

# Redes Sociales - Grafos

Santiago Breogán Pérez Pita - [santibreo@gmail.com](mailto:santibreo@gmail.com)

01 julio, 2019

## Contents

<b>Trabajo previo</b>	<b>2</b>
<b>Ejercicios</b>	<b>3</b>
1 - Los parámetros de una red aleatoria con parámetros N y p que tenga un número de nodos y aristas esperadas similar a la red USAir97 e indicar la ruta para hacerlo en <b>Pajek</b> (0.5 puntos) . . . . .	3
2 - Los parámetros de una red libre de escala que tenga un número de nodos y aristas esperadas similar a la red USAir97 e indicar la ruta para hacerlo en <b>Pajek</b> (0.5 puntos). . . . .	4
3 - Los parámetros de una red de pequeño mundo que tenga un número de nodos y aristas esperadas similar a la red USAir97 e indicar la ruta para hacerlo en <b>Pajek</b> (0.5 puntos). . . . .	5
4 - Para las cuatro redes (las tres de los apartados anteriores y la original USAir97), calcular su distribución del grado e indicar si son aleatorias o de libre escala cada una de ellas (1 punto). . . . .	6
5 - Para las cuatro redes (las tres de los apartados anteriores y la original USAir97), calcular su distancia media e indicar si son de pequeño mundo cada una de ellas (1 punto). . . . .	7
6 - Para la red USAir97 realizar un clustering por el método de Louvain e interpretar la modularidad alcanzada. Saca de forma separada cada una de las redes de cada cluster y la red donde todos los cluster se representan como un punto (1.5 puntos). . . . .	8
7 - Ver la importancia de cada cluster a los que pertenece alguno de los tres nodos seleccionados y la capacidad de cada uno de ellos para intermediar (1.5 puntos). . . . .	12
8 - Busca los cinco nodos más importantes en la red y los cinco nodos con una interpretación más diferente en función del criterio de centralidad usado (1.5 puntos). . . . .	13
9 - Indicar la medida de centralidad que nos parece mejor en este caso y la motivación de esta opinión en esta red (0.5 puntos). . . . .	13
10 - Para los tres nodos indicados para cada uno de vosotros, mira su centralidad en general en la red (de forma comparativa con el resto de la red), en su cluster y para intermediar entre los distintos cluster (1.5 puntos). . . . .	13
<b>Comentarios</b>	<b>14</b>

## Trabajo previo

La red que voy a trabajar es una red que contiene información sobre los aeropuertos de EEUU, y los vuelos que se produjeron entre ellos en el año 1997.

Al eliminar los 41 nodos que me correspondían queda una red con 291 nodos y 1634 aristas. Al representar los vuelos entre los aeropuertos podría trabajarse como una red direccional, usando arcos (no necesariamente los vuelos que van de un aeropuerto a otro son tan numerosos como los que hacen el recorrido contrario), sin embargo, el fichero no contiene esta información, así que la trabajaré como red **no** direccional. Siendo así sus parámetros son:

$$\text{número de nodos: } N = 291$$

$$\text{número de aristas: } L = 1634$$

$$\text{grado medio: } \langle k \rangle = \frac{1}{N} \sum_{i=1}^N k_i = \frac{2 \cdot L}{N} = 11'230$$

$$\text{probabilidad: } p = \frac{\text{aristas reales}}{\text{aristas posibles}} = \frac{L}{\frac{N \cdot (N-1)}{2}} = \frac{\langle k \rangle}{N-1} = 0'039$$

Con los parámetros calculados podemos pasar a realizar los ejercicios.

Antes una aclaración sobre la instalación de **Pajek**. Se trata de un software especializado para el tratamiento de redes, portátil y gratuito, existen tres versiones disponibles, que se adecuan al tamaño de la red limitando opciones. Nosotros hemos trabajado con la versión base (**Pjeck64**), que es la adecuada para redes de menos de 100 000 nodos y es la que más opciones permite. Se puede descargar en [aquí](#).

