
17	author	Salazar, Jean C	Weber, Philippe	Weber, Philippe	Nejari, Fatha	Nejari, Fatha
19	year	2017)				
20	author	Riba, Jordi-Roger	Casals-Torrens, Pau	Casals-Torrens, Pau	Bosch, Ricard	Bosch, Ricard
22	year	2014)				
23	author	van der Velde, Oscar A	Montanya, Joan	Montanya, Joan	Soula, Serge	Soula, Serge
25	year	2014)				
26	author	Roca, D	Lloberas-Valls, O	Lloberas-Valls, O	Cante, J	Cante, J
28	year	2018)				
29	author	Rodriguez-Contreras, Alej	ra	ra	{Soledad Marques-Calvo Maria	{Soledad Marques-Calvo Maria
31	year	2016)				
32	author	Salehifar, Mehdi	Moreno-Eguilaz, Manuel	Moreno-Eguilaz, Manuel		
34	year	2016)				
35	author	Xu, Feng	Tan, Junbo	Tan, Junbo	Wang, Xueqian	Wang, Xueqian
37	year	2017)				
38	author	Snasel, Vaclav	Nowakova, Jana	Nowakova, Jana	Xhafa, Fatos	Xhafa, Fatos
40	year	2017)				
41	author	Raush, G	Castilla, R	Castilla, R	Gamez-Montero, P J	Gamez-Montero, P J
43	year	2016)				
44	author	Ventura, Heura	Sorrentino, Luigi	Sorrentino, Luigi	Laguna-Gutierrez, Ester	Laguna-Gutierrez, Ester
46	year	2018)				
47	author	Purlys, V	Maigyte, L	Maigyte, L	Galleivicius, D	Galleivicius, D
49	year	2014)				
50	author	Turdujev, Mirbek	Giden, Ibrahim H	Giden, Ibrahim H	Babayigit, Ceren	Babayigit, Ceren
52	year	2017)				
53	author	Riba, Jordi-Roger	Mancini, Antonio-Giuseppe	Mancini, Antonio-Giuseppe	Abomallek, Carlos	Abomallek, Carlos
55	author					

Y siguiendo con la limpieza:

- Mediante la opción de reemplazar, eliminamos los símbolos “{” y “}”.
- Copiamos toda la base de datos a una hoja nueva únicamente con las columnas modificadas.

Por último, procedemos a la revisión y corrección de errores, además de completar en la medida de lo posible todos los nombres de los autores que únicamente disponen de la inicial con el apellido.

Ejemplo de la base de datos a rectificar:

C	D	E	F	G
Vilaseca, R	Zamora-Munt, J	Masoller, C		
Genesca, M	Romeu, J	Arcos, R		
Zhou, H	Chai, X	Johannes, D	Pope, R	Valls, C
Riba, J	Romeral, L			
Noguera, M	Pincus, D	Buela, G		
Arencon, D	Rodríguez, J	Salazar, A		
Nalawade, D	Zamora, J	Vilaseca, R	Masoller, C	
Ruderman, E	Francesko, A	Tzanov, T	Macedo, M M	Gedanken, A
Botey, M	Staliunas, K			
Roncero, M B	Vidal, T	Rojas, O J		
Cot, M A	Riva, M	Sánchez, R	Sanchez, M	Algaba, I
Codony, F	Sánchez, O	Martín, L	García, J	Mas, J
Xhafa, F				
Cañadas, I	Rodríguez, J			
Aracri, E	Gutiérrez, A	del Rio, J	Torres, A L	Vidal, T
Gorobets, A	Oliva, A	Borrell, R		
Sanz, G	Puig, V	Quevedo, J	Cugero-Escofet, M A	Nejjari, F
Chen, H	Cugero-Escofet, M A	Tino, P	Puig, V	García, D

Figura 20. Base de datos a rectificar.

En este momento, se planteó el uso de algún software tipo OpenRefine para la limpieza de las casi 3000 líneas de Excel. Pero debido a que se tenía que consultar en la mayoría de casos la web FUTUR para completar los nombres de los autores, y eso debía hacerse a mano, ya que podía darse el caso de autores con el mismo apellido e inicial pero distinto nombre, la corrección de los posibles errores se podía ir haciendo al mismo tiempo. Por lo tanto, se procedió a la revisión de línea por línea de manera manual.

3.2.1 Errores encontrados

Aquí explicaremos los principales errores encontrados y sus dificultades de tratamiento, pues ponen de manifiesto la gran problemática que implica no poder trabajar con nombres homogéneos o datos estandarizados.

1. La falta de comas. Muchos de los autores se encontraban escritos con nombre y apellido sin ninguna coma de separador.
2. Comas adicionales al final del último autor. Los artículos con varios autores, solían tener el último autor con una coma extra al final.
3. Autores de nombre compuesto. En estos casos podían pasar varias cosas:
 - a. Segundo nombre indicado como apellido.
 - b. Faltaba el apellido.
 - c. Faltaba el segundo nombre.
 - d. Faltaba la coma separadora.
4. Nombres o apellidos con acentos o "ñ". En este caso aparecían caracteres extra o el autor quedaba separado en dos columnas, lo que dificulta su búsqueda.
5. Mismos autores escritos de diferentes formas. A lo largo de la base de datos se han encontrado autores escritos de diferentes formas, por ejemplo, que en algunos casos se encuentre con los dos apellidos y en otras con solo uno, o que si tienen nombre compuesto se haya dejado únicamente el primero en algunos casos y el segundo en otros.

Cabe destacar también:

6. En artículos de pocos autores con apellidos comunes como Fernández o García, dificultaba su búsqueda.
7. En FUTUR muchos de los autores no tienen ficha propia, lo que da lugar a otras dificultades:
 - a. Imposibilita su búsqueda mediante la plataforma FUTUR.
 - b. En diferentes artículos se puede ver el mismo autor escrito de diferentes formas, como por ejemplo que en algunos artículos solo salga su inicial y apellido, pero en otro si aparezca su nombre.

3.2.2 Soluciones adoptadas

Explicación de las soluciones tomadas para los errores previamente comentados.

- Error 1 y 2: de forma manual, revisando cada línea, se han eliminado o añadido las comas necesarias.
- Errores 3, 4 y 5: en el momento de encontrarse con este tipo de errores, se vio rápidamente que la opción de ir revisando línea por línea e ir modificándolas progresivamente no era eficiente por varios motivos:
 - Podía darse el caso de haber modificado un autor de una manera, y más tarde encontrarlo escrito de otra forma (casos de nombres compuestos o con dos apellidos).
 - Haber encontrado artículos con autores no fichados en FUTUR, y por lo tanto de difícil búsqueda, y más tarde tener que buscar artículos con prácticamente los mismos coautores.

Por estos motivos, y con la intención de agilizar el proceso, y asegurarse de modificar a un autor de la misma forma en toda la base de datos, durante el procedimiento de ir revisando línea por línea, una vez completado el nombre, mediante el comando (ctrl + F) se buscaba el mismo apellido, y se completaba con el nombre encontrado en todos los artículos de la base de datos.

Con este método, otra de las ventajas es que en la mayoría de los casos los coautores eran los mismos en la mayoría de artículos, y, por lo tanto, aparte de completar ese nombre, se podían ir completando algunos de los coautores. Una de las precauciones al hacer esto, era asegurarse de que era el mismo autor el que completábamos con ese nombre, ya que en casos de apellidos comunes como hemos comentado, podía darse el caso de que tuvieran la misma inicial, pero fueran nombres distintos.

Por este motivo, una de las soluciones era comprobar si el resto de coautores eran los mismos, y en caso de no serlos o tener dudas, dejarlo incompleto y proceder a buscar ese artículo en FUTUR.

- Error 7: en casos de artículos donde todos los coautores no tuvieran ficha propia en FUTUR, y tuvieran pocos artículos publicados, lo cual dificultaba su búsqueda, se optó por buscarlos directamente en Google indicando 1 o 2 de los apellidos de los coautores conjuntamente con las palabras “upc” y “futur”. De esta forma, en las primeras posiciones de los resultados de la búsqueda solía salir el enlace a los posibles artículos de los autores, o el acceso a la ficha del autor en FUTUR, en caso de que la tuviese.

3.2.3 Creación de aristas

Una vez limpia la base de datos, se procedió a crear una nueva pestaña "Arestes_03"¹⁴ donde se reflejaban todas las coautorías de los autores por pares, es decir en dos columnas, indispensable para poder trasladarlo posteriormente al NodeXL y crear la red. A parte de esto, se añadió una tercera columna con el año perteneciente a cada colaboración. Esto se hizo así para más tarde, poder crear las redes de coautoría de cada año, acumulando en cada uno la información del anterior, y así poder estudiar la evolución de la red.

Destacar la creación de la pestaña "TotsAutors" donde se reflejaban todos los nombres de la lista. Aquí nos dimos cuenta que aún había errores en los nombres, la mayoría duplicados por la existencia de "espacios" al final del nombre en algunas celdas. Por ello, tuvimos que hacer una segunda limpieza de datos revisando toda esta pestaña, y corrigiéndola en la pestaña de "Arestes_03", que es la que finalmente utilizaríamos para implementar en NodeXL.

¹⁴https://drive.google.com/drive/folders/1ddBHz_HIMoGCR-QWaCAXzW3a6S0vkaH?usp=sharing

3.3 CREACIÓ DE LA RED MEDIANTE NODEXL

Una vez tuvimos la base de datos lista, pasamos al NodeXL para crear la red. En la pestaña de "Edges" ponemos las dos columnas de autores (filtradas previamente por el año 2014), como vértices 1 y 2. Recordamos que en esta red, los autores hacen de vértices o nodos, y sus relaciones de coautoría son las conexiones o aristas que los unen. Una vez implementada esta lista, el grafo obtenido es un grafo aleatorio tal como podemos ver en la siguiente imagen:

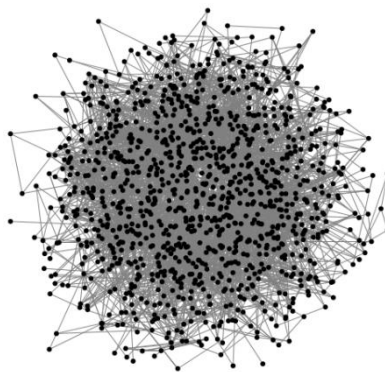


Ilustración 1. Grafo obtenido de la red de coautorías de artículos científicos del año 2014. Elaboración propia.

A partir de aquí, realizamos varios tratamientos a la red de tal manera que sea más fácil analizarla:

1. **"Count and merge duplicate edges"**: en este paso, contamos y fundimos todas las aristas duplicadas. Esto quiere decir que, en el caso de que dos autores hayan colaborado varias veces entre ellos, en vez de tener tantas aristas como colaboraciones tengan, se deja únicamente una arista, y mediante la opción de "insert the counts into a an Edge Weight column" se añade una columna indicando el peso de esa colaboración. El peso es directamente el número de colaboraciones entre esos dos autores. De esta manera, tenemos ya un filtro para observar los autores que más veces colaboran entre ellos, y hacemos que la visualización del grafico final no sea tan cargada.

2. **"Autofill Columns"**: mediante esta herramienta hacemos que las columnas referentes a las características del grafo, como pueden ser el color o la visibilidad entre otras, tanto de los vértices como de los nodos, sean "rellenadas" con los valores o parámetros que nosotros asignemos. En nuestro caso hemos hecho lo siguiente:
- a. **Edge opacity > Edge Weight**: escogemos que la opacidad de cada arista sea proporcional al peso de esta, en una escala del 50 al 100% de opacidad. Como hemos explicado antes, los autores que hayan colaborado más veces entre ellos, tendrán la arista con un 100% de opacidad, y a partir de ahí, cada vez menos hasta el 50% los que tengan únicamente 1 colaboración. Lo hacemos así ya que si escogiéramos una escala del 0 al 100, los autores con una colaboración entre ellos tendrían una arista invisible.
 - b. **Vertex size > Degree**: asignamos al tamaño de cada vértice, su valor en grado, en una escala del 1 al 5. De esta manera, podremos visualizar fácilmente los autores con más grado al tener un tamaño mayor.
 - c. **VertexOpacity > Betweenness Centrality**: asignamos a la opacidad del vértice, el parámetro de centralidad de intermediación, en una escala nuevamente de 50 a 100. Recordemos que este parámetro mide la frecuencia en la que un nodo sirve de puente dentro de la ruta más corta entre otros dos nodos de la red. De esta manera podemos ver como los vértices más oscuros, representan una mayor importancia dentro de la red.
 - d. **Vertex Tooltip > Vertex**: aquí únicamente asignamos a la herramienta "tooltip" el vértice. Esto hace que cuando nos posicionemos con el ratón encima de un vértice, salga una etiqueta de información emergente con lo que queramos, en nuestro caso, el nombre del autor.

3. Group by cluster: este paso es fundamental a la hora de visualizar nuestra red. En nuestro caso es evidente que nos encontramos con un grafo no conexo, es decir, no todos los pares de nodos están conectados entre sí.

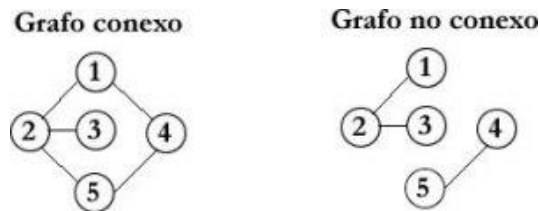


Figura 21. Grafo conexo y no conexo. Fuente¹⁵.

Esto es lógico ya que resulta evidente pensar que haya autores de un departamento que hayan colaborado entre ellos y no se hayan involucrado en colaborar con otros autores de otros campos de estudio. Por lo tanto, al agrupar por "clusters" (o módulos), visualizaremos todos los grupos por separado, teniendo diferentes grafos, ahora sí, conexos.

Como resultado final de estos tratamientos a la red, obtenemos el conjunto de grafos de la figura inferior:

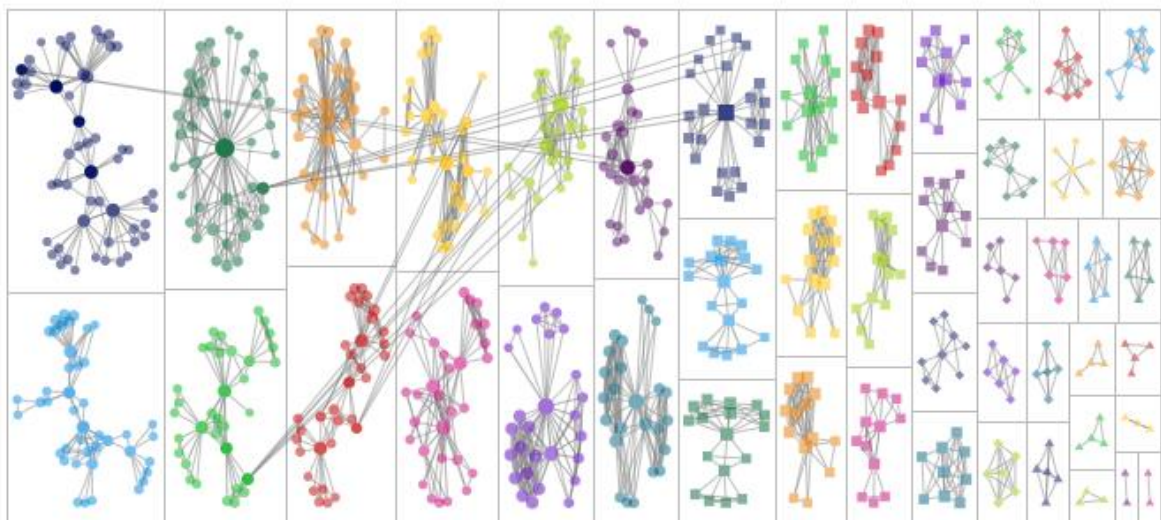


Figura 22. Red de coautorías de artículos científicos del año 2014. Elaboración propia mediante NodeXL.

¹⁵<https://matematicadicreta.wordpress.com/matematica-discreta-2/caracterizacion-de-los-grafos/>

Una vez hemos tratado visualmente la red, mediante "Graph Metrics" hemos calculado todas las métricas que NodeXL nos proporciona. Con esto, obtenemos las métricas para la red de coautorías de artículos científicos para el año 2014. A partir de aquí, hemos repetido todo este proceso para el resto de años hasta 2018, haciendo que en cada año se acumulara la información de las redes anteriores. Es decir, en nuestro siguiente grafo de 2015, añadimos las coautorías de 2014 y 2015, y así sucesivamente, hasta tener un grafo final de 2018 con las coautorías de los 5 años estudiados.

Con toda esta información, podemos crear una tabla con la evolución de todas estas métricas (pestaña de Overall Metrics) y proceder a comentarlas en nuestro siguiente apartado de resultados a partir de los gráficos lineales extraídos de cada parámetro en función de los años.

3.4 ANÁLISIS DE LAS DISTRIBUCIONES DE PROBABILIDAD

De momento, se han tratado todas las métricas de forma más superficial, observando su progreso durante estos cinco años usando principalmente su valor promedio, sus valores máximos o su mediana en el caso del PageRank.

Ahora, profundizaremos más en cada métrica por separado, utilizando la función de distribución. Con este análisis pretendemos entender mejor la evolución de este parámetro en la red a lo largo de los años, y ver qué tipo de distribución tiene cada parámetro. Como hemos visto antes, las redes aleatorias de Erdős–Rényi y las de pequeño mundo (Watts–Strogatz) siguen una distribución de Poisson, mientras que las de libre escala siguen la ley de potencia. Al estudiar cada parámetro con la función de distribución, podremos tener más información a la hora de determinar qué tipo de estructura tiene nuestra red y determinaremos si tiene sentido hablar de promedio o no a la hora de tratar algunos parámetros.

Como podemos ver en el Excel¹⁶, contamos con una pestaña para cada parámetro analizado: "Degree", "Betweenness", "Closeness", "EigenvectorCentrality", "PageRank" y "Clustering". El proceso que hemos seguido para cada una de estas métricas es el siguiente:

1. Seleccionar del grafo de 2014 todos los valores pertenecientes al parámetro que queremos analizar. En el caso del grado, seleccionamos todos los valores de grado de todos los autores existentes en el grafo, y los introducimos en una columna del Excel ordenados de menor a mayor.

¹⁶https://drive.google.com/drive/folders/1ddBHz_HIMoGCR-QWaCAXzW3a6S0vkaH?usp=sharing

2. Para sacar la distribución de cada valor, en una columna inmediatamente a la derecha a los valores, y mediante la función "contar.si" de excel, seleccionamos todo este rango de valores extraídos del grafo, y establecemos como condición que sea mayor o igual que el valor único que tenemos en aquella fila. Toda esta función, la dividiremos entre el número total de valores analizados para obtener un resultado entre 0 y 1.
3. Repetimos este proceso para el resto de años, con los grafos obtenidos anteriormente.
4. Extraemos los gráficos para ver la función de distribución de cada año, recordando que en cada año posterior, se almacena la información de los años anteriores.
5. Añadimos los cinco gráficos en uno, poniendo los ejes en escala doble logarítmica (para ver si obtenemos una recta), y así ver la evolución de todas las distribuciones de una vez, teniendo como resultado, en el caso del grado:

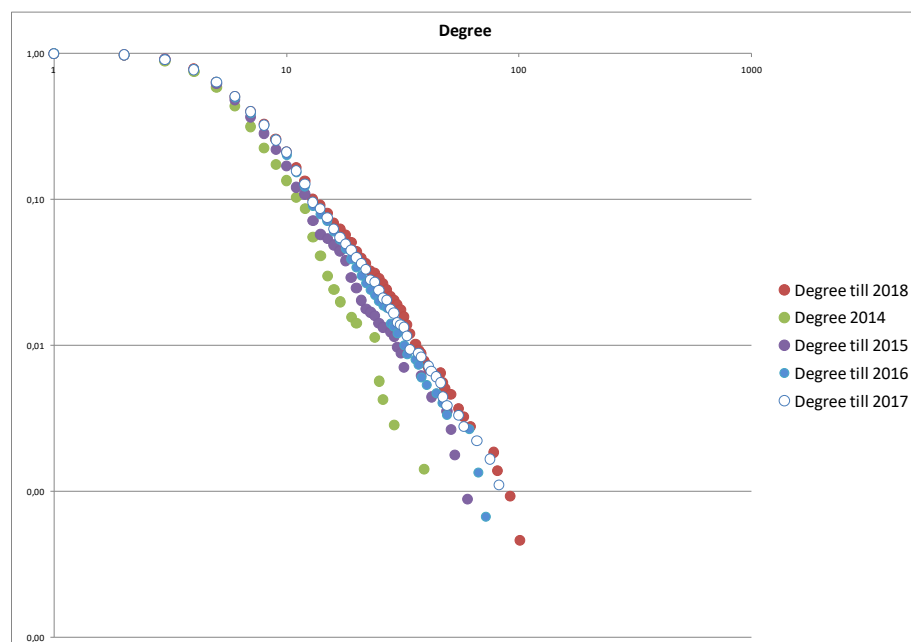


Figura 23. Gráfico de la función de distribución para el grado en los diferentes años. Elaboración propia en Excel.

6. Este proceso lo repetimos para el resto de métricas. Los resultados de este análisis serán comentados más tarde en nuestro apartado de resultados.