---

*Como curiosidad: Las aristas posibles en una red no direccional son la suma de sus nodos menos uno, ya que el primer nodo puede estar conectado a los  $N - 1$  nodos restantes, el siguiente a los  $N - 2$  (ni él mismo, ni el primero) y así. La suma de los  $N - 1$  primeros números naturales es la expresión usada:  $\frac{N \cdot (N-1)}{2}$ .*

*Si fuera una red direccional el número de arcos posibles sería el doble  $N \cdot (N - 1)$  ya que cada nodo puede conectarse a cualquier otro. Visto esto se hace evidente que una red no direccional es como una direccional con todos sus arcos dobles ( $L$  aristas  $\rightarrow$   $2L$  arcos), es decir, cada arista equivale a dos arcos.*

---

## Ejercicios

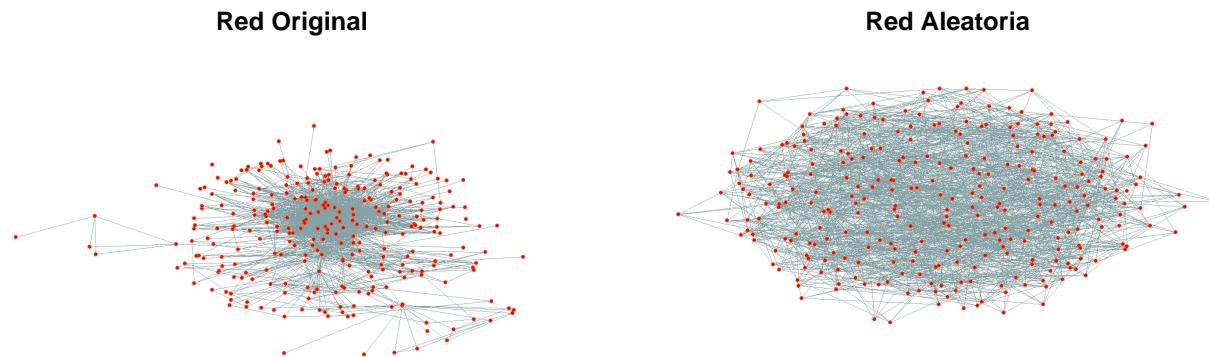
Ayudándote con **Pajek** y cualquier otra herramienta de tratamiento de datos, **calcular**:

**1 - Los parámetros de una red aleatoria con parámetros N y p que tenga un número de nodos y aristas esperadas similar a la red USAir97 e indicar la ruta para hacerlo en Pajek (0.5 puntos)**

Generar una red aleatoria en **Pajek** mediante el algoritmo de Erdos - Renyi es sencillo:

1. Vamos a la barra de herramientas y hacemos click en el menu desplegable **Network**.
2. A partir de aquí navegamos:
  1. **Create Random Network > Bernoulli/Poisson > Undirected > General.**
  3. Nos pide unos parámetros:
    - *Number of Vertices: 292*
    - *Average Degree of Vertices: 11'230*

Haciendo esto genera de forma prácticamente inmediata una red aleatoria que podemos comparar con la original:



Ya vemos que mientras la Red Original presenta pocas conexiones en la periferia del grafo y cada vez más al aproximarnos al centro, la Red Aleatoria tiene una densidad más homogénea. La red generada presenta los siguientes parámetros:

- Número de nodos:  $N = 292$
- Número de aristas:  $L = 1599$
- Grado promedio:  $\langle k \rangle = 10'990$  (esta desviación respecto al *input* se debe al algoritmo de construcción)
- Probabilidad:  $p = 0'038$

Debemos notar que al fijar nosotros el número de nodos y el grado promedio el resto de parámetros van a ser muy similares a los de nuestra red.

**2 - Los parámetros de una red libre de escala que tenga un número de nodos y aristas esperadas similar a la red USAir97 e indicar la ruta para hacerlo en Pajek (0.5 puntos).**

Generar una red libre de escala en **Pajek** mediante el algoritmo de Barabási – Albert tampoco es complicado:

1. Vamos a la barra de herramientas y hacemos click en el menu desplegable **Network**.

2. A partir de aquí navegamos:

1. **Create Random Network > Scale Free > Undirected.**

3. Nos pide unos parámetros:

- *Number of Vertices:* **292**
- *Number of Lines:* **1634**
- *Average Degree of Vertices:* **11'230**
- *Number of Vertices in initial Erdos - Renyi Network:* **50** (número inicial de nodos)
- *Initial Probability of lines:* **0'0038**
- *Alpha:* **0'18**

La probabilidad de que un nodo nuevo forme enlace en las redes libres de escala generadas mediante el algoritmo de Barabási - Albert es:

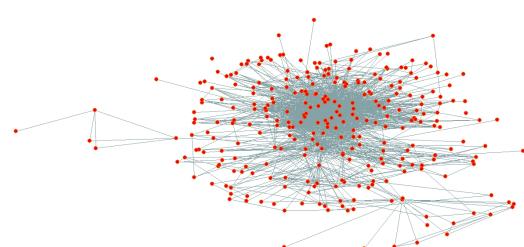
$$p_i = \frac{k_i}{\sum_j k_j}$$

Por lo tanto es proporcional a la contribución del nodo al grado medio. Esto da lugar a que nodos con más aristas son propensos a obtener nuevas aristas sentando las bases del algoritmo.

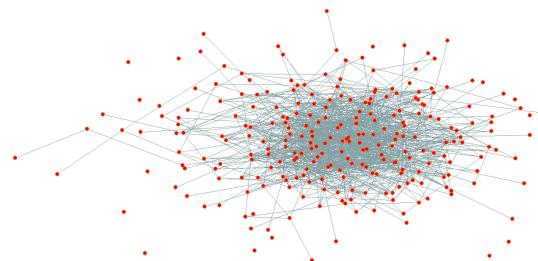
Los últimos parámetros que nos pide **Pajek** son propios del algoritmo y afectarán a la red resultante, yo he elegido un número elevado de nodos inicial para poder usar como probabilidad inicial una similar a la de nuestra red, garantizándome que se produjera algún enlace, luego he notado que aumentar **alpha** implica más nodos y menor grado promedio, así que he empleado un valor cuyos resultados me parecían adecuados

Haciendo esto genera de forma prácticamente inmediata una red de libre escala que podemos comparar con la original:

**Red Original**



**Red Libre de Escala**



En este caso si observamos ese aumento de densidad de enlaces en la zona central, común a ambas redes.

Los parámetros de la red de libre escala resultante son:

- Número de nodos:  $N = 263$
- Número de aristas:  $L = 1364$
- Grado promedio:  $\langle k \rangle = 10.373$  (esta desviación respecto al *input* se debe a la forma de construir la red del algoritmo)
- Probabilidad:  $p = 0.040$

Igual que antes el fijar nosotros varios parámetros garantiza cierta similitud entre los parámetros de la red original y la obtenida.

**3 - Los parámetros de una red de pequeño mundo que tenga un número de nodos y aristas esperadas similar a la red USAir97 e indicar la ruta para hacerlo en Pajek (0.5 puntos).**

Generar una red de pequeño mundo en **Pajek** según el algoritmo de Watts y Strogatz es igual de sencillo que en los casos anteriores:

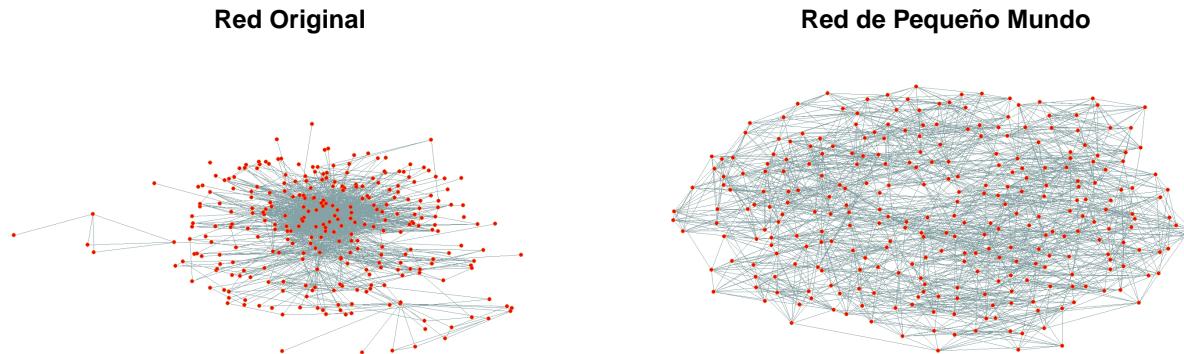
1. Vamos a la barra de herramientas y hacemos click en el menu desplegable **Network**.
2. A partir de aquí navegamos:
  1. **Create Random Network > Small World.**
3. Nos pide unos parámetros:
  - *Number of Vertices:* **292**
  - *Number of Linked Neighbors on each Side of a Vertex:* **6**
  - *Replacement probability :* **0.15**