3.6 REDES EGO

Las redes ego (*in.*, *ego networks*) consisten en un nodo focal y los nodos a los cuales está directamente conectado este nodo principal o focal, con sus respectivas conexiones o aristas. Es decir, partiendo de un nodo focal, se representarán únicamente sus vecinos directos. Estos vecinos son llamados normalmente "alters". Y obviamente, cada uno de estos "alters" pueden tener su propia "ego network".

Siguiendo con nuestro estudio, y una vez entendido que los resultados en cuanto a "degree" y "PageRank" no parece que sigan una distribución potencial, pero sí de larga cola, es decir, logarítmica normal, y por lo tanto, signifique la existencia de pocos nodos altamente conectados (*hubs*), procedemos a analizar las redes ego de los 10 nodos con más grado en 2018 (teniendo en cuenta que la información que extraemos de nuestro grafo de 2018 acumulado incluye todas las redes de coautoría desde 2014 hasta 2018). Los nodos o autores que resultaron ser los mayores en cuanto a grado son los siguientes:

2018	Degree	BTW	Closeness	Eigenvector	PageRank	Clustering
Puig, Vicenc	101	105748	0,000	0,000	12,71	0,06
Xhafa, Fatos	92	135214	0,000	0,000	11,22	0,06
Tzanov, Tzanko	81	642711	0,000	0,039	7,16	0,11
Staliunas, Kestutis	78	214114	0,000	0,000	8,29	0,09
Masoller, Cristina	62	189531	0,000	0,000	8,99	0,07
Morato, Jordi	62	256706	0,000	0,021	6,23	0,12
Pujol, Jaume	58	35498	0,000	0,000	6,39	0,11
Montanya, Joan	55	48733	0,000	0,001	4,95	0,16
Oliva, Asensio	51	89972	0,000	0,000	5,75	0,13
Vilaseca, Meritxell	51	139663	0,000	0,000	5,12	0,15

Tabla 3. Clasificación de los 10 autores con más grado en 2018 acumulado. Elaboración propia en Excel.

Cabe destacar que de estos 10 autores, 8 de ellos también son los autores con mayor PageRank en 2018. Esta clasificación, junto a las clasificaciones para el resto de los años, podemos encontrarlas en Excel en las pestañas de "Class.Degree" y "Class.PageRank"¹⁷.

¹⁷ https://drive.google.com/drive/folders/1ddBHoz_HIMoGCR-QWaCAXzW3a6S0vkaH?usp=sharing

3.6.1 Análisis de las redes ego de los 10 nodos con más grado

Como hemos comentado, procedemos a analizar a los 10 autores con más grado en 2018 acumulado, es decir, a los autores con mayores conexiones o enlaces. En este análisis, nos centraremos únicamente en sus valores para el grado y PageRank, ya que son los parámetros verificados después del estudio con la función de distribución. Para hacerlo, necesitamos hacer el estudio de sus redes ego para los 5 años por separado, para más tarde, ver su evolución.

Para realizar este paso, el procedimiento a seguir ha sido el siguiente:

1. En el grafo de 2014, y para el primer autor, en este caso Vicenç Puig, seleccionar su red ego. Para hacer esto, seleccionamos el subgrafo del autor con un nivel de 1. Esto significa que únicamente seleccionará los vecinos directos a este autor (o de grado 1). Las redes ego, son exactamente eso, la red de sus vecinos directos.

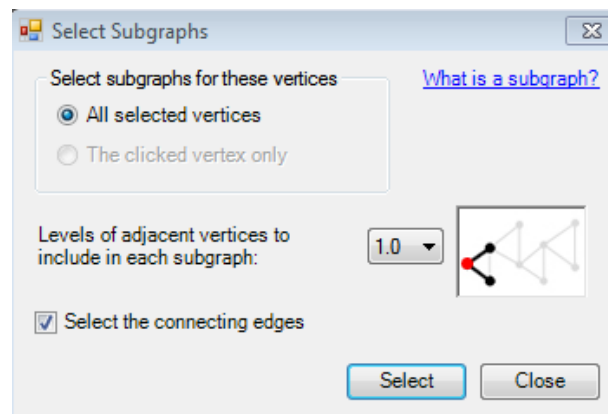


Figura 24. Selección de sub grafos en NodeXL.

2. Pasar a Excel los datos de la red ego de este autor, seleccionando únicamente los nombres de sus enlaces, y sus valores de degree y PageRank, ordenados de mayor a menor.
3. Repetir el proceso para el resto de años, y para los 9 autores restantes.

- Una vez con las tablas, creamos los gráficos de todas las redes ego para cada año por separado, y para todos los autores, para visualizar la evolución del grado en función del PageRank.

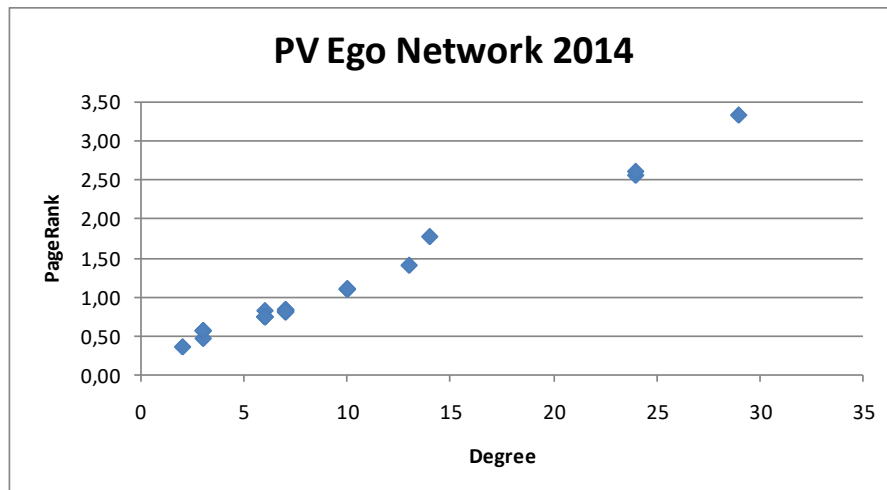


Figura 25. Evolución del grado en función del PageRank para la red ego de Vicenç Puig en 2014.

- Por último, para cada año, hemos querido ver el comportamiento del nodo focal (refiriéndonos al autor del que hacemos el estudio) frente a su segundo autor con más grado y PageRank. Para ello, simplemente hemos visualizado la diferencia de grado del nodo focal frente al segundo (DeltaDeg) para cada año, y la diferencia de PageRank de estos mismos nodos (DeltaPR) para cada año. Estos resultados los hemos incluido en una tabla, para la creación de un gráfico dinámico.

SUMMARY			
PDI	Any	DeltaDeg	DeltaPR
Puig, Vicenc	2014	5	0,770
Puig, Vicenc	2015	23	3,293
Puig, Vicenc	2016	40	5,771
Puig, Vicenc	2017	50	6,802
Puig, Vicenc	2018	65	8,763

Tabla 4. Visualización de la diferencia de grado y de PageRank entre el nodo focal Vicenç Puig y su segundo nodo Joseba Quevedo para los años desde 2014 hasta 2018.

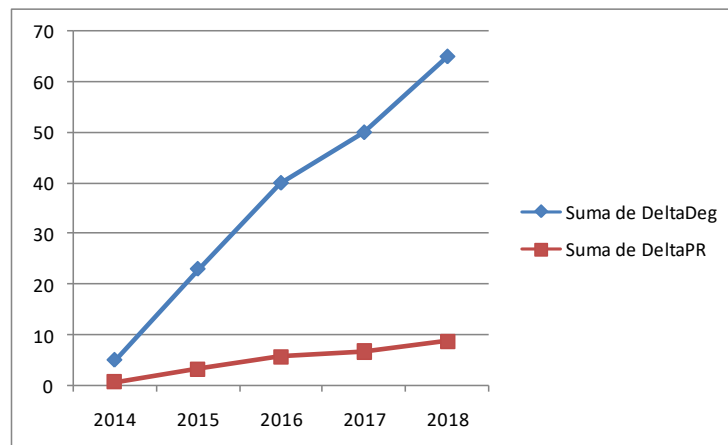


Figura 26. Gráfico para la visualización de la diferencia de grado y de PageRank entre el nodo focal Vicenç Puig y su segundo nodo para los años desde 2014 hasta 2018.

Todas estas tablas se pueden encontrar en el Excel, en las pestañas de cada autor. La interpretación de los resultados extraídos será comentada más tarde en el apartado de resultados.

3.7 ANÁLISIS DE LAS TOPOLOGÍAS DE RED

Siguiendo con el estudio, después de analizar las redes ego, pasamos a estudiar la topología de las redes. La topología de red es el patrón de interconexión de los nodos, es decir, el mapa físico o lógico de la red o la forma en que está diseñada. A continuación tenemos un ejemplo de las posibles topologías que podemos encontrar en una red.

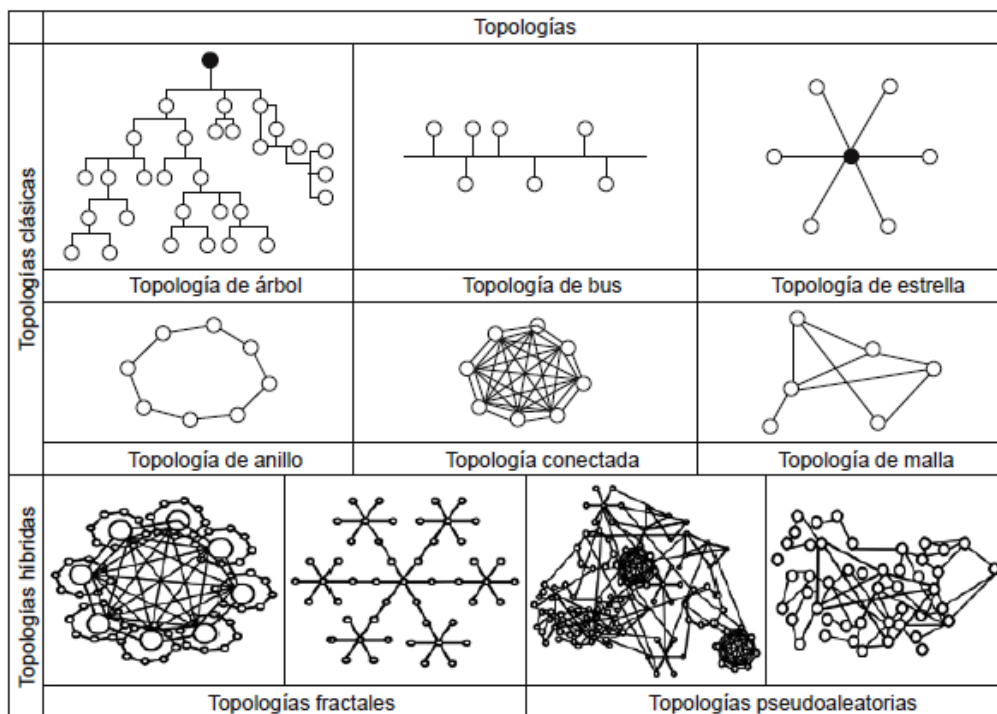


Figura 27. Topologías en una red. Fuente¹⁸.

Lo que queremos saber ahora, es si el tipo de topología de cada red, guarda relación con los resultados obtenidos para las redes ego ya analizadas, y como afecta la topología de red a otros grupos. Analizaremos a continuación las topologías de las redes ego ya tratadas, y del resto de grupos o "clusters" de mayor importancia, siempre con los datos acumulados del grafo de 2018.

¹⁸<https://revistas.urosario.edu.co/xml/3596/359638976005/index.html>

3.7.1 Topología de las redes ego analizadas

En el anexo "7.1 Topología de las 10 redes ego con más grado en 2018" y en el "7.2 Topología del resto de grupos de 2018" podemos ver las topologías de todas las redes analizadas. Como ejemplo, tenemos la topología de Vicenç Puig.

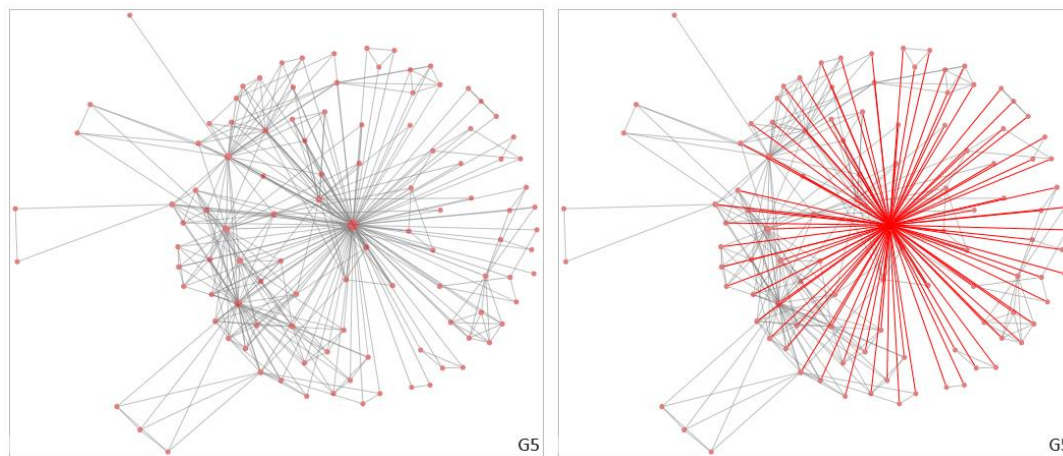


Figura 28. Grupo G5 a la izquierda y la red ego de Vicenç Puig iluminada a la derecha. Elaboración propia en NodeXL

Este tipo de topología estrellada que presenta esta red, se presenta en la mayoría de redes analizadas. Esta topología presenta el nodo altamente conectado o "hub" en el medio, conectado al resto de nodos.

Una vez con las imágenes de todas las topologías, hemos vuelto a hacer el análisis presentado en el apartado "3.6.1 Análisis de las redes ego de los 10 nodos con más grado" para el resto de grupos de 2018 (evitando los grupos con pocos nodos y de menor importancia). Hemos partido en este caso del grupo del cual tenemos su topología, hemos buscado el autor con más grado de ese grupo, y hemos hecho su seguimiento estudiando la red ego de este autor en los años anteriores. En el Excel se pueden ver estos análisis a partir de la pestaña "G4", y en las pestañas acabadas en "-S" el resumen de la tendencia de esta evolución.

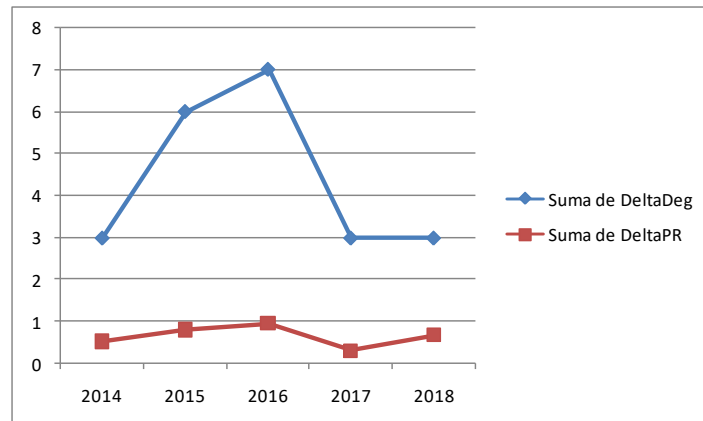


Figura 29. Gráfico para la visualización de la diferencia de grado y de PageRank entre el nodo focal Jasmina Casals Terre y su segundo nodo para los años desde 2014 hasta 2018.

En el apartado de resultados haremos un resumen de las tendencias encontradas, y en el apartado de conclusiones, analizaremos si existe alguna relación entre la topología de red y las tendencias encontradas en las gráficas.

4. RESULTADOS Y DISCUSIÓN

4.1 MÉTRICAS GENERALES

A continuación, vemos las principales métricas obtenidas mediante NodeXL de los grafos obtenidos y su evolución a lo largo de los años escogidos.

OVERAL METRICS	YEARS ACCUMULATED				
	2014	2015	2016	2017	2018
Vertices	702	1124	1500	1800	2154
Unique Edges	2100	3762	5313	6547	8028
Edges With Duplicates	0	0	0	0	0
Total Edges	2100	3762	5313	6547	8028
Self-Loops	0	0	0	0	1
Connected Components	42	41	38	37	28
Single-Vertex Connected Components	0	0	0	0	0
Maximum Vertices in a Connected Component	80	382	1027	1457	1964
Maximum Edges in a Connected Component	229	1341	3813	5577	7574
Maximum Geodesic Distance (Diameter)	8	18	19	15	18
Average Geodesic Distance	3,05	6,08	7,86	7,38	7,62
Graph Density	0,008	0,006	0,005	0,004	0,003
Modularity	0,95	0,94	0,93	0,90	0,88
Minimum Degree	1	1	1	1	1
Maximum Degree	39	60	72	82	101
Average Degree	6,0	6,7	7,1	7,3	7,5
Median Degree	5,0	5,0	5,5	6,0	6,0
Minimum Betweenness Centrality	0	0	0	0	0
Maximum Betweenness Centrality	1599	39752	233677	443520	642711
Average Betweenness Centrality	40	451	2464	3779	5939
Median Betweenness Centrality	0	0	0	0	0
Minimum Closeness Centrality	0	0	0	0	0
Maximum Closeness Centrality	1	1	1	1	1
Average Closeness Centrality	0,053	0,032	0,026	0,022	0,015
Median Closeness Centrality	0,017	0,002	0,000	0,000	0,000
Minimum Eigenvector Centrality	0	0	0	0	0
Maximum Eigenvector Centrality	0,072	0,058	0,039	0,043	0,039
Average Eigenvector Centrality	0	0	0	0	0
Median Eigenvector Centrality	0	0	0	0	0
Minimum PageRank	0,28	0,26	0,24	0,24	0,24
Maximum PageRank	4,89	7,63	9,20	10,23	12,71
Average PageRank	1	1	1	1	1
Median PageRank	0,92	0,88	0,85	0,85	0,84
Minimum Clustering Coefficient	0	0	0	0	0
Maximum Clustering Coefficient	1	1	1	1	1
Average Clustering Coefficient	0,86	0,85	0,85	0,85	0,85
Median Clustering Coefficient	1	1	1	1	1

Tabla 5. Métricas generales. Elaboración propia en Excel.

4.1.1 Componentes conectados

Aquí apreciamos la evolución de los componentes del grafo conectados. Entendemos por componentes conectados o conexos, aquellos subgrafos donde para cada par de nodos de éste hay un camino por el cual están conectados. Se entiende también que, en este caso, para una red de coautorías de artículos científicos de una universidad, los diferentes subgrafos o componentes, se dan cuando se tratan de diferentes ámbitos de estudio o departamentos.

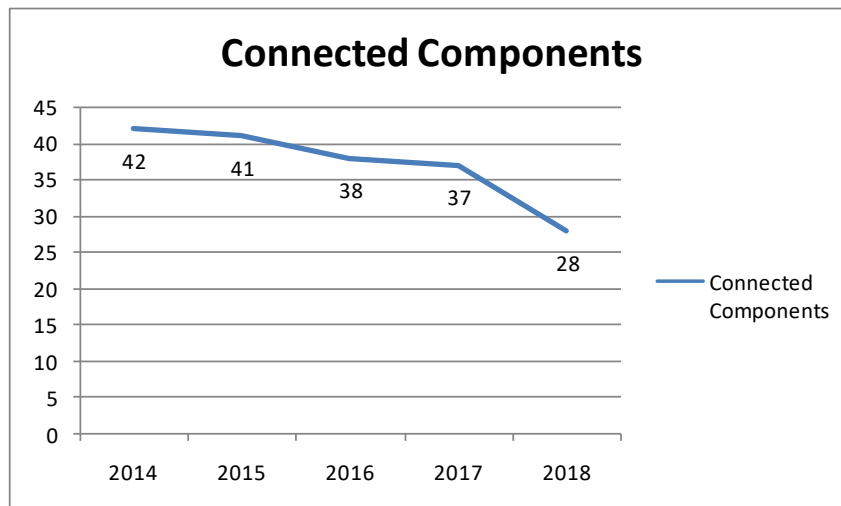


Figura 30. Componentes conectados. Elaboración propia en Excel.

Vemos como la tendencia es a que estos componentes se reduzcan con el paso del tiempo. Es decir, a medida que pasa el tiempo, los componentes o subgrafos aislados, pasan a unirse a un grafo mayor o principal, quedando de esta manera, menos componentes desconectados. La conclusión lógica a esta tendencia en una red de coautoría de artículos científicos nos indica que a medida que pasa el tiempo, y contando con cada vez una base de datos más amplia, teniendo constancia de todas las colaboraciones que se han hecho por parte de estos autores, existe una colaboración mayor entre departamentos o ámbitos de estudio. Es significativa la disminución del 24% entre los años 2017 y 2018.

Es decir, en el momento que por ejemplo alguien del departamento de ingeniería química colabora con alguien del departamento de ciencia e ingeniería de los materiales, estos dos subgrafos independientes donde solo se colaboraba entre los miembros del mismo, pasan a estar conectados, reduciendo así una unidad en este parámetro.

Este parámetro, nos indica también la importancia y la evolución de la interdisciplinariedad en la educación y en la investigación. Cada vez, se da más importancia a la aplicación de todos los conocimientos adquiridos en diversas materias a un proyecto conjunto. La educación actual en muchos colegios está evolucionando a una educación basada en proyectos (ABP), donde esta metodología ayuda a potenciar las habilidades sociales, el pensamiento crítico, la creatividad, el aprendizaje autónomo y la aplicación de los conocimientos adquiridos a situaciones reales. En la investigación, y viendo el bajón significativo de 2017 a 2018, vemos cómo a medida que evolucionamos, necesitamos cada vez más la unión de los conocimientos en diversas aéreas para dar paso a nuevas teorías y oportunidades.¹⁹

4.1.2 Vértices y aristas.

Aquí vemos la evolución de los vértices y las aristas durante estos años. Se ve claramente una relación lineal positiva constante. A cada año que pasa, se incrementan las publicaciones de artículos científicos, lo que hace que incrementen las autorías y las relaciones entre ellos.

¹⁹Un claro ejemplo es la búsqueda de una teoría del todo, que unifique la teoría de la relatividad y la teoría cuántica, que pueda explicar las leyes físicas tanto en el mundo macroscópico como microscópico.

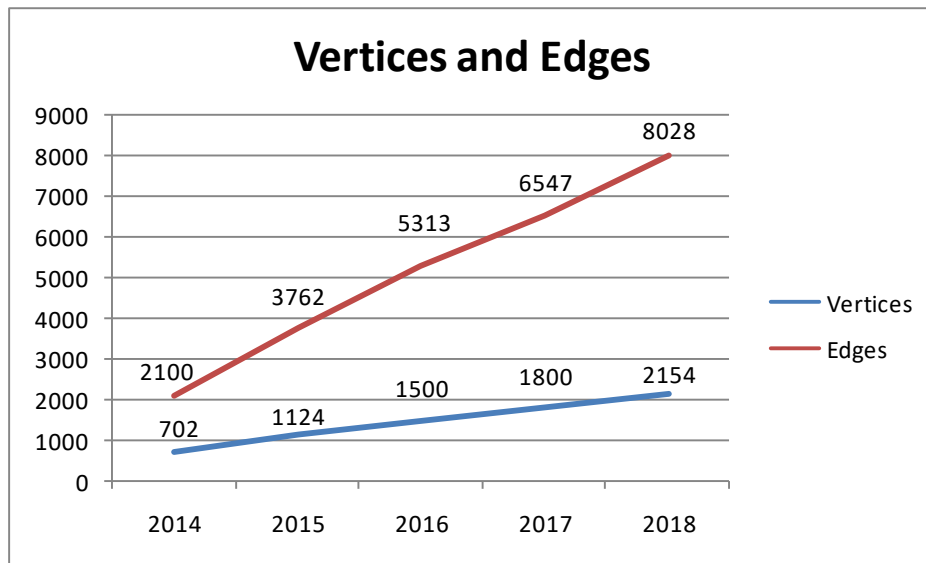


Figura 31. Vértices y aristas. Elaboración propia en Excel.

4.1.3 Grado promedio

Cuando queremos hablar de centralidad en un grafo, el primer parámetro que se debe analizar es la centralidad de grado. Esta medida nos indica el número de enlaces o conexiones conectados al vértice, es decir, el número de aristas pertenecientes a ese nodo. Eso quiere decir que a mayor valor de grado, más importancia tendrá ese nodo o autor dentro de la red. En nuestro caso, el valor mínimo de grado es 1 (como podemos ver en la Tabla 5 de Métricas generales), ya que previamente se han descartado los artículos realizados por únicamente una persona (siendo un número bastante reducido) ya que no aportaban información para el análisis de la red social del PDI. El valor de grado máximo por año también podemos verlo en la Tabla 5, aunque para ver la tendencia que sigue con el tiempo, analizamos el grado promedio.

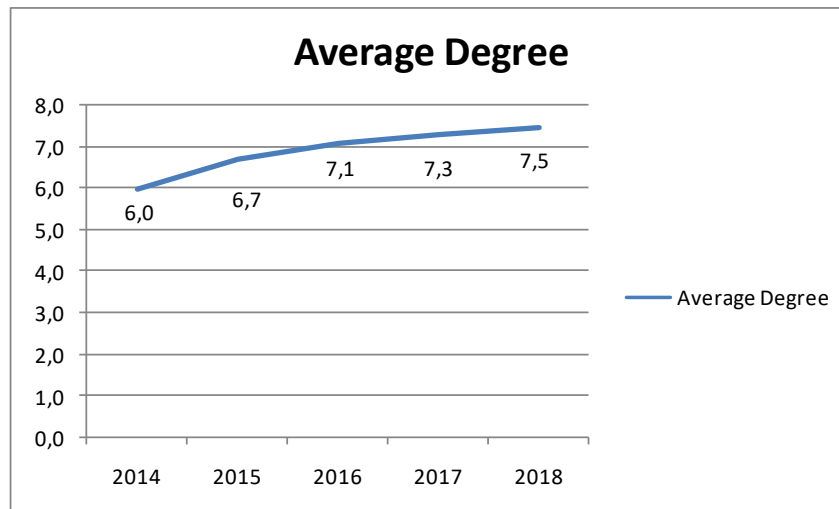


Figura 32. Grado promedio. Elaboración propia en Excel.

En este gráfico se puede ver la evolución del grado promedio de nuestra red social durante los años seleccionados. Se aprecia, como es lógico, que siga una relación lineal positiva, aunque como vemos, la tendencia es a que se estabilice con el tiempo. Esto es coherente, sabiendo que durante la creación de los grafos hemos fusionado todas las aristas duplicadas.

Es decir, los pares de autores que han trabajado más de una vez conjuntamente, no suman más aristas o grados a sus vértices, simplemente se fusionan, y se contabiliza en otra columna como peso de la arista. De esta manera, se puede ver fácilmente que par de autores han trabajado más a menudo entre ellos. Por lo tanto, sabiendo esto, y que en la red social analizada los autores suelen ser prácticamente los mismos durante unos cuantos años, añadiéndose pocos nuevos autores cada año, la tendencia es que el grado siga aumentando por cada vez tener más información en la base de datos, pero cada vez aumente en menor medida.

Cabe decir que a pesar del mayor aumento de aristas respecto al de nodos, y que el grado promedio tan sólo incrementa de 6.0 a 7.5 (Tabla 5), este mínimo aumento no es significativo como valor medio, pues la distribución de grado aparece con cola larga, hecho que implica una acumulación de aristas no generalizada sino en ciertos nodos (o *hubs*).

4.1.4 Modularidad

La modularidad mide la fuerza de la división de una red en módulos o grupos entre los actores de los cuales existen más conexiones de las que se esperarían de forma puramente accidental. Es decir, cuando la modularidad de una red sea alta, significará que existen conexiones densas entre los nodos pertenecientes a cada módulo o grupo, pero habrá escasas conexiones entre los nodos o vértices pertenecientes a diferentes módulos.

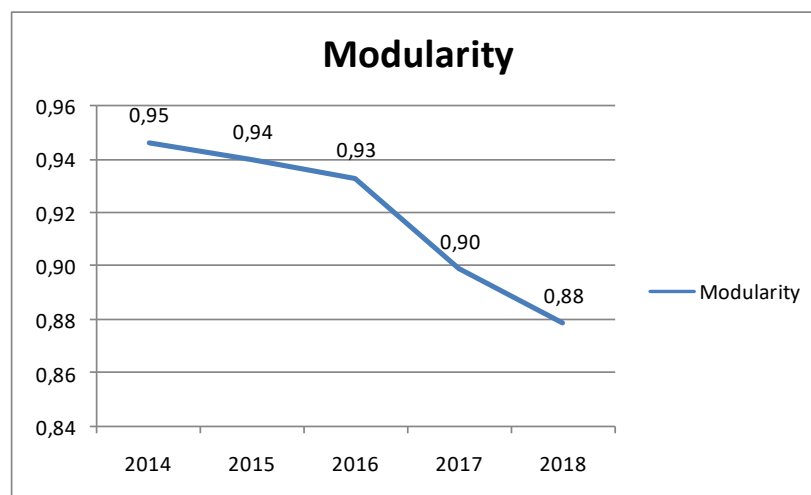


Figura 33. Modularidad. Elaboración propia en Excel.

En este caso vemos como la modularidad en 2014 es de 0,95, siendo 1 el valor máximo al que puede llegar. A partir de aquí, la tendencia de este parámetro es a disminuir poco a poco, aun manteniéndose en unos valores altos.

Esto nos dice que la red analizada del profesorado docente de investigación está bastante segmentada en subgrafos o módulos en los cuales hay una fuerte colaboración entre ellos, pero pocas conexiones entre los diferentes grupos. Cuando hablamos de grupos o módulos, podemos entender que hablamos de departamentos dentro de la universidad. Es lógico pensar que, dentro de un departamento, el número de colaboraciones sea alto, es decir, que las conexiones sean densas entre los diferentes nodos o personas, ya que se trabaja dentro de un mismo ámbito.

Ahora bien, cuando se trata de colaborar entre diferentes departamentos, las conexiones son más escasas. A pesar de esto, la tendencia a medida que pasan los años, nos indica que las colaboraciones entre diferentes departamentos aumentan, y las existentes dentro de estos grupos se reducen. Esto hace que, en una visión general, pasemos de ver una red muy segmentada en la que en cada grupo vemos conexiones muy densas, a un grafo más equilibrado, donde la densidad de estas conexiones ya no se concentra por módulos sino que se reparte por todas las personas de la red. Esto nos confirma una vez más la importancia y la evolución de la interdisciplinariedad en la investigación, de la cual ya habíamos hablado en el apartado de componentes conectados de la red.

4.1.5 Densidad de la red

Cuando hablamos de densidad de la red, nos referimos a la relación entre el número de aristas del grafo y el número de aristas máximo que este grafo tendría si todos los actores estuvieran conectados entre ellos. Recordemos que este valor puede ir desde 0, cuando es grafo es disperso, a un valor de densidad máximo de 1, cuando el grafo es completo.

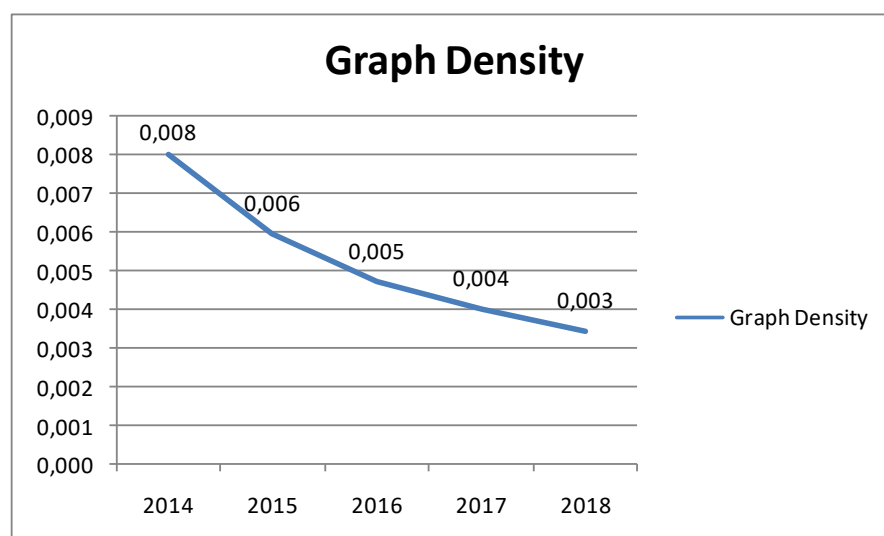


Figura 34. Densidad de la red. Elaboración propia en Excel.

En este caso, el valor general para la red estudiada es muy bajo, con tendencia a ir disminuyendo con el tiempo. No sorprende, ya que para que el grafo fuera completo o con una densidad alta, toda persona del profesorado de investigación presente en la red debería colaborar con prácticamente todo el resto para generar las aristas o conexiones que densificaran la red, y esto no es así. Si nos damos cuenta, en el último año, tenemos 2154 personas en la red, y el valor máximo de grado encontrado es de 101, como podemos ver en la tabla de métricas generales. Esto quiere decir que el grado de la persona de la red que cuenta con más colaboraciones o aristas es de 101, cuando su máximo podría ser de 2153.

Sabemos también que el promedio de grado en todo el grafo durante estos es de un valor aproximado a 7, es decir, que de media, todo el PDI colabora con aproximadamente 7 personas, siendo un valor muy bajo para un grafo en el que hay desde 702 personas en el primer año estudiado, a 2154 en 2018. Por lo tanto, es coherente que el grafo sea catalogado como disperso y de baja densidad.

4.1.6 Centralidad del vector propio y PageRank

Analizamos a continuación la centralidad del vector propio y el PageRank. La centralidad del vector propio o (in., *eigenvector centrality*) mide la influencia o conectividad de un nodo en la red. Es decir, en nuestro caso, una persona o autor que tenga pocas conexiones, puede poseer un valor alto de centralidad del vector propio, si estas pocas conexiones que tiene a su vez están muy bien conectadas. Dicho de otra manera, este parámetro calcula el grado de los vértices vecinos conectados al vértice principal.

El algoritmo PageRank, por otro lado, es una variante del "eigenvector centrality" creado por Google en 1999 y utilizado en su motor de búsqueda. La diferencia en este caso, de manera muy simplificada, es que aparte de medir la cantidad de enlaces que posee una web o en nuestro caso un autor, mide también la "calidad" de estos enlaces. Este valor se dejó de actualizar en 2013 para el público, ya que daba lugar a confusiones como la de ser el único parámetro válido para medir el posicionamiento de una web (cosa que no es cierta), pero el PageRank interno sigue usándose en Google, y sigue teniendo su importancia a la hora de medir las métricas de una red.

En ambos casos los dos parámetros se valoran igual: a valor más alto, mayor importancia.

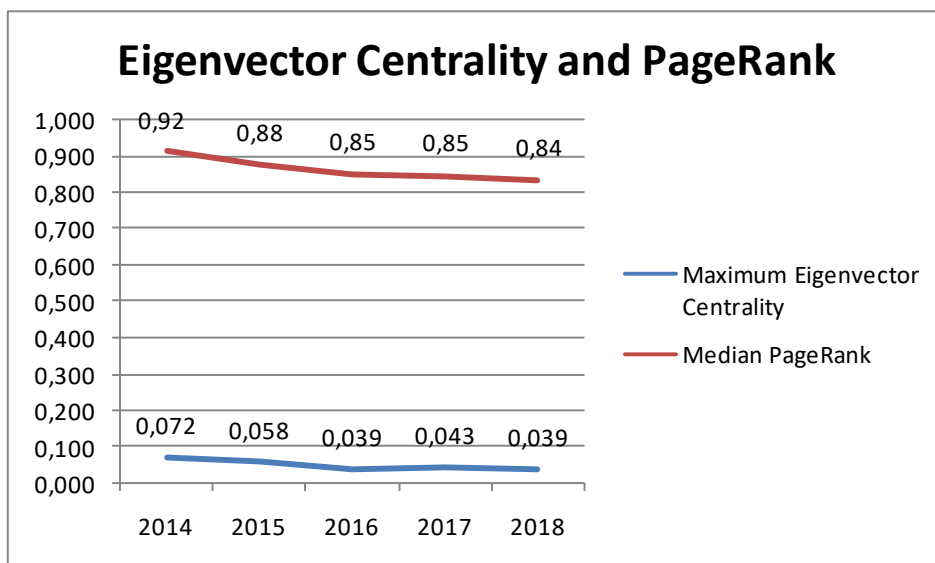


Figura 35. Centralidad del vector propio y PageRank. Elaboración propia en Excel.

Debido a que el promedio de los dos parámetros nos daba muy poca información sobre la tendencia que tenían estos valores a lo largo de los años, hemos escogido el valor máximo para la centralidad del vector propio (in., *eigenvector centrality*) y la mediana del PageRank. Lo que encontramos es que en ambos casos la tendencia que siguen estos parámetros es a disminuir.

4.1.7 Distancia Geodésica

Otro de los parámetros a estudiar es la distancia entre dos vértices. Llamamos distancia o distancia geodésica al mínimo número de nodos que hay que recorrer para unirlos. Es decir, el camino más corto que existe entre dos vértices, entendiendo que puede haber más de uno.

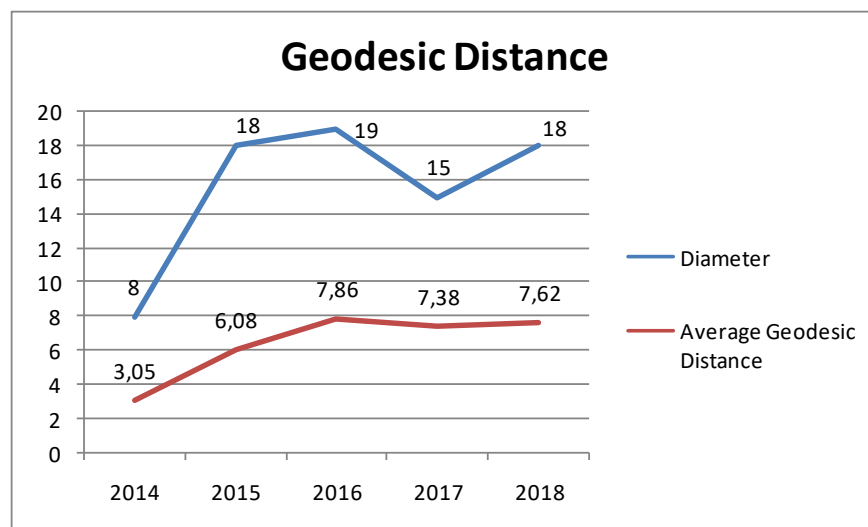


Figura 36. Distancia Geodésica. Elaboración propia en Excel.

En este caso se han representado en la Figura 36 la distancia geodésica máxima del grafo, a la que llamamos diámetro, y la distancia geodésica media entre todos los nodos. En ambos casos, se aprecia como la tendencia de estos dos parámetros es la de aumentar con los años. Esto es lógico si entendemos que a medida que pasan los años, se introducen nuevos autores en la red de forma jerárquica, es decir, siguiendo un único camino que hace aumentar esta distancia.

Pero en realidad, tenemos esta tendencia debido a que en el año 2014 la distancia era muy inferior a la que podemos encontrar en los 4 años siguientes. Si obviamos este primer año, lo que observaremos sería un parámetro que se mantiene constante con el tiempo, lo que da a pensar que los nuevos autores se relacionan de una manera colaborativa en esta red.

4.1.8 Intermediación

Cuando hablamos de intermediación (in., *betweenness centrality*), seguimos hablando de distancia en la red. En este caso, este parámetro mide con qué frecuencia un vértice particular está situado en el camino más corto (o lo que es lo mismo en la distancia geodésica) entre otros dos vértices de la red. Podemos imaginarnos este parámetro como una medida de la capacidad de unir o de estar “en medio” de los otros nodos. Si este nodo permite unir muchos vértices de la red por su camino más corto, este parámetro aumentará para este nodo, y nos indicará una mayor importancia de este nodo en la red, ya que si este vértice desapareciera, la unión del resto de vértices debería hacerse por otros caminos más largos.

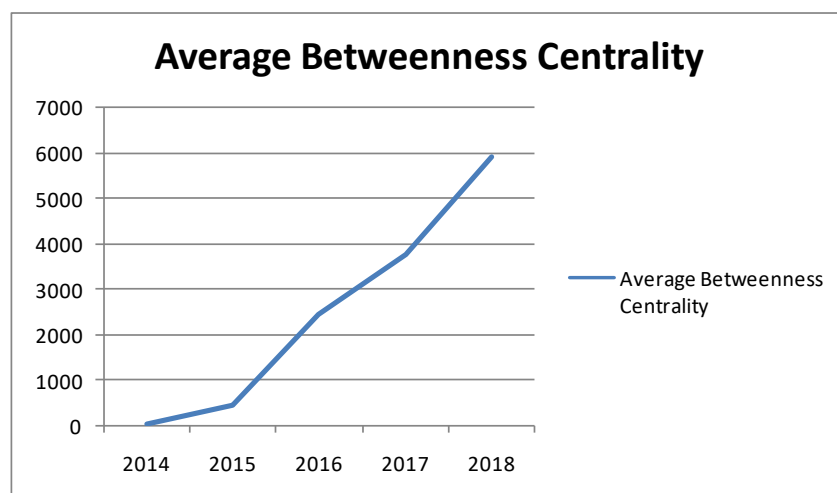


Figura 37. Intermediación. Elaboración propia en Excel.

En este caso, hemos representado la media de intermediación. Lo que apreciamos es claramente una tendencia a aumentar considerablemente cada año que pasa. Recordemos que este parámetro nos indica la media de intermediación para todos los nodos del grafo por cada año. Es decir, si cada año crece notablemente este valor, lo que nos dice es que en general la mayoría de autores incrementan su "betweenness centrality". En nuestra red, indicará que los autores ganarán importancia al ser una persona "puente" para el resto de autores. Esto quiere decir que la persona con el valor más alto en intermediación, es el mejor nexo al que puedes acudir para llegar de la forma más rápida a la mayoría del profesorado docente de investigación.