El algoritmo consiste en escribir todas las aristas posibles  $(n_i, n_j)$  y conservar solo aquellas en las que se cumpla:

$$0 < |i - j| \mod \left( N - 1 - \frac{\langle k \rangle}{2} \right) \leq \frac{\langle k \rangle}{2}$$

Y luego de las que sobrevivan, reemplazar el nodo destino con una probabilidad igual a  $\beta$  que se llama probabilidad de reemplazamiento. Cuanto mayor sea más se asemejará nuestra red a una red aleatoria, de ahí su valor

Haciendo esto genera de forma prácticamente inmediata una red de pequeño mundo que podemos comparar con la original:



En este caso observamos aumentos de densidad en varias regiones deslocalizadas, común a ambas redes.

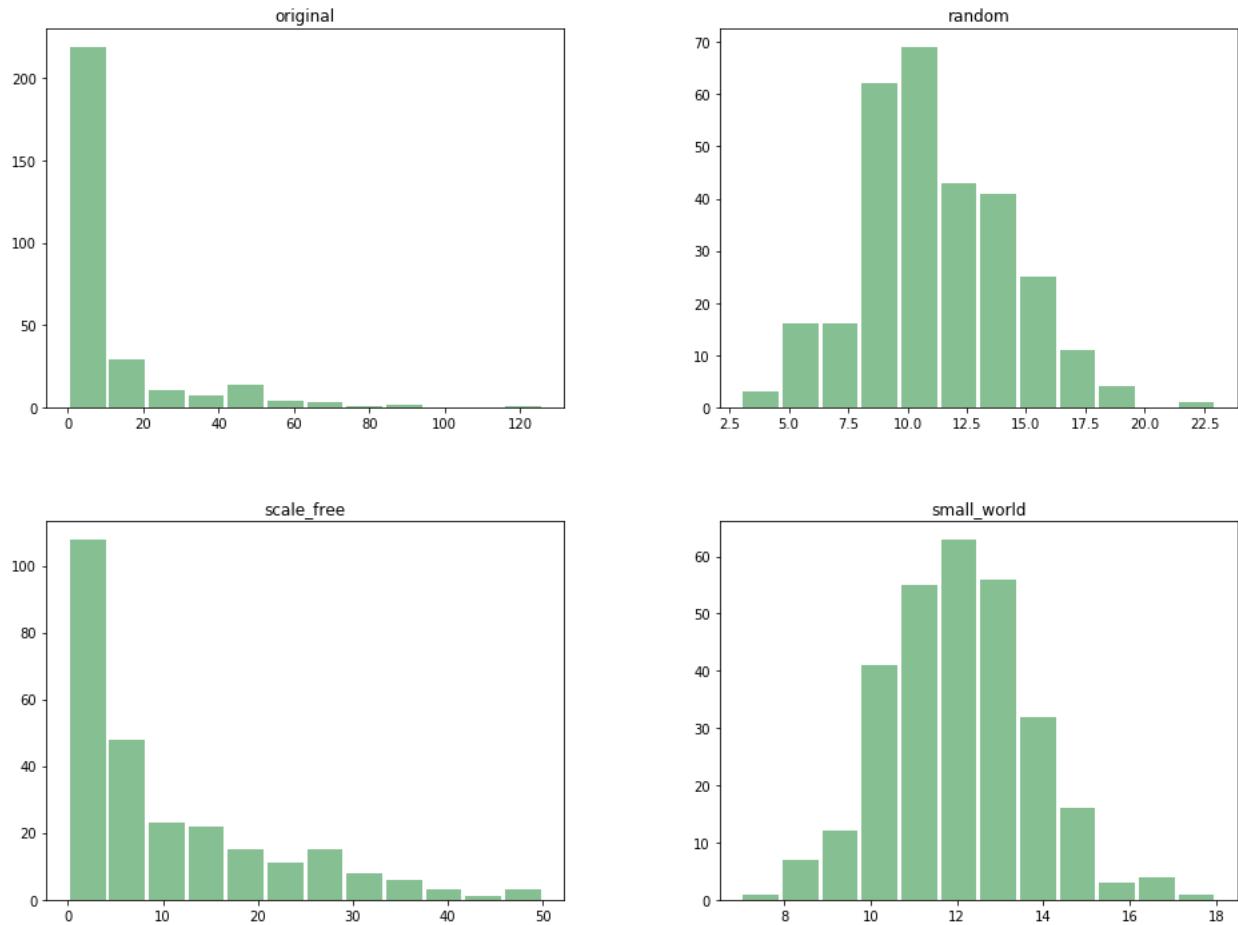
Los parámetros de la red de pequeño mundo resultante son:

- Número de nodos:  $N = 291$
- Número de aristas:  $L = 1746$
- Grado promedio:  $\langle k \rangle = 12$  (esta desviación respecto al *input* se debe a la forma de construir la red del algoritmo)
- Probabilidad:  $p = 0'041$

Como siempre nuestros parámetros nos procuran una red similar a la red original.

**4 - Para las cuatro redes (las tres de los apartados anteriores y la original USAir97), calcular su distribución del grado e indicar si son aleatorias o de libre escala cada una de ellas (1 punto).**

Usando **Pajek** he obtenido fácilmente los vectores con el grado de cada red, y los he representado, obteniendo los siguientes gráficos:

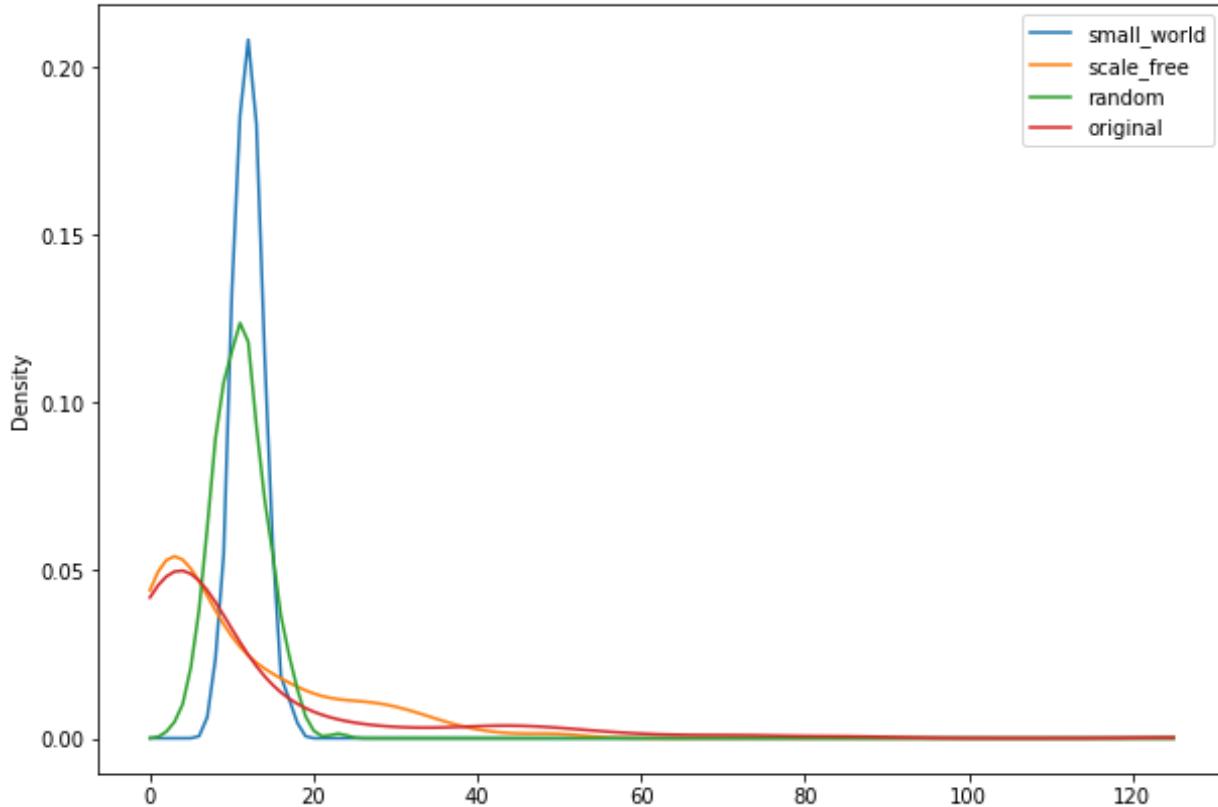


Como puede verse las distribuciones de la red aleatoria y de pequeño mundo se asemejan más a una distribución de Poisson de forma que la mayoría de nodos muestran un grado similar al grado promedio. No obstante hay que señalar que la red aleatoria presenta grado promedio menor, y por lo tanto aparece menos centrada, aún así se asemeja ya más a una distribución de Poisson (esta aproximación es válida en el límite  $\langle k \rangle \ll N$  dándose  $\ll N$ ) que a una binomial (que presenta una cola menor). Por lo tanto **la Red Aleatoria y la de Pequeño Mundo son Redes Aleatorias**.

Por su parte tanto la red libre de escala y la red original (es como he llamado a USAir97) presentan distribuciones con un marcado decaimiento exponencial, lo cual es un rasgo propio de las redes libres de