Volviendo a la tendencia que vemos en el gráfico, y para que esto ocurra, es lógico pensar en un grafo de estructura colaborativa, al menos de manera global, ya que al crecer en cuanto a autores en la red de manera distribuida, la mayoría de ellos aumentarían este parámetro. No podría ser así en un grafo claramente jerárquico, ya que, en este, los autores o nodos que aumentarían este parámetro serían unos pocos, y el resto mantendrían un valor bajo de intermediación, por lo que al hacer la media en un grafo que como en el nuestro suma 8028 vértices en su último año, se vería claramente afectado este valor.

4.1.9 Cercanía y coeficiente de agrupamiento

Cuando hablamos de cercanía, nos referimos a la distancia media entre un vértice y el resto de los vértices de la red. En este caso, cuanto menor sea su "closeness centrality", significa que aquella persona o autor está directamente conectado o a una pequeña distancia del resto de personas en la red. Es decir, que a menor valor en este parámetro, más importancia tendrá la persona en la red.

Por otro lado, el coeficiente de agrupamiento es una medida del grado en que los nodos en un gráfico tienden a agruparse en forma de triángulo en una red. En nuestro caso, para el profesorado docente de investigación, podemos decir que si los vecinos o coautores de una persona en particular también han trabajado o publicado algún artículo entre ellos (y por lo tanto también están conectados entre ellos), se tendrá un coeficiente de agrupamiento alto. Por el contrario, si los coautores de una persona dada no han trabajado entre ellos, este pequeño subgrafo contará con menos conexiones, y su coeficiente de agrupamiento será menor.

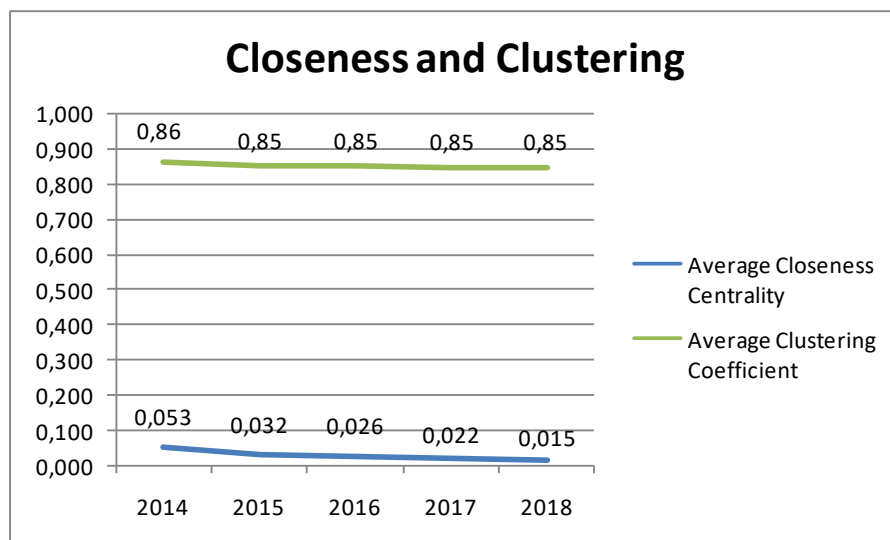


Figura 38. Cercanía y coeficiente de agrupamiento. Elaboración propia en Excel.

En este gráfico vemos la media de cercanía y la media del coeficiente de agrupamiento durante los años. La tendencia que nos muestran estos dos parámetros es muy similar, una tendencia bastante constante en el tiempo con una ligera disminución a medida que pasa el tiempo. Los dos valores se evalúan dentro del mismo rango, que va desde 0 hasta 1.

En cuanto a la media de cercanía vemos que su valor es prácticamente 0. Es decir, que la mayoría de las personas que forman la red están directamente conectadas o muy cercanas al resto.

Para la media del coeficiente de agrupación vemos que el valor es bastante alto, con una media de 0'85 para los 5 años estudiados, siendo el 1 el valor máximo. Esto nos indica, como hemos comentado anteriormente, que en general los coautores de cada persona que forma la red también trabajan entre ellos, o dicho de otra manera: ¡@s amig@s de mis amig@s son también mis amig@s.

El resultado de estos dos valores, indica que se trata de una red que, a nivel global, tiende a ser más colaborativa que no jerárquica.

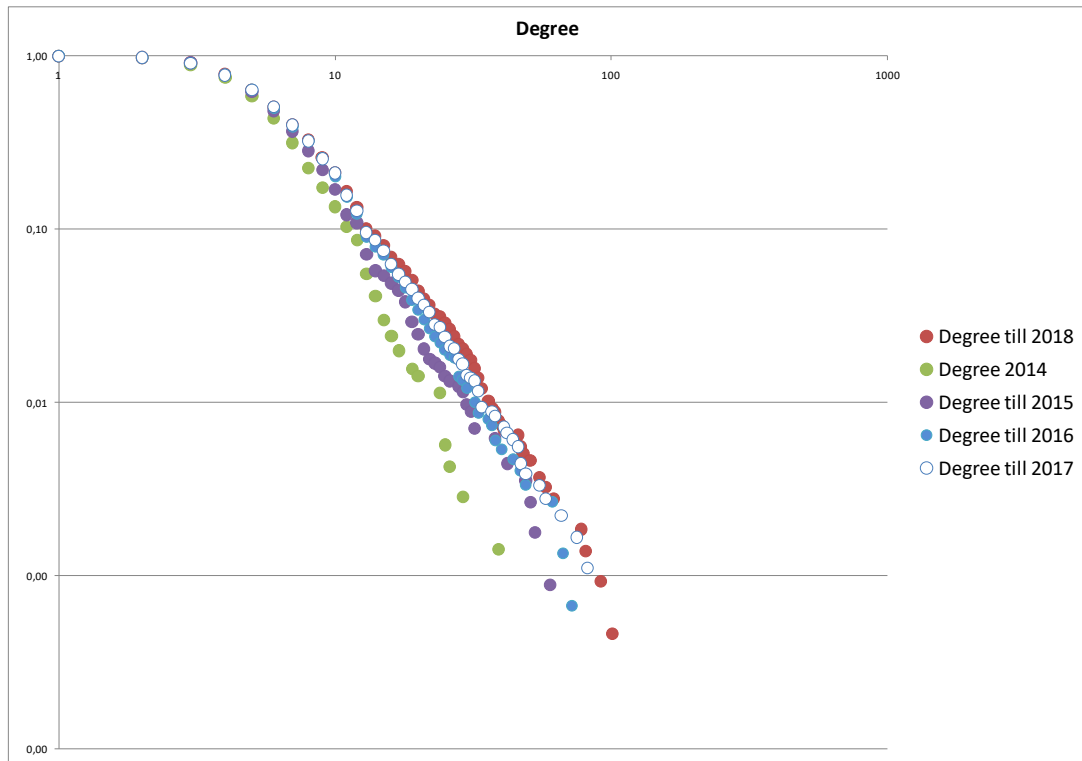
4.2 ESTRUCTURA GLOBAL DE LA RED

Teniendo ya una base de qué tipo de estructuras pueden presentarse en las redes, y una vez analizadas las principales métricas del análisis de la red de coautorías de artículos científicos, procedemos a analizar la estructura global de la red.

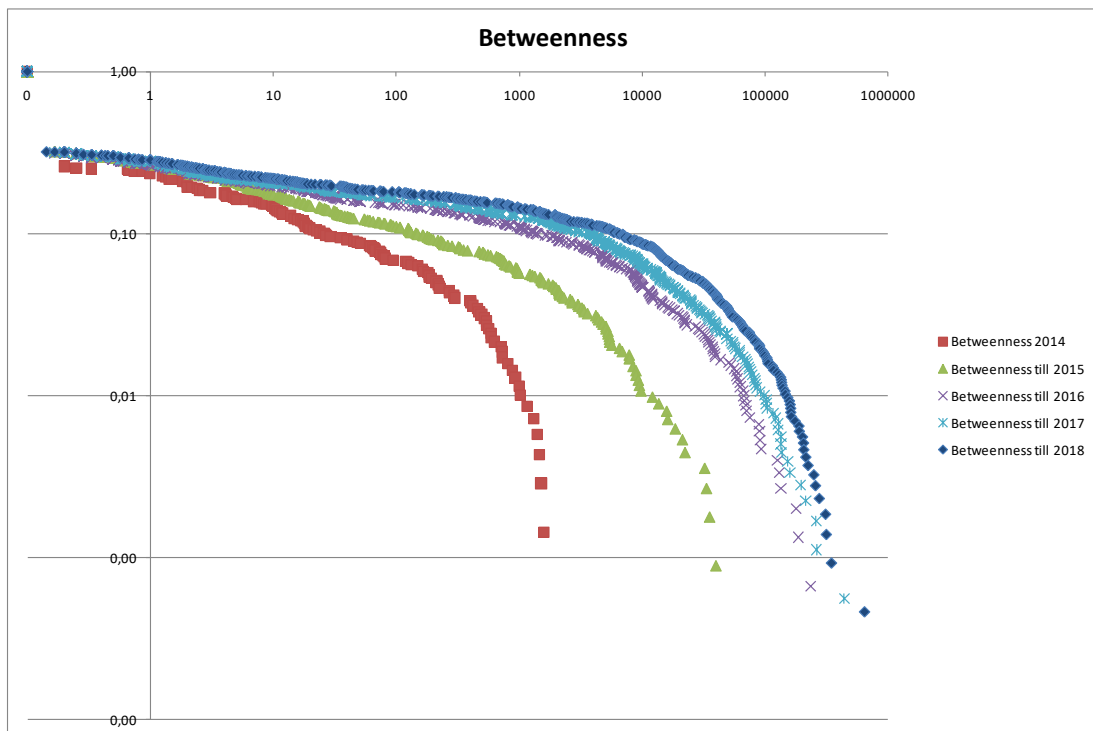
En este sentido, una forma de aproximarnos a la estructura global de la red es a partir de la firma estadística que generan las distribuciones de probabilidad de las diferentes medidas de centralidad. Como se ha mencionado con anterioridad, la distribución de probabilidad (sea histograma o histograma acumulado), especialmente la de grado, es una herramienta poderosa a la hora de determinar si nuestra red es de tipo aleatorio, o bien de tipo libre de escala, aspecto este último que implica (a) la presencia de “hubs” o nodos muy enlazados, los cuales dominan la conectividad global del sistema, y (b) mecanismos de generación de la red no triviales (i.e., “el rico se hace más rico”).

Los resultados del análisis de las distribuciones de probabilidad acumulada de las diferentes medidas de centralidad (i.e., "Degree", "Betweenness", "Closeness", "Eigenvector Centrality", "PageRank" y "Clustering") se muestran a continuación.

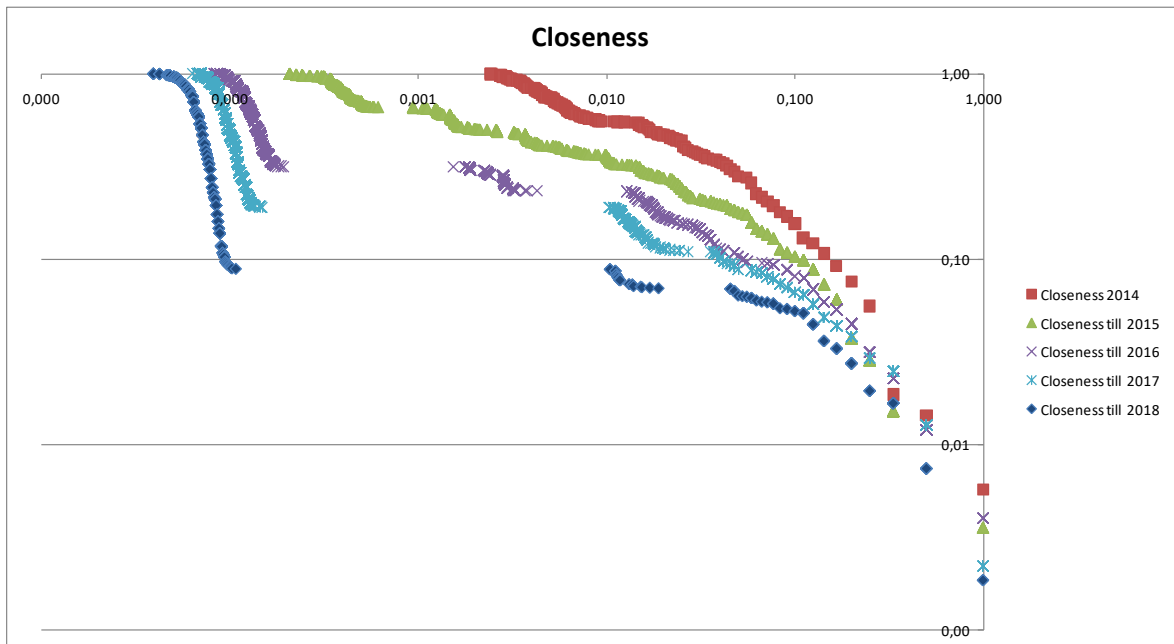
a)



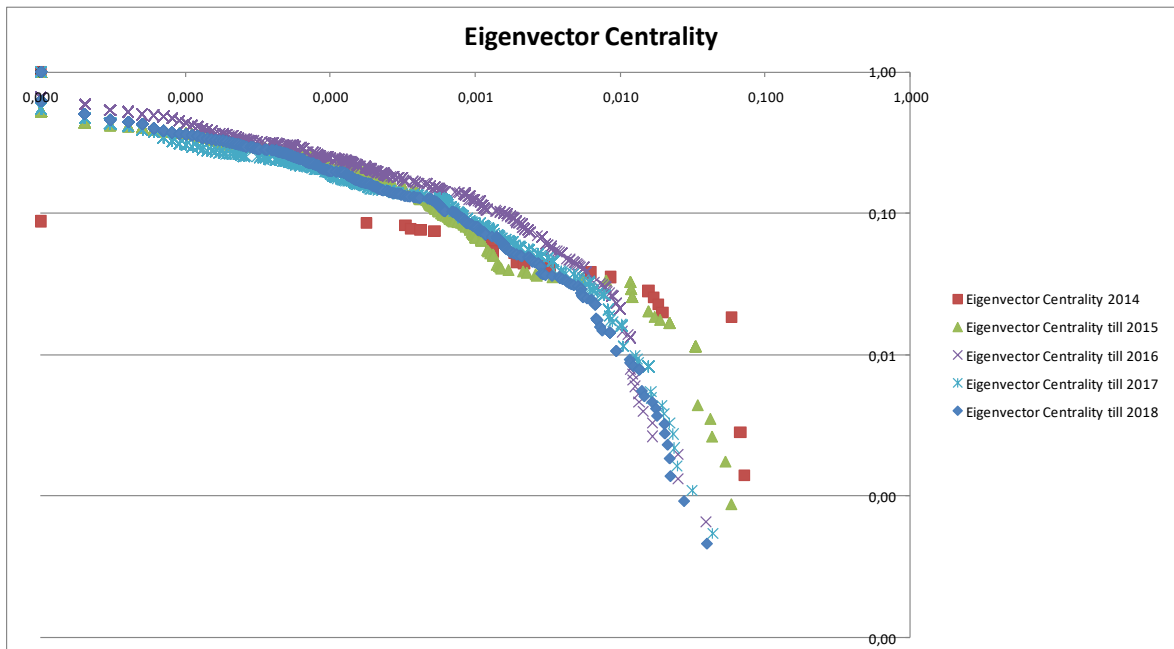
b)



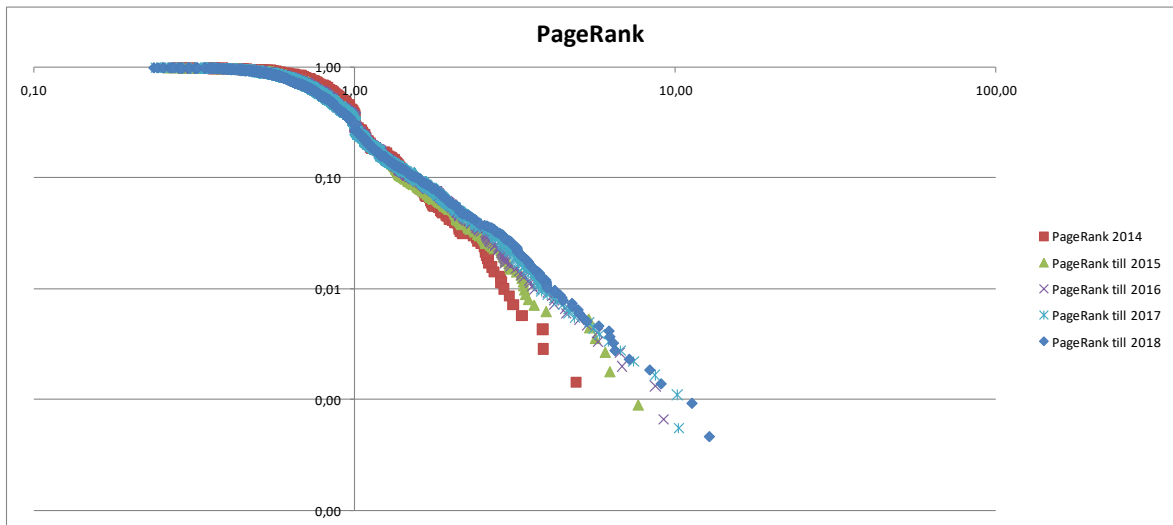
c)



d)



e)



f)

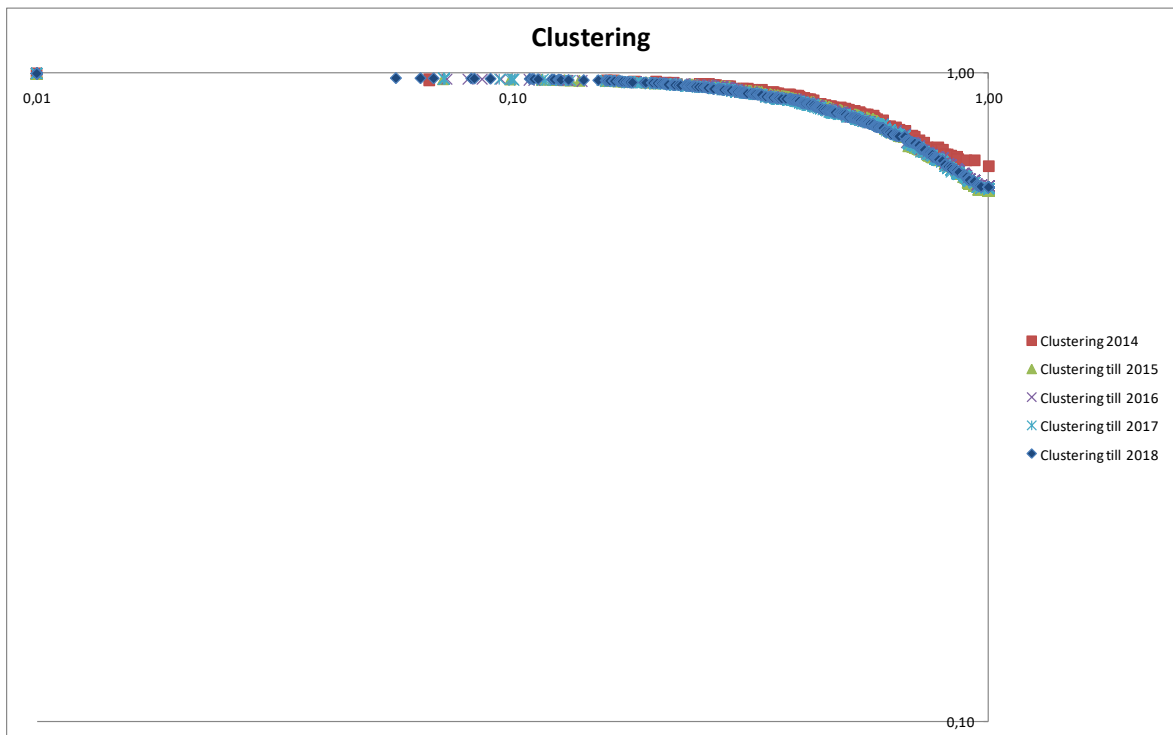


Figura 39. Distribuciones de probabilidad acumulada para las medidas de centralidad (a) "Degree", (b) "Betweenness", (c) "Closeness", (d) "EigenvectorCentrality", (e) "PageRank" y (f) coeficiente de "Clustering".

Como podemos observar, las únicas distribuciones de probabilidad que ofrecen una “cola larga” y que podrían sugerir una función de potencia característica de las redes libres de escala son las correspondientes a “degree” y “PageRank”. Para verificar la bondad del ajuste de estas funciones, aplicamos metodologías estadísticas contrastadas basadas en técnicas de muestreo aleatorio (Clauset et al., 2009). Los resultados, mostrados en la Tabla 6, permiten determinar los siguientes valores:

- **Alpha:** exponente de la función potencial ajustada (pendiente de la función según ecuaciones 9 y 10).
- **Error:** error en la determinación de Alpha.
- **Xmin:** valor mínimo a partir del cual se puede ajustar la “cola” de la función potencial.
- **N(tail):** número de observaciones en la “cola” anteriormente citada.
- **N(tot):** número total de observaciones en la función.
- **N(tail)/N(tot):** porcentaje de observaciones en la “cola” de la función. Este resultado permite determinar la importancia del ajuste en función del porcentaje de “cola” correctamente ajustado. Como podemos observar, para “degree” nos movemos entre el 16% y el 44% de los valores observados. La variabilidad es mayor en el caso del “PageRank” (3% y 27%).
- Para un grupo de funciones de cola larga característico (i.e., **powerlaw con cut-off, lognormal, exponencial y stretched exponential**), las columnas siguientes muestran la bondad del ajuste a partir del estadístico R según Clauset et al. (2009) y el correspondiente p-value. Valores de R positivos favorecen la función potencial, pero para valores de $p < 0.05$, estos resultados no son fiables.

	Año	alpha	error	xmin	n(tail)	n(tot)	n(tail)/n(tot)
Degree	2014	3,93	0.16	6	307	702	0,44
	2015	3,88	0,16	8	320	1124	0,28
	2016	3,85	0,14	9	385	1500	0,26
	2017	3,61	0,15	11	284	1800	0,16
	2018	3,48	0,1	9	560	2154	0,26
PageRank	2014	4,22	0,3	1,29	109	702	0,16
	2015	3,78	0,16	1,03	306	1124	0,27
	2016	3,37	0,13	1,09	323	1500	0,22
	2017	4,04	0,39	2,65	60	1800	0,03
	2018	3,17	0,1	1,12	426	2154	0,20

	Año	PL + cutoff		Lognormal		Exponential		Stretchedexp.	
		R	p	R	p	R	p	R	p
Degree	2014	-0,279	0,714	0,237	0,812	2,707	0,006	0,956	0,338
	2015	0,0005	0,999	1,14	0,25	4,08	6,11	2,32	0,02
	2016	0,003	0,993	-2,18	0,02	4,69	2,6	2,94	0,003
	2017	0,0014	0,998	-2,08	0,03	4,06	4,9	2,21	0,02
	2018	0,005	0,993	-2,01	0,04	5,35	8,68	2,87	0,004
PageRank	2014	-0,85	0,22	-0,681	0,49	0,312	0,754	-0,699	0,48
	2015	-0,027	0,97	-1,72	0,08	3,29	0,00009	1,38	0,16
	2016	-0,928	0,18	-0,649	0,516	2,15	0,03	-0,43	0,66
	2017	-0,214	0,82	-0,019	0,98	1,47	0,14	0,34	0,72
	2018	-1,09	0,11	-0,7	0,48	2,68	0,007	-0,45	0,65

Tabla 6. Resultados obtenidos de la aplicación Powerlaw²⁰ de Jupyter notebooks²¹ para degree y PageRank.

Como podemos ver, los resultados son poco concluyentes excepto en el hecho de descartar la función exponencial en la mayoría de los casos y presentar la función lognormal como candidata para los años 2016 a 2018 en el caso del “degree”, y lognormal en 2015 para PageRank. A pesar de ello, la dominancia de la conectividad de la red por “hubs” altamente conectados se hace patente al observar la disminución del exponente Alpha a medida que pasan los años, hecho que “aplana” la cola de la distribución y hace que la probabilidad de encontrar un nodo muy enlazado sea cada vez mayor.

²⁰<https://github.com/jeffalstott/powerlaw>

²¹<https://jupyter.org/>

Una de las preguntas que nos hacíamos al principio de este trabajo era si esta red se podía considerar un mundo pequeño. Recordemos antes, dos de las propiedades fundamentales de las redes de mundo pequeño. Por un lado, este tipo de red implica que cualquiera de los nodos pertenecientes a ésta, se comuniquen a cualquier otro nodo de la red a través de un camino de nodos intermedios relativamente corto. Por otro lado, este tipo de red posee valores de coeficiente de agrupación bastante altos, ya que este valor nos indica como de conectados están nuestros vecinos o en este caso los coautores del artículo. Esta característica es fácil de comprobar, ya que como hemos visto, la media de este parámetro para los cinco años analizados es de 0.85, siendo 1 el valor máximo que puede alcanzar. Por lo tanto, podemos decir que cumple con este requisito. En nuestro caso, la longitud de camino característica debería tener el mismo orden de magnitud que en el caso aleatoria, mientras que el coeficiente de clustering debería ser órdenes de magnitud mayor que el de su homóloga aleatoria del mismo tamaño (i.e, número de nodos y aristas). Con estos datos, y para el año final de 2018, la Tabla 7 permite corroborar nuestra hipótesis de que se trata, efectivamente, de un mundo pequeño.

Red	Distancia geodésica media	Coficiente de clustering
Aleatoria	$\ln N = \ln(2154) = 7.67$	$\frac{\langle k \rangle}{N} = \frac{7.5}{2154} = 0.003$
Co-autores UPC	7.62	0.85

Tabla 7. Comparativa de valores para el mundo pequeño.²²

²² Las ecuaciones para el cálculo analítico de los valores de distancia geodésica media y coeficiente de clustering para un grafo aleatorio se pueden encontrar en (Newman, 2010) y referencias en él.

4.3 ESTRUCTURAS LOCALES DE LA RED

Desde un punto de vista global, la red social del PDI del Campus de Terrassa de la UPC pareciera que es de tipo colaborativo, un mundo pequeño dominado por algunos nodos muy productivos. ¿Pero qué ocurre a nivel local? ¿Cómo es la estructura local de la red de colaboraciones de estos nodos tan productivos? ¿Actúan a lo largo de los años con las mismas dinámicas? ¿Estas redes tienen evoluciones estructurales diferentes?

Para responder a estas preguntas, analizamos y comparamos las 10 redes ego con más grado en 2018 más las 17 redes ego analizadas del resto de grupos obtenidos mediante el proceso de detección de grupos o modularidad presentado al inicio del trabajo (Figura 22), para comprobar si existen patrones estructurales similares entre las diferentes redes ego. Estos 27 resultados finales pueden observarse en los anexos a este trabajo. Exceptuando los grupos 20, 21, 22 y 23 (Figura 40) ya que no hemos podido completar el seguimiento.

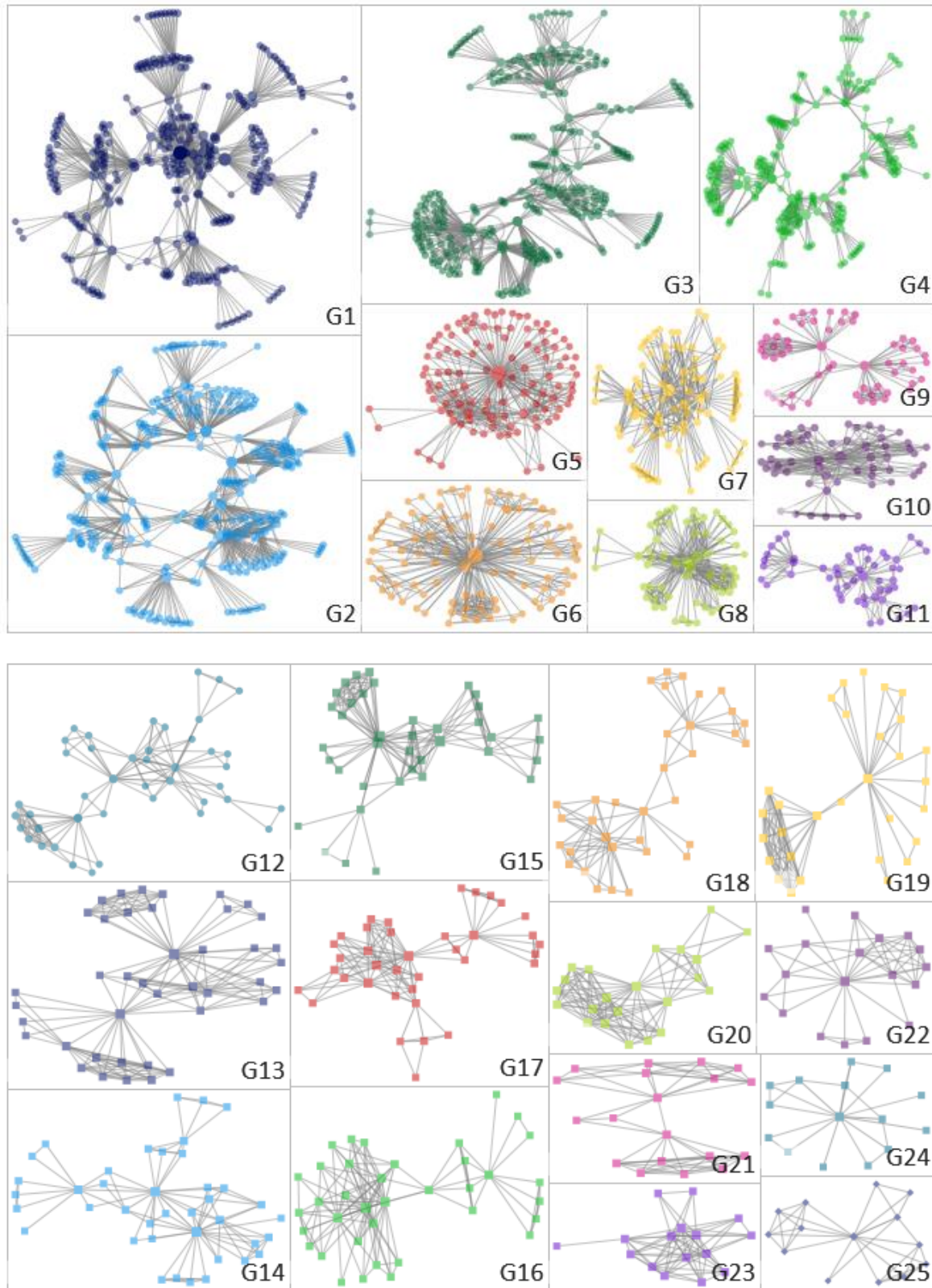


Figura 40. Red de coautorías de artículos científicos de 2018 por grupos. Elaboración propia mediante NodeXL.

El análisis de las diferentes redes ego se ha completado con el de la evaluación de la “distancia” que, a medida que pasan los años, estos nodos más enlazados de cada clúster ganan respecto de su primer vecino de mayor grado. Como presentamos anteriormente en las secciones 3.6 y 3.7, nuestro objetivo aquí es comprobar a la “velocidad” con la que los nodos más conectados aumentan su grado y su PageRank a medida que pasan los años. Para ello, evaluamos la diferencia absoluta entre su grado y su PageRank (i.e., DeltaDeg y DeltaPR) y el de su primer vecino con segundo mayor grado. En este sentido:

- La variación absoluta de grado con el tiempo nos permite detectar aumentos de generación de producción científica diferentes para los diferentes actores y sugerir aquellos aumentos realmente posibles, de aquellos que, a lo mejor, implicarían la utilización del estatus de jerarquía para imponer la aparición de ese actor como autor de una publicación científica.
- La variación absoluta del PageRank con el tiempo nos permite detectar si las coautorías son de calidad. Recordemos que el PageRank genera un valor elevado cuando el grado del nodo analizado es elevado, pero al mismo tiempo tiene en cuenta el valor de grado de sus primeros vecinos: si éste es elevado (i.e., primeros vecinos que a su vez tienen muchas coautorías), el PageRank aumentará; en sentido contrario, aunque mi grado sea elevado, si el grado de mis primeros vecinos es bajo (i.e., primeros vecinos que no tienen coautorías, como estudiantes de doctorado o postdocs de menor nivel académico), el PageRank disminuirá.

Para cuantificar esta variación hemos implementado una metodología esencialmente binaria que permite, grosso modo, evaluar la dinámica de DeltaDeg y DeltaPR para cada actor (Figura 41). La figura 41 muestra la evolución de DeltaDeg y DeltaPR para dos autores diferentes y la tabla 8 el resumen de estas tendencias para los autores de todas las redes ego analizadas. Para cada uno de los segmentos temporales presentados en la Figura 41, se considera el cuantificador “1” si la evolución es de pendiente positiva, “-1” si la evolución es de pendiente negativa, y “0” en cualquier otro caso. Como podemos observar, el actor (a) tendría una evolución de 2 y 3 con respecto a DeltaDeg y DeltaPR, respectivamente, mientras que el actor (b) la tendría de 1 y 2 respectivamente.

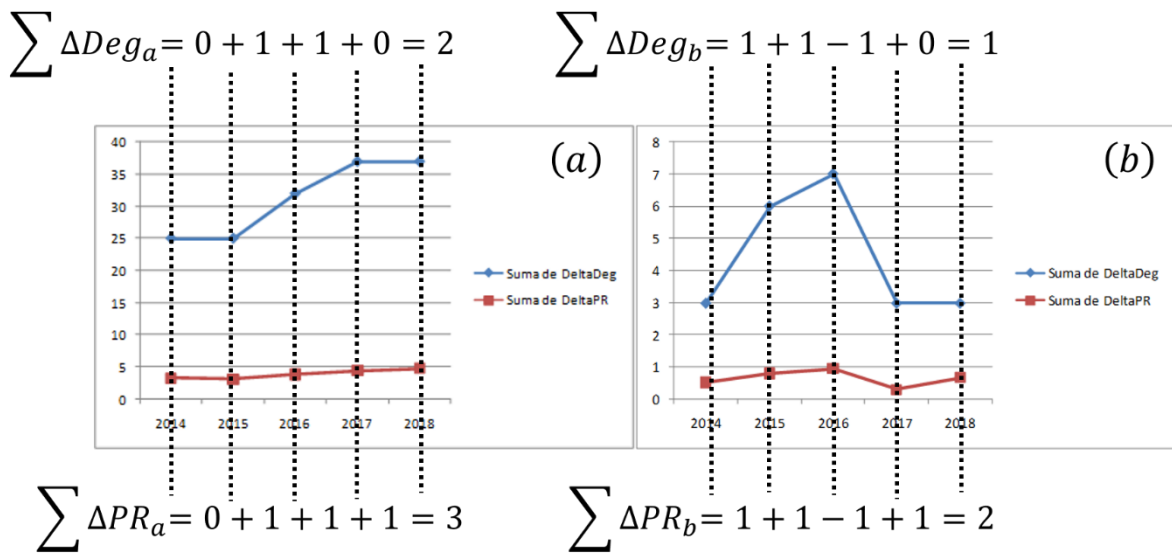


Figura 41. Ejemplo de cálculo del cuantificador de la evolución de DeltaDeg y DeltaPR para dos actores de la red analizada en este TFG.

Esta cuantificación nos permite clasificar los actores más conectados de los diferentes módulos (i.e., grupos, en la figura 40) estudiados mediante tres ejes con los valores de $\sum \Delta Deg$, $\sum \Delta PR$ y grado absoluto en 2018 (Figura 42). Este tercer eje nos permite segregar los diferentes actores no tan sólo en función de cómo rápidamente se diferencian de sus primeros vecinos en grado y PageRank, sino también en cómo han acumulado enlaces en estos pocos años de trabajo analizados. Para unos mismos valores de $\sum \Delta Deg$ y $\sum \Delta PR$, un mayor grado absoluto en 2018 permite establecer diferentes (y mayores) capacidades de trabajo y generación de conocimiento para estos actores, en comparación con los de menor grado.

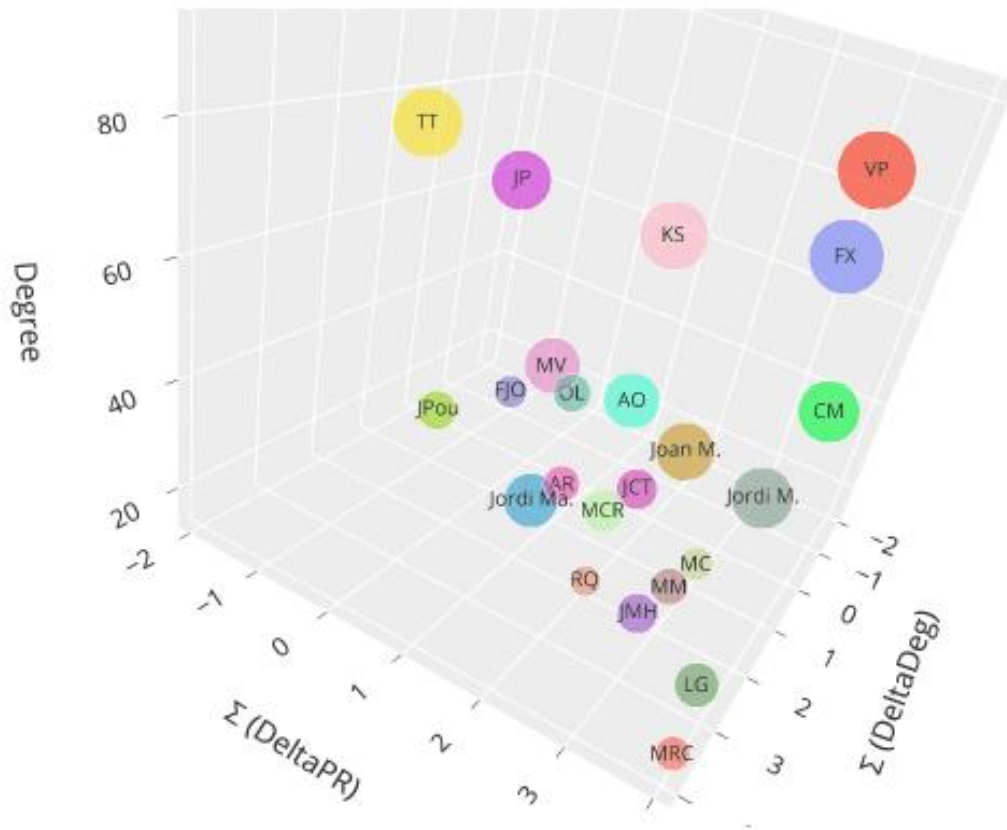


Figura 42 Diagrama de dispersión de los actores analizados referenciados mediante sus iniciales, según $\Sigma\Delta Deg$, $\Sigma\Delta PR$ y grado absoluto en 2018. Elaboración propia mediante Chart Studio Plotly²³.

Esta última variable es la que más se correlaciona con la estructura local de la red, siendo las redes en forma de estrella las más correlacionadas con un perfil de valores altos de $\Sigma\Delta Deg$, $\Sigma\Delta PR$ y grado absoluto en 2018. Por ejemplo, los casos de Jordi Morato (Jordi M.), Fathos Xafa (FX) y Vicenç Puig (VP), se sitúan en este último grupo. Mientras que Martí Rosas Casals (MRC) o Lluís Gil (LG), con los mismos valores de $\Sigma\Delta Deg$ y $\Sigma\Delta PR$, aparecen con un menor grado absoluto en 2018.

Joan Montanya (Joan M.), Kestutis Staliunas (KS), Cristina Masoller (CM), Asensio Oliva (AO), Meritxell Vilaseca (MV) y Jordi Martorell (Jordi Ma.) se encuentran en una zona intermedia, con valores medios de $\Sigma\Delta Deg$, $\Sigma\Delta PR$ y grado absoluto en 2018, todos ellos con una topología estrellada también.

²³ <https://chart-studio.plotly.com/~CGplot/1>

Por la zona baja de la dispersión nos encontramos con nodos que, a pesar de compartir estructura estrellada, tienen capacidades de colaboración menos notables y capacidad de producción menor, como son Jaime Martí Herrero (JMh), María Carmen Riva (MCR), Meritxell Martí (MM) o Ramón Quintanilla (RQ).

NOMBRE	GRUPO	Tend. DeltaDeg					Tend. DeltaPR					GRADO 2018
		1	2	3	4	Suma	1	2	3	4	Suma	
Vicenç Puig	G5	1	1	1	1	4	1	1	1	1	4	101
Cristina Masoller	G3	1	1	1	-1	2	1	1	1	1	4	62
Fatos Xhafa	G6	1	1	1	1	4	1	1	1	1	4	92
Kestutis Staliunas	G3	0	1	1	0	2	-1	1	1	1	2	78
Jaume Pujol	G2	1	-1	-1	-1	-2	1	-1	-1	-1	-2	58
Asensio Oliva	G7	1	1	0	0	2	1	1	0	0	2	51
Tzanko Tzanov	G1	1	1	1	-1	2	1	-1	0	-1	-1	81
Joan Montanya	G8	1	0	1	1	3	1	0	1	1	3	55
Jordi Morato	G1	1	1	1	1	4	1	1	1	1	4	62
Meritxell Vilaseca	G2	1	1	-1	1	2	1	1	-1	0	1	51
Jasmina Casals Terre	G4	1	1	-1	0	1	1	1	-1	1	2	27
Jaime Martí Herrero	G9	1	1	1	0	3	1	1	1	0	3	26
Jordi Martorell	G10	1	1	1	1	4	-1	1	1	1	2	47
Lluís Gil	G11	1	1	1	1	4	1	1	1	1	4	32
Miquel Casals	G12	0	1	-1	1	1	0	1	1	1	3	18
Anna Rius	G13	0	1	-1	1	1	0	1	-1	1	1	21
Oriol Lordan	G14	-1	0	1	-1	-1	-1	1	1	-1	0	23
María Carmen Riva	G15	1	1	0	0	2	1	1	0	0	2	30
Josep Pou	G16	0	0	1	0	1	-1	-1	1	0	-1	24
Meritxell Martí	G17	0	1	0	1	2	1	1	0	1	3	22
Francisco Javier Oliver	G18	-1	1	-1	0	-1	-1	1	-1	0	-1	17
Martí Rosas Casals	G19	1	1	1	1	4	1	1	1	1	4	19
Jose María Baldasano	G20					0					0	19
Alejandra R. Contreras	G21					0					0	12
Leonardo Acho	G22					0					0	19
Fayyaz Ahmad	G23					0					0	15
Ramón Quintanilla	G24	1	1	1	-1	2	1	1	1	-1	2	14

Tabla 8. Tabla resumen de las tendencias de las redes analizadas.

5. CONCLUSIONES

Una vez realizado todo el estudio, procedemos a continuación con las conclusiones finales de la red de coautorías de artículos científicos, la verificación o no de las hipótesis planteadas al inicio del trabajo, y comentarios adicionales acerca del trabajo.

Primeramente, con la discusión de las métricas generales, vemos ciertos comportamientos a destacar. El primero, la importancia de la interdisciplinariedad tanto en la educación como en la investigación. Vemos en esta red de coautorías de artículos científicos que a medida que pasan los años existe una mayor colaboración entre departamentos o ámbitos de estudio, con mayor énfasis en el último año analizado 2018, además de una mayor producción de artículos científicos.

Vemos también un aumento sutil del grado promedio, pero gracias a la distribución de grado en forma de cola larga, sabemos que este aumento no es generalizado, sino que está centrado en ciertos "hubs" o autores con un gran número de conexiones, que pasan a tener aún más (el rico se hace más rico).

Finalmente, con una clara tendencia de aumento en su "betweenness centrality", unos valores muy bajos en cuanto a "closeness centrality" y una distancia geodésica estable en el tiempo, podemos concluir que **en su estructura global**, la red tiene un **comportamiento colaborativo** y no jerárquico.

Por otro lado, siguiendo en su estructura global, y viendo los resultados de la Tabla 7, concluimos también que estamos ante una **red de mundo pequeño de escala libre**, ya que hemos podido detectar algunos "hubs" o autores con un gran número de conexiones, mientras que el resto de nodos o autores (la gran mayoría de la red) poseen pocas conexiones, visible en las distribuciones de cola larga.

5.1 COMPROBACIÓN DE HIPÓTESIS

En cuanto a las hipótesis de jerarquía. Planteamos la primera hipótesis: "Existen aristas con un peso mayor de lo que se podría esperar, por ejemplo, con una colaboración más intensa entre un jefe (de departamento o grupo de investigación) y un subordinado (i.e., alumno de doctorado, postdoc, etc.)". Para ello, tenemos la siguiente figura:

Vertex 1	Vertex 2	Edge Weight
Lehmkuhl, Oriol	Oliva, Asensio	26
Candela, Jose Ignacio	Rodriguez, Pedro	23
Puig, Vicenc	Nejjari, Fatiha	22
Luna, Alvaro	Rodriguez, Pedro	22
Xhafa, Fatos	Barolli, Leonard	21
Casals, Miquel	Gangolells, Marta	21
Rotondo, Damiano	Puig, Vicenc	21
Rotondo, Damiano	Nejjari, Fatiha	19
Roncero, Maria Blanca	Vidal, Teresa	19
Roncero, Maria Blanca	Valls, Cristina	19
Herrero-Simon, Ramon	Staliunas, Kestutis	19
Botey, Muriel	Staliunas, Kestutis	19
Candela, Jose Ignacio	Luna, Alvaro	18
Herrero-Simon, Ramon	Botey, Muriel	18
Francesko, Antonio	Tzanov, Tzanko	18

Figura 43. Las 15 colaboraciones más intensas para los años desde 2014 hasta 2018.

Vemos aquí como existen varias aristas, con intensas colaboraciones entre autores, que destacan también por pertenecer al mismo grupo cada par de autores reflejados. Por lo tanto, confirmamos esta hipótesis.

Planteamos la segunda hipótesis: "Destaca un nodo (i.e., jefe de departamento o PDI con antigüedad notable) que aumenta su grado y su coeficiente de agrupación de manera no lineal, mientras que su nodo (subordinado) repite el mismo patrón con un cierto retraso".

En este caso, hemos visto que el aumento de grado a medida que pasan los años es claro en estos "hubs", y por consiguiente, de sus subordinados. Pero el aumento de su coeficiente de agrupamiento no se cumple, de hecho, es lo contrario. En la pestaña de "Ego Networks DEGREE"²⁴ por ejemplo, eligiendo dos autores considerados como "hubs" como pueden ser Fatos Xhafa, vemos que su coeficiente de "clustering" pasa de ser de 0'3 en 2014 a 0'06 en 2018, y Asensio Oliva, pasa de tener un coeficiente de agrupamiento de 0'21 en 2014 a 0'13 en 2018. Ya que este coeficiente nos indica la tendencia que tienen los nodos a agruparse en forma de triángulo en la red, en el momento en que este coeficiente disminuye, vemos que este hecho no es así, y lo que realmente pasa es que estos "hubs" o jefes de departamento, colaboran cada vez más con nuevos autores (que entre ellos no se relacionan). De esta manera aumentan su grado, pero disminuyen su coeficiente de agrupamiento.

En cuanto a las hipótesis de colaboración. Planteamos la primera hipótesis: "A medida que pasa el tiempo, el número de componentes conexas disminuye para cada año acumulado." Se cumple como podemos ver en la Figura 30.

Planteamos también que algunos nodos aumentan su "betweenness". Esto también es cierto y reflejado en la Figura 37.

En la tercera, dijimos que el grado y el "clustering coefficient" de los nodos colaborativos aumentan de manera lineal con el tiempo, pero con valores absolutos que pueden ser órdenes de magnitud por debajo del caso jerárquico. En este caso sí que vemos que el aumento de grado se refleja en toda la red, como vemos en la Figura 32, pero en cuanto a su "clustering coefficient" no podemos estar seguros, ya que si escogemos a un nodo como colaborativo por su valor alto en grado por ejemplo, como hemos visto anteriormente, no se cumple este aumento.

²⁴ <https://drive.google.com/file/d/1U10a4KG7HkFAyj7CEk3Hcei6tnGs9lYZ/view?usp=sharing>

Por último, tenemos que el valor de cercanía es muy bajo, con tendencia a disminuir. Este hecho si se cumple, y se refleja en la Figura 38.

Con esta verificación de las hipótesis realizadas, podemos concluir que en una visión global, la red se comporta de manera colaborativa, pero en una visión local, centrándonos en el funcionamiento interno de los departamentos, podemos intuir cierta jerarquía a la hora de gestionar la producción de artículos científicos.

5.2 COMENTARIOS ADICIONALES

A la hora de estudiar las tendencias de evolución de ciertos autores en las redes ego analizadas, y centrándonos en los comportamientos locales de la red, nos hemos dado cuenta que algunos autores presentan algunas irregularidades y creemos que pueden estar justificadas de algún modo, sin ser éstas un reflejo del funcionamiento de la red. Por ello, creemos que profundizar en estas irregularidades, mediante posibles futuras entrevistas a dichos autores, podría darnos un conocimiento mayor en cuanto los motivos de estos cambios de comportamiento, y a su vez, un conocimiento mayor del funcionamiento interno del departamento al cual pertenecen, además de una verificación de los análisis realizados en este trabajo.

6. BIBLIOGRAFÍA

- Barabási, A.-L., & Albert, R. (1999). Emergence of Scaling in Random Networks. *Science*, 286, 509–512. [internal-pdf://emergencerandom_science-3540984321/EmergenceRandom_Science.pdf](#)
- Bavelas, A. (1950). Communication Patterns in Task-Oriented Groups. *The Journal of the Acoustical Society of America*, 22(6), 725–730. <https://doi.org/10.1121/1.1906679>
- Bettencourt, L. M. A., Kaiser, D. I., & Kaur, J. (2009). Scientific discovery and topological transitions in collaboration networks. *Journal of Informetrics*, 3(3), 210–221. <https://doi.org/10.1016/j.joi.2009.03.001>
- Bonacich, P., & Lu, P. (2012). *Introduction to Mathematical Sociology*. Princeton University Press.
- Chen, H.-Q., Wang, X., He, L., Chen, P., Wan, Y., Yang, L., & Jiang, S. (2016). Chinese energy and fuels research priorities and trend: A bibliometric analysis. *Renewable and Sustainable Energy Reviews*, 58, 966–975. <https://doi.org/10.1016/j.rser.2015.12.239>
- Clauset, A., Shalizi, C. R., & Newman, M. E. J. (2009). Power-Law Distributions in Empirical Data. *SIAM Review*, 51(4), 661–703. <https://doi.org/10.1137/070710111>
- Freeman, L. C., Borgatti, S. P., & White, D. R. (1991). Centrality in valued graphs: A measure of betweenness based on network flow. *Soc. Networks*, 13, 141–154.
- Girvan, M., & Newman, M. E. J. (2001). Community structure in social and biological networks. *Proc. Natl. Acad. Sci. USA*, 99(12), 7821–7826.
- Mitchell, M. (2009). *Complexity: A Guided Tour*. Oxford University Press. <https://books.google.com/books?id=bbN-6aDFrAC&pgis=1>
- Newman, M. E. J. (2010). *Networks. An introduction*. Oxford University Press.
- Newman, M. E. J., Strogatz, S. H., & Watts, D. J. (2002). Random graphs models of social networks. *Proceedings of the National Academy of Science of the United States of America*, 99, 2566–2572.



- Small, H. (1973). Co-citation in the scientific literature: a new measure of the relationship between two documents. *Journal of the American Society for Information Science*, 24, 265–269.
- Small, H. G. (1978). Cited Documents as Concept Symbols. *Social Studies of Science*, 8(3), 327–340. <https://doi.org/10.1177/030631277800800305>
- Watts, D. J., & Strogatz, S. H. (1998). Collective dynamics of “small-world” networks. *Nature*, 393(6684), 440–442. <https://doi.org/10.1038/30918>

7. ANEXOS

7.1 Topología de las 10 redes ego con más grado en 2018

7.1.1 "Ego network" de Vicenç Puig en el grupo G5.

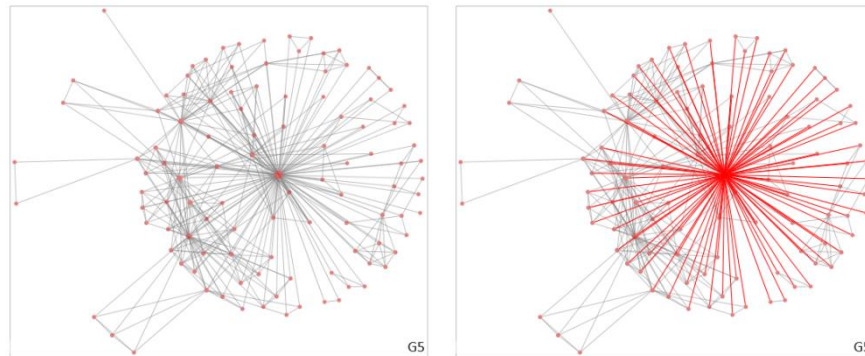


Figura 44. Grupo G5 a la izquierda y la red ego de Vicenç Puig iluminada a la derecha. Elaboración propia en NodeXL.

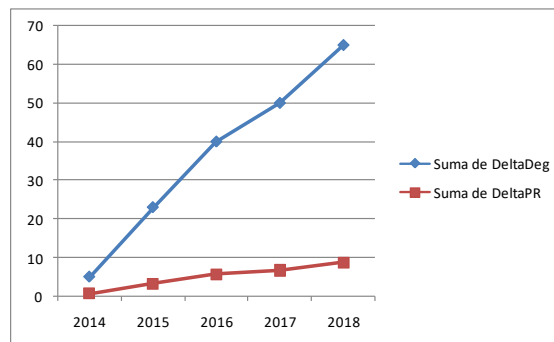


Figura 45. Gráfico para la visualización de la diferencia de grado y de PageRank entre el nodo focal Vicenç Puig y su segundo nodo para los años desde 2014 hasta 2018.

7.1.2 "Ego network" de Fatos Xhafa en el grupo G6.

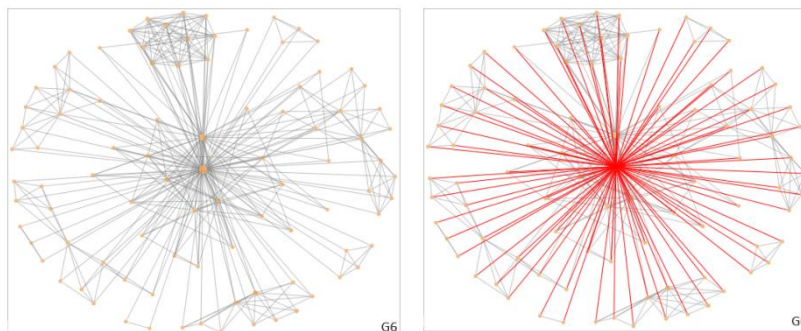


Figura 46. Grupo G6 a la izquierda y la red ego de Fatos Xhafa iluminada a la derecha. Elaboración propia en NodeXL.

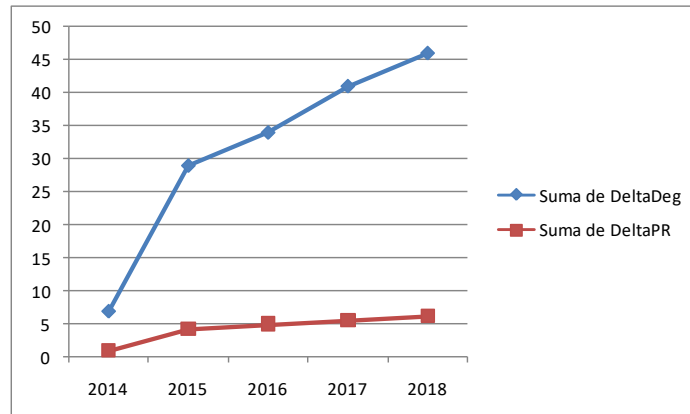


Figura 47. Gráfico para la visualización de la diferencia de grado y de PageRank entre el nodo focal Fatos Xhafa y su segundo nodo para los años desde 2014 hasta 2018.

7.1.3 "Ego network" de Tzanko Tzanov en el grupo G1

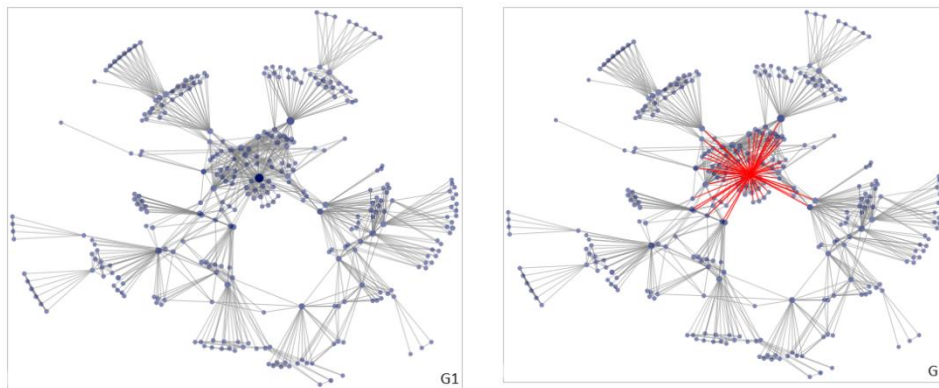


Figura 48. Grupo G1 a la izquierda y la red ego de Tzanko Tzanov iluminada a la derecha. Elaboración propia en NodeXL.

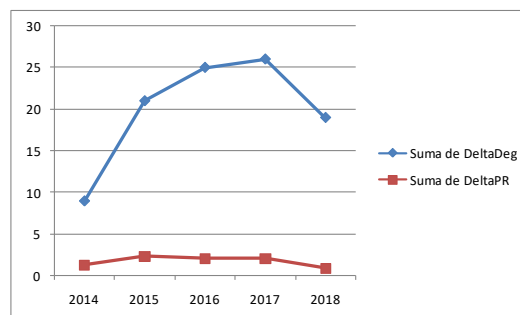


Figura 49. Gráfico para la visualización de la diferencia de grado y de PageRank entre el nodo focal Tzanko Tzanov y su segundo nodo para los años desde 2014 hasta 2018.

7.1.4 "Ego network" de Kestutis Staliunas en el grupo G3.

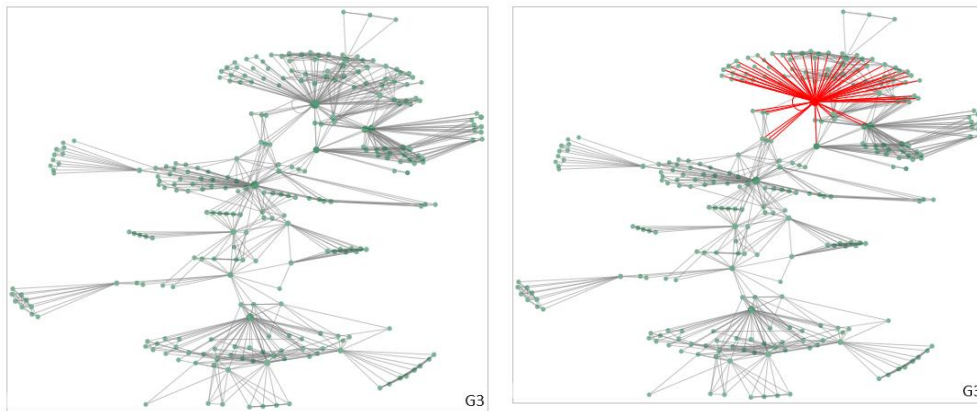


Figura 50. Grupo G3 a la izquierda y la red ego de Kestutis Staliunas iluminada a la derecha. Elaboración propia en NodeXL.

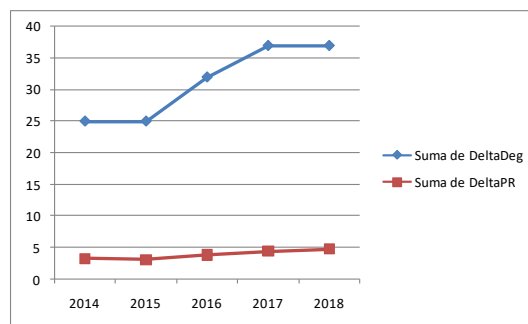


Figura 51. Gráfico para la visualización de la diferencia de grado y de PageRank entre el nodo focal Kestutis Staliunas y su segundo nodo para los años desde 2014 hasta 2018.

7.1.5 "Ego network" de Cristina Masoller en el grupo G3.

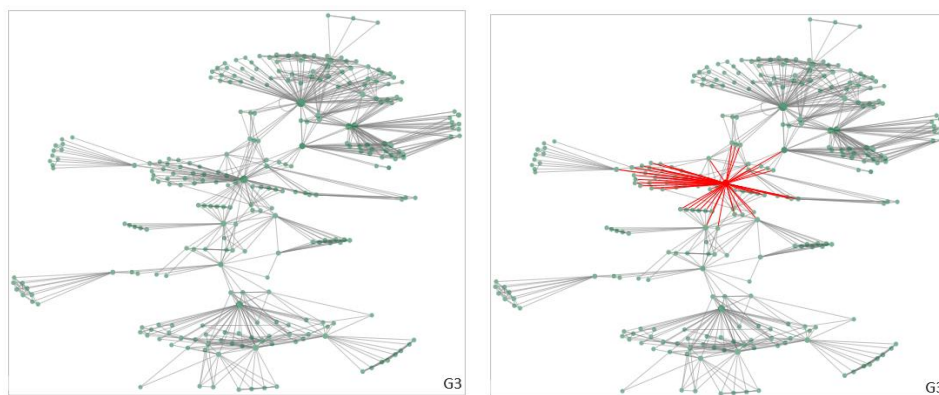


Figura 52. Grupo G3 a la izquierda y la red ego de Cristina Masoller iluminada a la derecha. Elaboración propia en NodeXL.

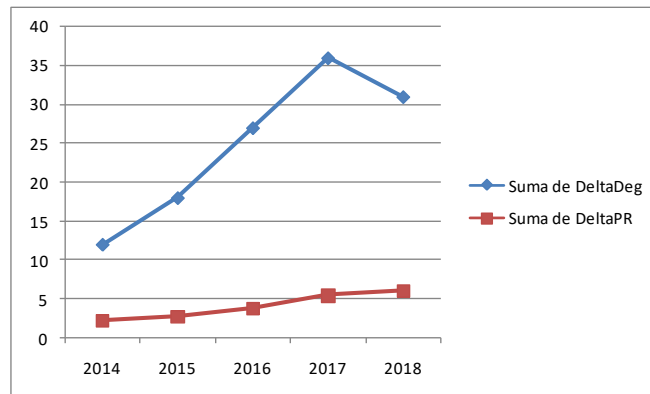


Figura 53. Gráfico para la visualización de la diferencia de grado y de PageRank entre el nodo focal Cristina Masoller y su segundo nodo para los años desde 2014 hasta 2018.

7.1.6 "Ego network" de Jordi Morato en el grupo G1.

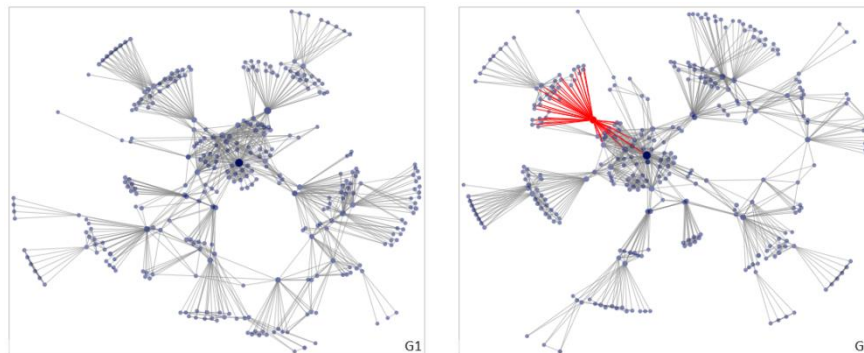


Figura 54. Grupo G1 a la izquierda y la red ego de Jordi Morato iluminada a la derecha. Elaboración propia en NodeXL.

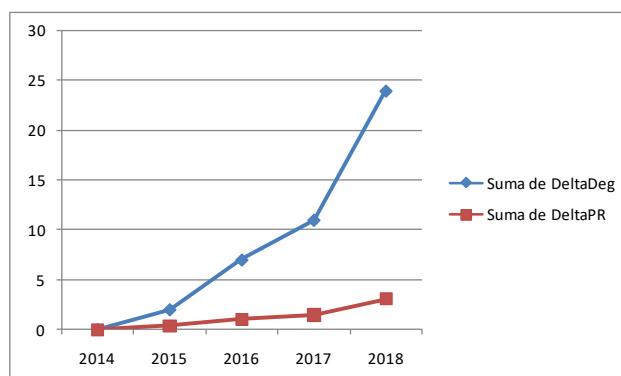


Figura 55. Gráfico para la visualización de la diferencia de grado y de PageRank entre el nodo focal Jordi Morato y su segundo nodo para los años desde 2014 hasta 2018.

7.1.7 "Ego network" de Jaume Pujol en el grupo G2.

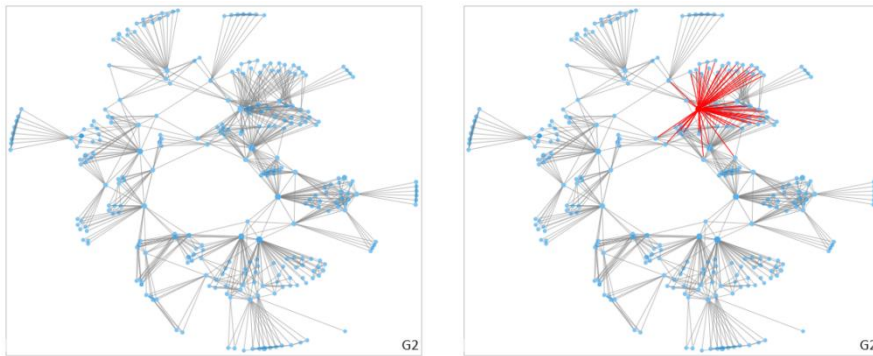


Figura 56. Grupo G2 a la izquierda y la red ego de Jaume Pujol iluminada a la derecha. Elaboración propia en NodeXL.

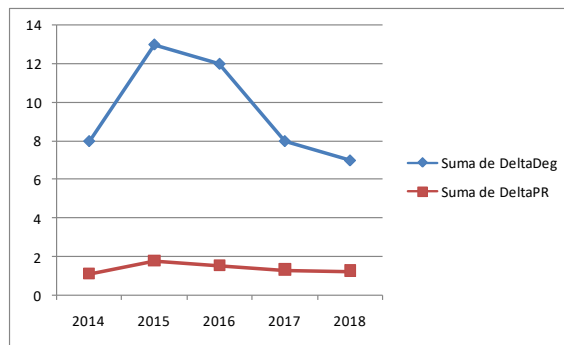


Figura 57. Gráfico para la visualización de la diferencia de grado y de PageRank entre el nodo focal Jaume Pujol y su segundo nodo para los años desde 2014 hasta 2018.

7.1.8 "Ego network" de Joan Montanya en el grupo G8.

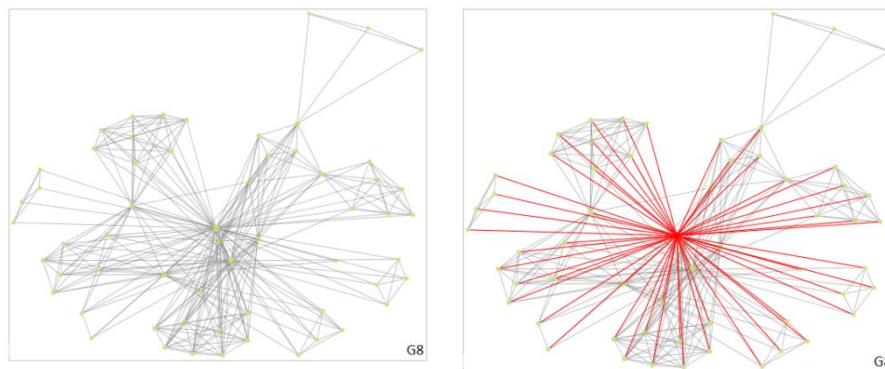


Figura 58. Grupo G8 a la izquierda y red ego de Joan Montanya iluminada a la derecha. Elaboración propia en NodeXL.

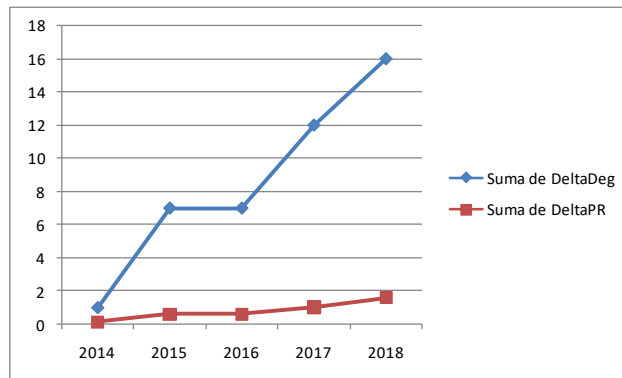


Figura 59. Gráfico para la visualización de la diferencia de grado y de PageRank entre el nodo focal Joan Montanya y su segundo nodo para los años desde 2014 hasta 2018.

7.1.9 "Ego network" de Oliva Asensio en el grupo G7.

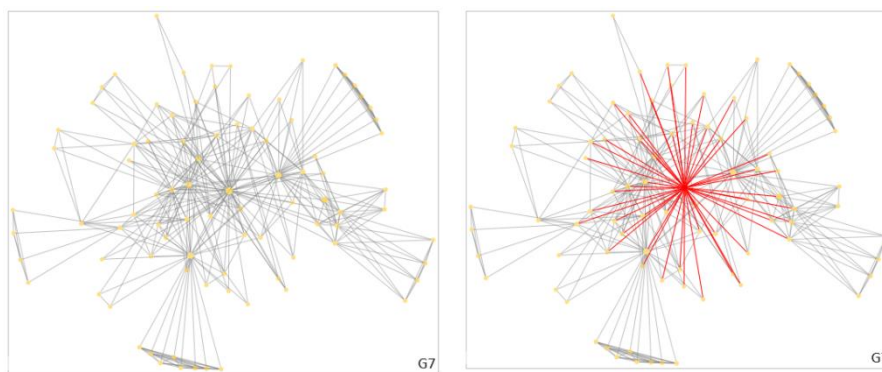


Figura 60. Grupo G7 a la izquierda y la red ego de Oliva Asensio a la derecha. Elaboración propia en NodeXL.

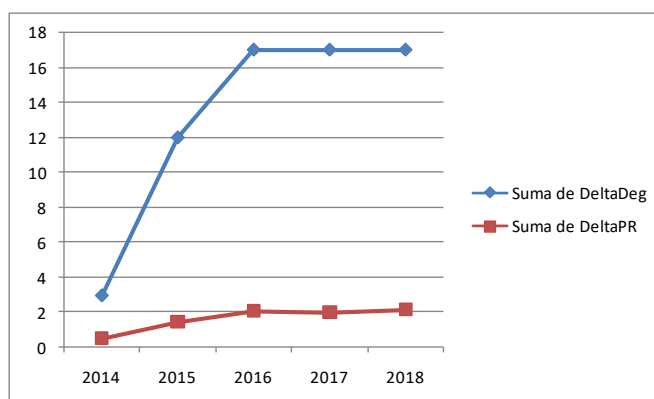


Figura 61. Gráfico para la visualización de la diferencia de grado y de PageRank entre el nodo focal Oliva Asensio y su segundo nodo para los años desde 2014 hasta 2018.

7.1.10 "Ego network" de Meritxell Vilaseca en el grupo G2.

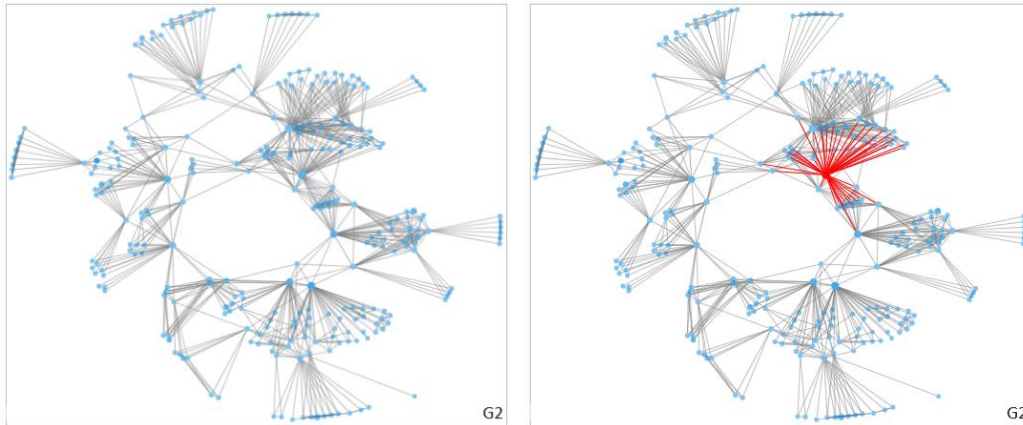


Figura 62. Grupo G2 a la izquierda y la red ego de Meritxell Vilaseca iluminada a la derecha. Elaboración propia en NodeXL.

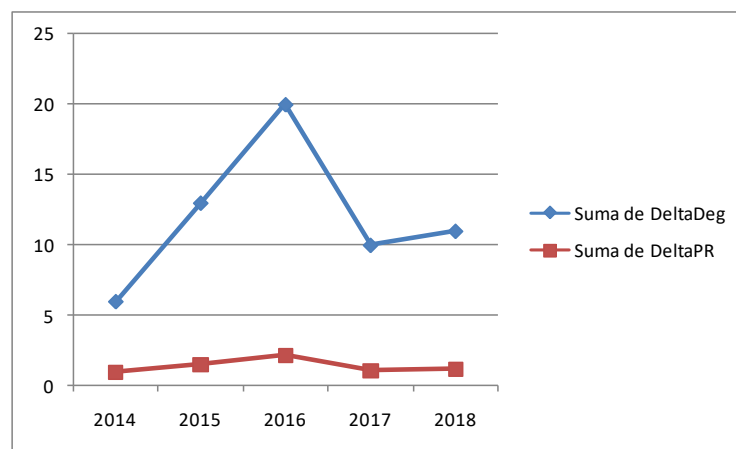


Figura 63. Gráfico para la visualización de la diferencia de grado y de PageRank entre el nodo focal Meritxell Vilaseca y su segundo nodo para los años desde 2014 hasta 2018.

7.2 Topología del resto de grupos de 2018

7.2.1 "Ego network" de Jasmina Casals Terre en el grupo G4

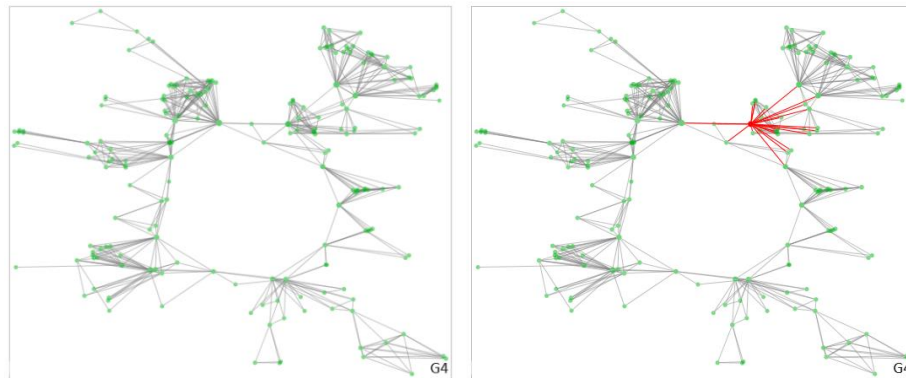


Figura 64. Grupo G4 a la izquierda y la red ego Jasmina Casals Terre iluminada a la derecha. Elaboración propia en NodeXL.

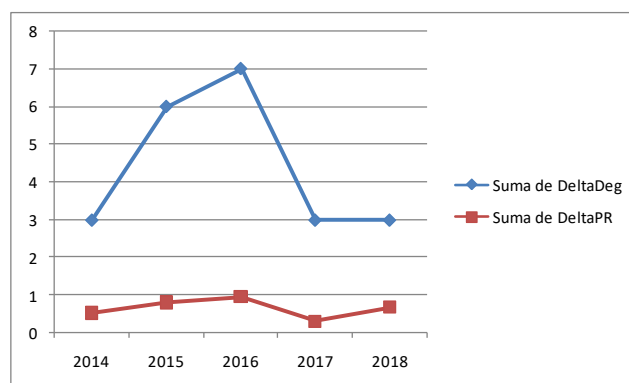


Figura 65. Gráfico para la visualización de la diferencia de grado y de PageRank entre el nodo focal Jasmina Casals Terre y su segundo nodo para los años desde 2014 hasta 2018.

7.2.2 "Ego network" de Jaime Martí Herrero en el grupo G9

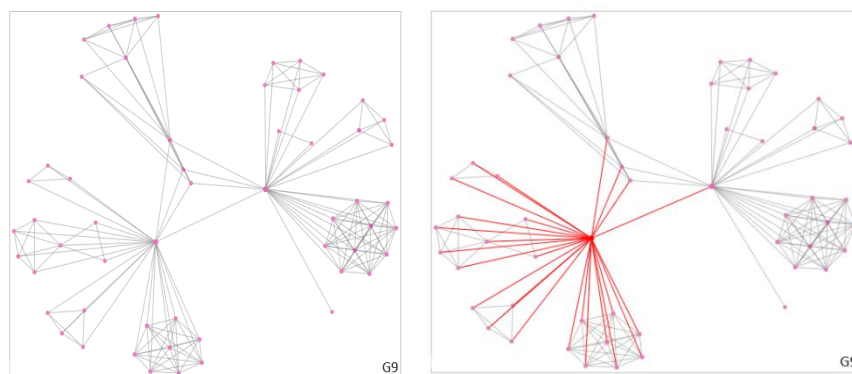


Figura 66. Grupo G9 a la izquierda y la red ego Jaime Martí Herrero iluminada a la derecha. Elaboración propia en NodeXL.

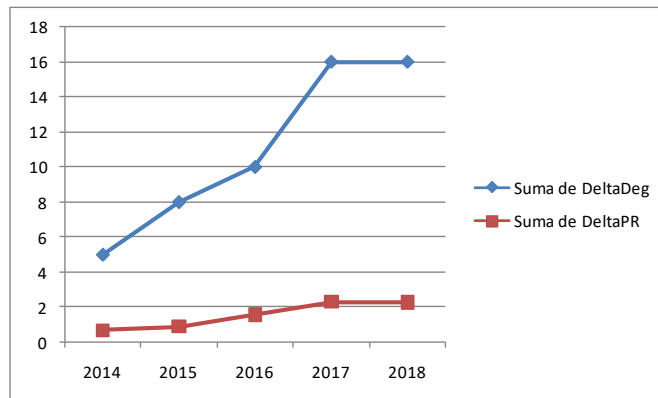


Figura 67. Gráfico para la visualización de la diferencia de grado y de PageRank entre el nodo focal Jaime Martí Herrero y su segundo nodo para los años desde 2014 hasta 2018.

7.2.3 "Ego network" de Jordi Martorell en el grupo G10

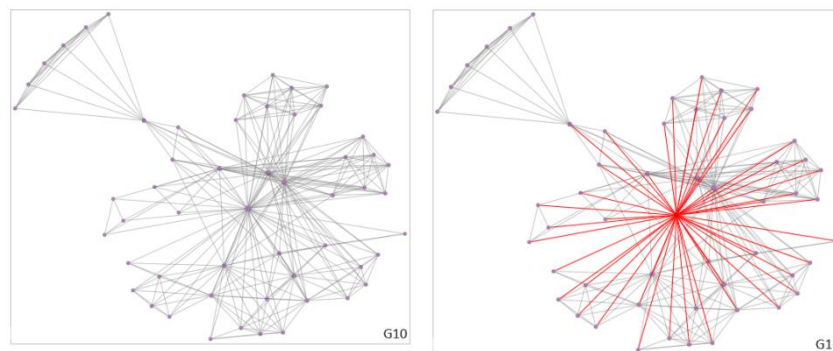


Figura 68. Grupo G10 a la izquierda y la red ego Jordi Martorell iluminada a la derecha. Elaboración propia en NodeXL.

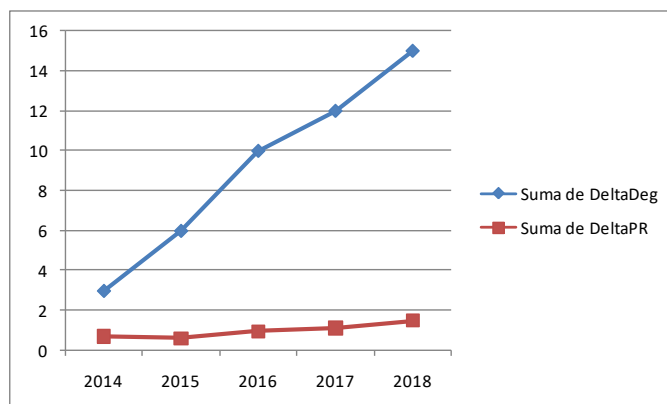


Figura 69. Gráfico para la visualización de la diferencia de grado y de PageRank entre el nodo Jordi Martorell y su segundo nodo para los años desde 2014 hasta 2018.

7.2.4 "Ego network" de Lluís Gil en el grupo G11

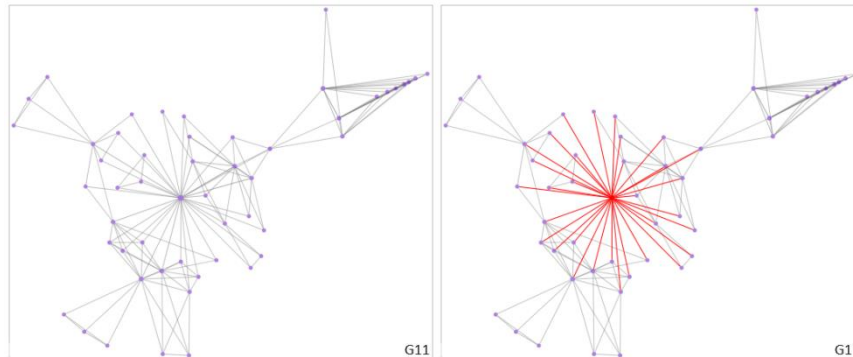


Figura 70. Grupo G11 a la izquierda y la red ego Lluís Gil iluminada a la derecha.
 Elaboración propia en NodeXL.

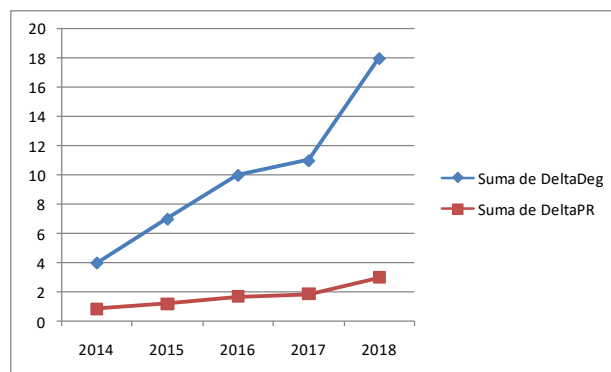


Figura 71. Gráfico para la visualización de la diferencia de grado y de PageRank entre el nodo Lluís Gil y su segundo nodo para los años desde 2014 hasta 2018.

7.2.5 "Ego network" de Miquel Casals en el grupo G12

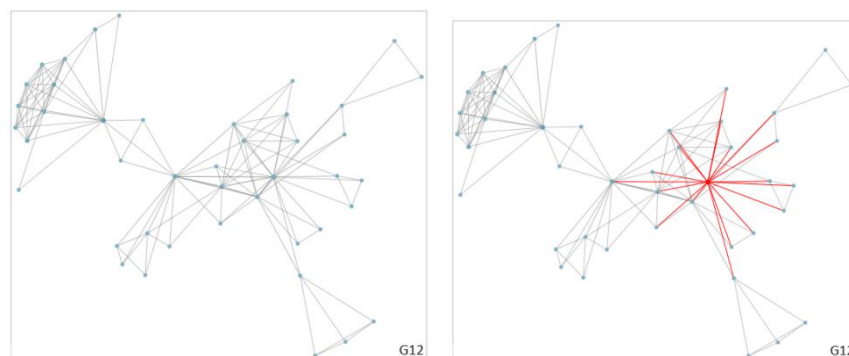


Figura 72. Grupo G12 a la izquierda y la red ego Miquel Casals iluminada a la derecha.
 Elaboración propia en NodeXL.

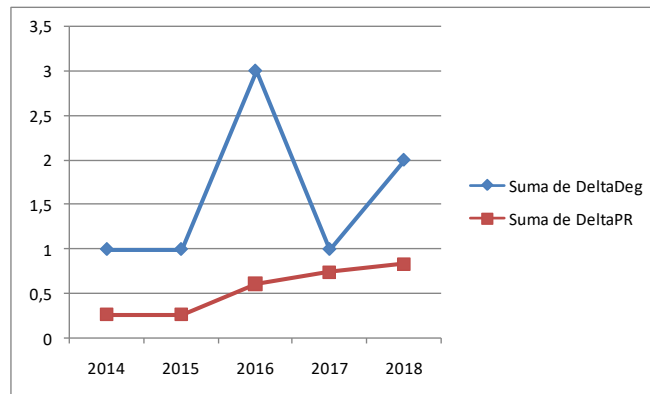


Figura 73. Gráfico para la visualización de la diferencia de grado y de PageRank entre el nodo Miquel Casals y su segundo nodo para los años desde 2014 hasta 2018.

7.2.6 "Ego network" de Anna Rius en el grupo G13

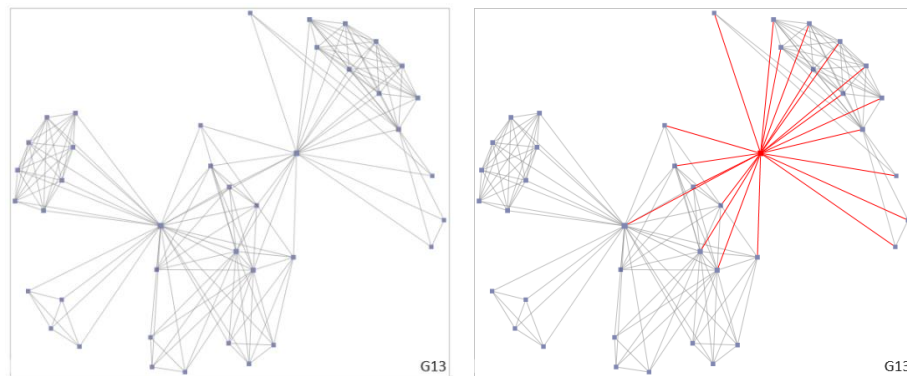


Figura 74. Grupo G13 a la izquierda y la red ego Anna Rius iluminada a la derecha.
Elaboración propia en NodeXL.

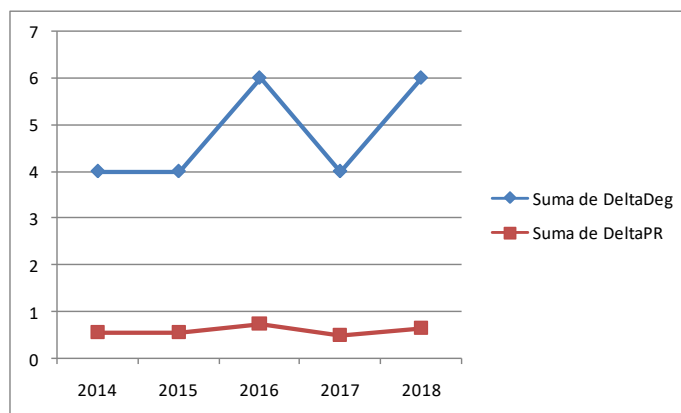


Figura 75. Gráfico para la visualización de la diferencia de grado y de PageRank entre el nodo Anna Rius y su segundo nodo para los años desde 2014 hasta 2018.

7.2.7 "Ego network" de Oriol Lordan en el grupo G14

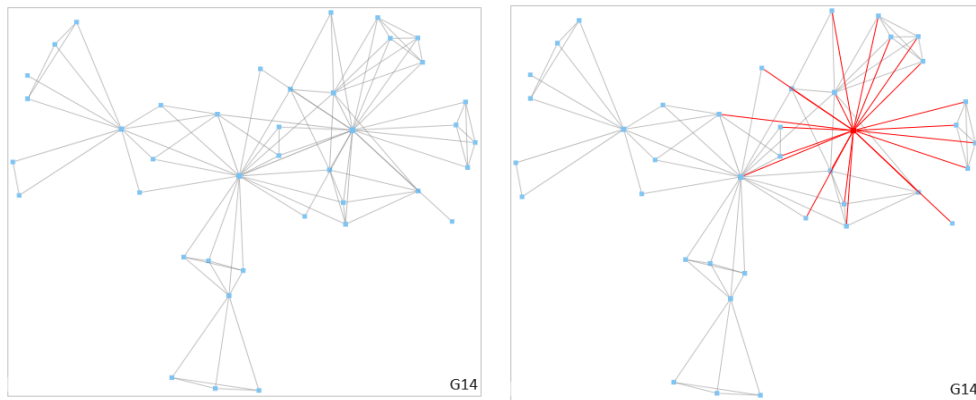


Figura 76. Grupo G14 a la izquierda y la red ego Oriol Lordan iluminada a la derecha. Elaboración propia en NodeXL.

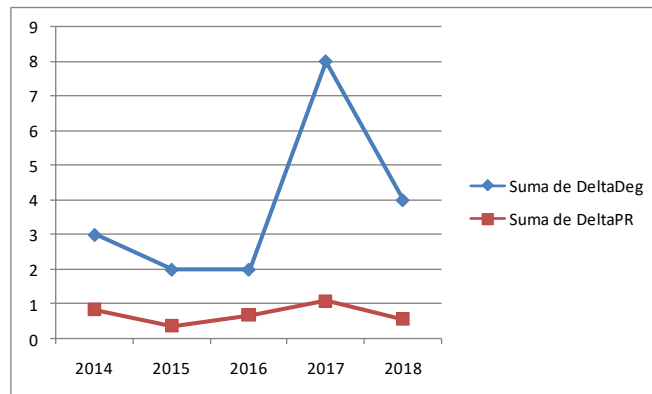


Figura 77. Gráfico para la visualización de la diferencia de grado y de PageRank entre el nodo Oriol Lordan y su segundo nodo para los años desde 2014 hasta 2018.

7.2.8 "Ego network" de María Carmen Riva en el grupo G15

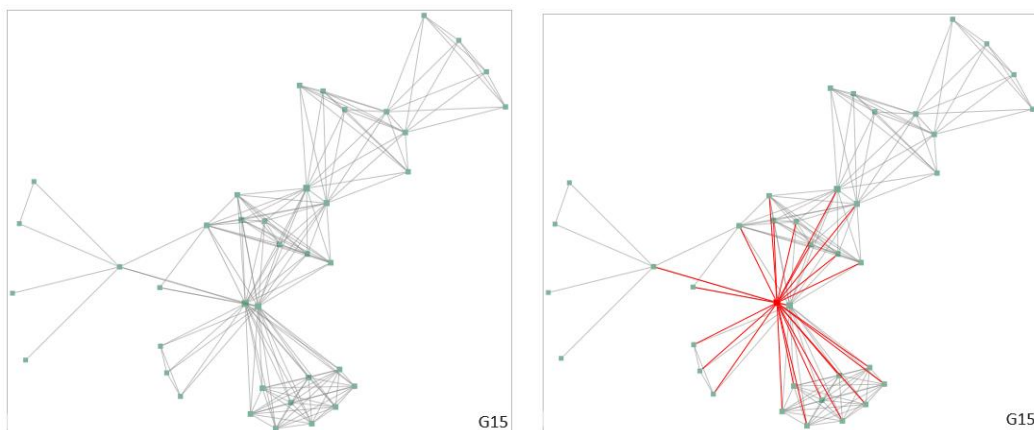


Figura 78. Grupo G15 a la izquierda y la red ego María Carmen Riva iluminada a la derecha. Elaboración propia en NodeXL.

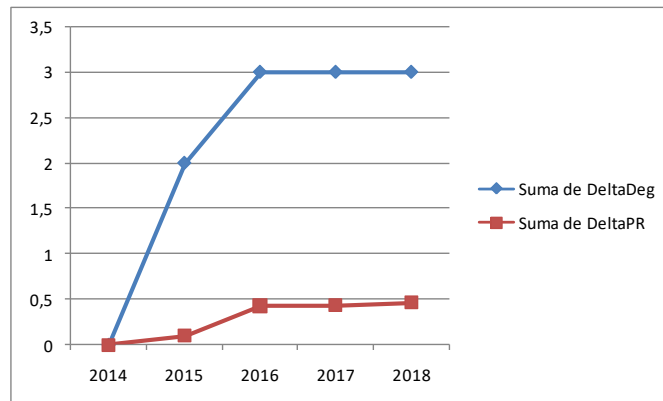


Figura 79. Gráfico para la visualización de la diferencia de grado y de PageRank entre el nodo María Carmen Riva y su segundo nodo para los años desde 2014 hasta 2018.

7.2.9 "Ego network" de Josep Pou en el grupo G16

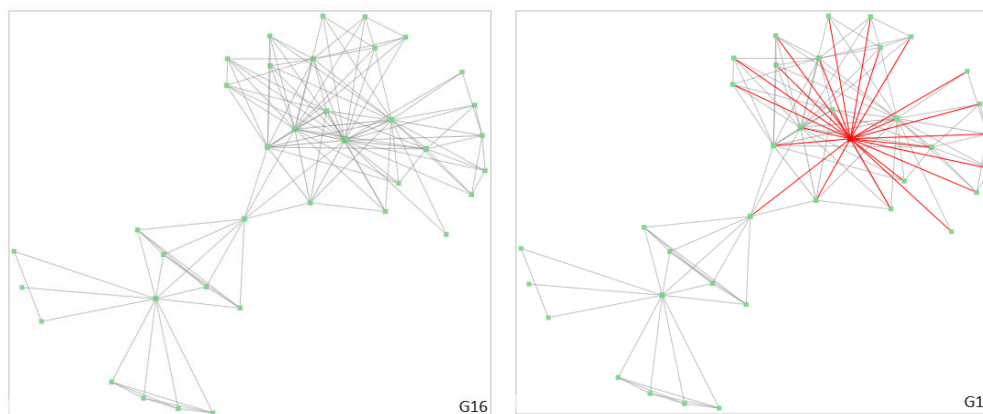


Figura 80. Grupo G16 a la izquierda y la red ego Josep Pou iluminada a la derecha. Elaboración propia en NodeXL.

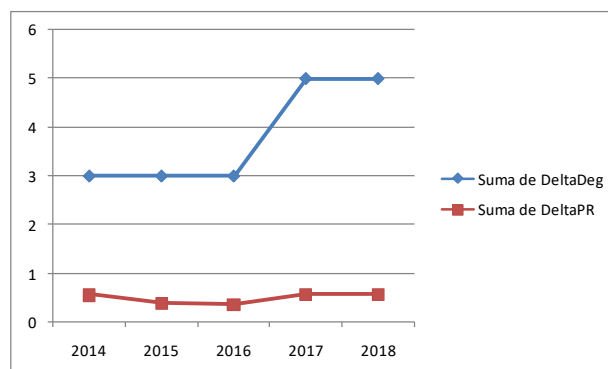


Figura 81. Gráfico para la visualización de la diferencia de grado y de PageRank entre el nodo Josep Pou y su segundo nodo para los años desde 2014 hasta 2018.

7.2.10 "Ego network" de Meritxell Martí en el grupo G17

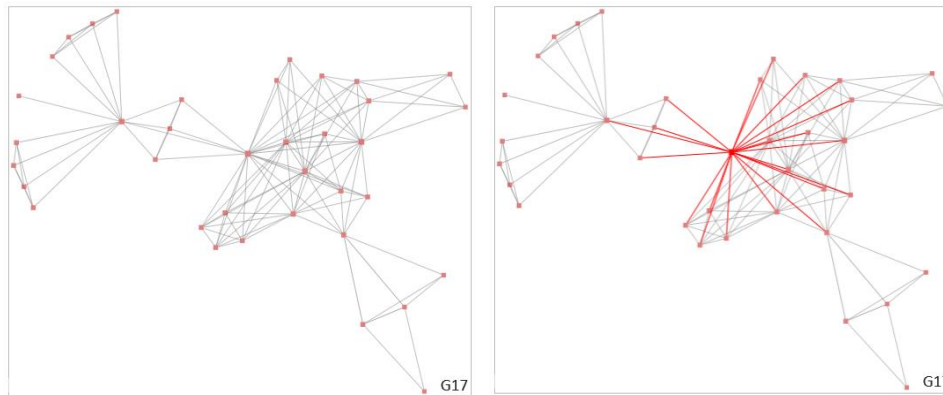


Figura 82. Grupo G17 a la izquierda y la red ego Meritxell Martí iluminada a la derecha.
Elaboración propia en NodeXL.

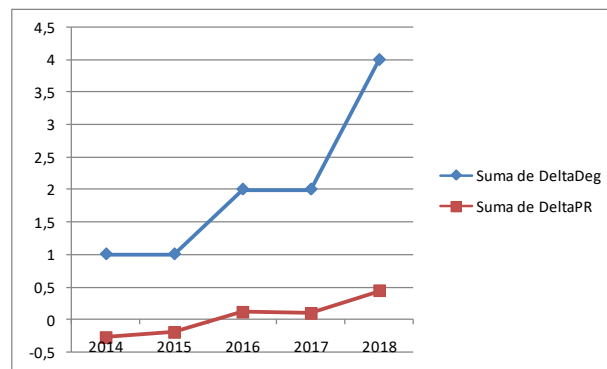


Figura 83. Gráfico para la visualización de la diferencia de grado y de PageRank entre el nodo Meritxell Martí y su segundo nodo para los años desde 2014 hasta 2018.

7.2.11 "Ego network" de Francisco Javier Oliver en el grupo G18

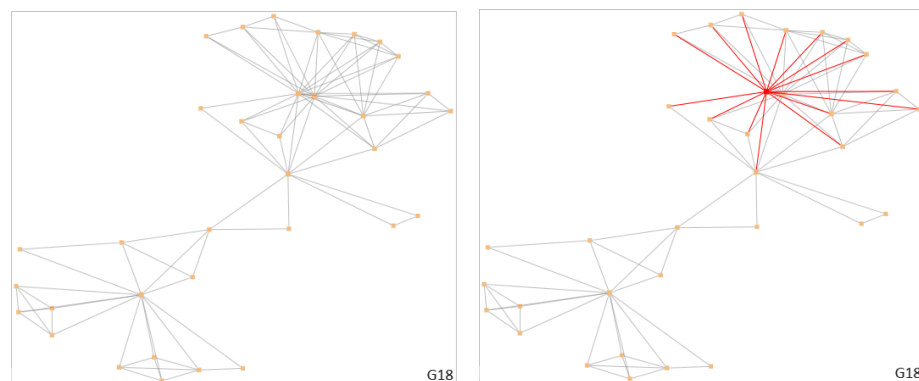


Figura 84. Grupo G18 a la izquierda y la red ego Francisco Javier Oliver iluminada a la derecha.
Elaboración propia en NodeXL.

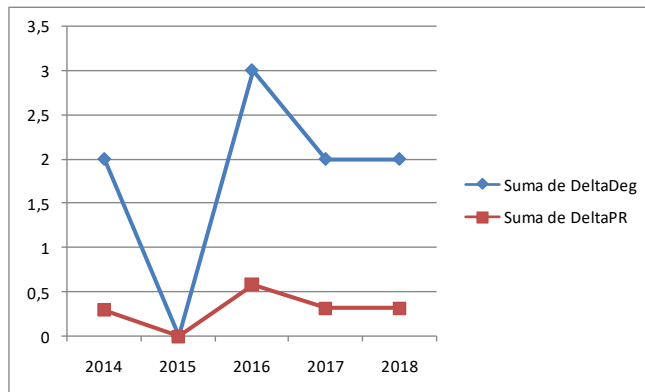


Figura 85. Gráfico para la visualización de la diferencia de grado y de PageRank entre el nodo Francisco Javier Oliver y su segundo nodo para los años desde 2014 hasta 2018.

7.2.12 "Ego network" de Martí Rosas Casals en el grupo G19

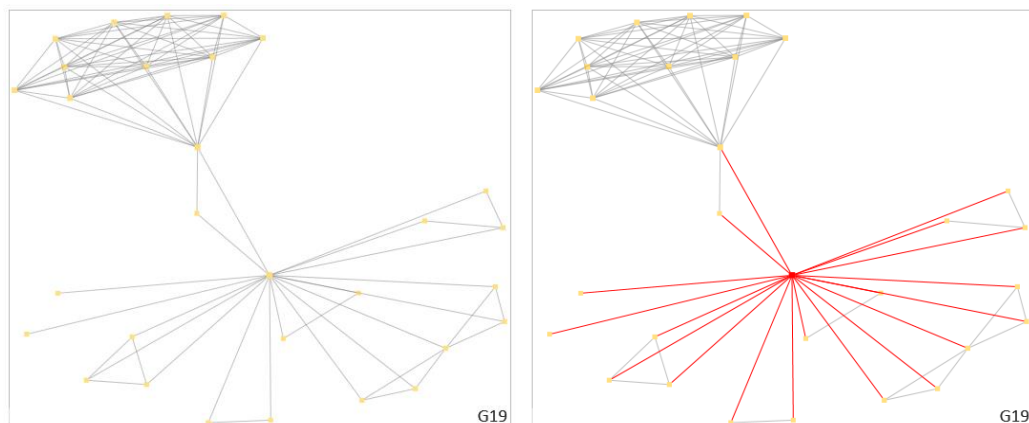


Figura 86. Grupo G19 a la izquierda y la red ego Martí Rosas Casals iluminada a la derecha. Elaboración propia en NodeXL.

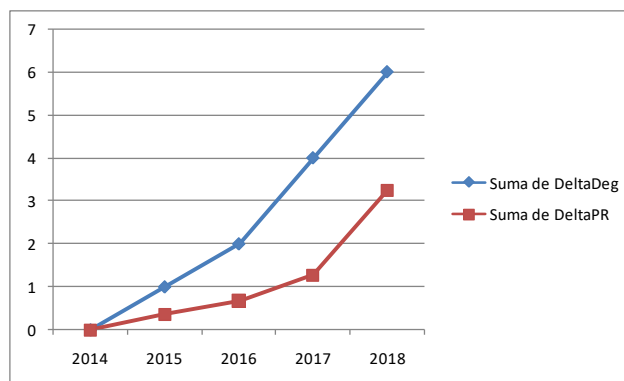


Figura 87. Gráfico para la visualización de la diferencia de grado y de PageRank entre el nodo Martí Rosas Casals y su segundo nodo para los años desde 2014 hasta 2018.

7.2.17 "Ego network" de Ramón Quintanilla en el grupo G24

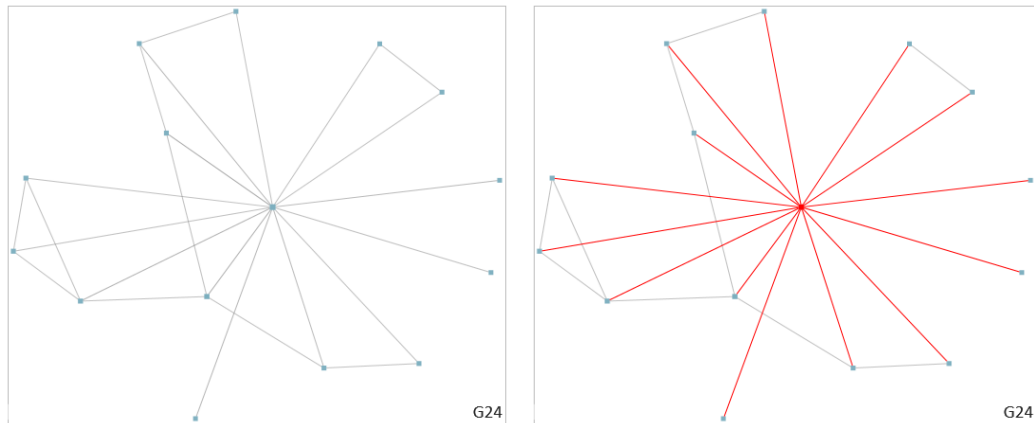


Figura 88. Grupo G24 a la izquierda y la red ego Ramón Quintanilla iluminada a la derecha.
Elaboración propia en NodeXL.

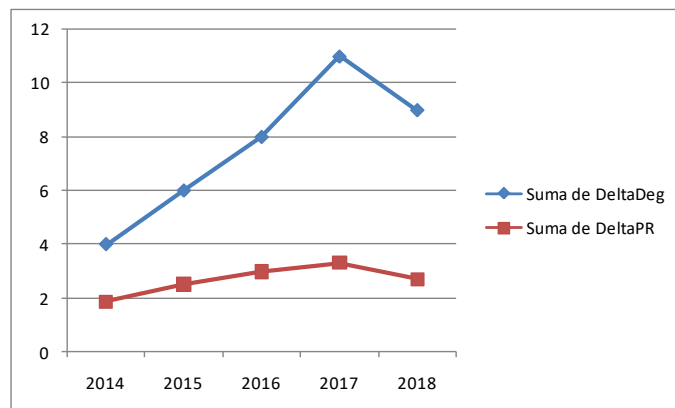


Figura 89. Gráfico para la visualización de la diferencia de grado y de PageRank entre el nodo Ramón Quintanilla y su segundo nodo para los años desde 2014 hasta 2018.