escala. Cabe notar que el decaimiento de la red original es mucho más rápido que el de la libre de escala, lo que implica muchos nodos de grado bajo y algunos de grado muy alto, ya que en todas las redes el número de nodos es aproximadamente el mismo y su grado promedio también. Son, en consecuencia, **la Red Libre de Escala y la Red USAir97 Redes Libres de Escala**.

Otra información interesante la podemos sacar comparando las distribuciones juntas:



En este caso vemos como la original y la libre de escala alcanzan valores mayores y como la de pequeño mundo y la aleatoria presentan valores muy concentrados en torno al grado promedio, sobretodo la de pequeño mundo.

**5 - Para las cuatro redes (las tres de los apartados anteriores y la original USAir97), calcular su distancia media e indicar si son de pequeño mundo cada una de ellas (1 punto).**

Para obtener la distancia media de una red podemos seguir dos caminos:

1. Empleando la definición de Closeness: Este método consiste en aplicar la definición de *Closeness*, que es uno de los criterios de centralidad, para obtener la distancia promedio de cada nodo:

$$\text{Closeness del nodo } i : \quad C_i = \frac{1}{\sum_{j \neq i} d(i,j)} \rightarrow \langle d_i \rangle = \frac{\sum_{j \neq i} d(i,j)}{n_i \text{ (nodos alcanzables desde } i)}$$

$$\text{Distancia media : } \langle d \rangle = \frac{\sum_i \langle d_i \rangle}{N} = \frac{\sum_i \frac{1}{n_i} \sum_{j \neq i} d(i,j)}{N} = \frac{\sum_i \frac{1}{C_i \cdot n_i}}{N}$$

De esta forma, y si la red fuera completamente conexa (todos los nodos acceden a  $N - 1$  nodos), bastaría con calcular el vector de *Closeness*, invertirlo, hacer la media y dividirla por  $N-1$ . Plantearme el ejercicio de esta

forma me ha servido para notar que lo que devuelve Pajek cuando le pedimos que calcule *Closeness* no es el valor absoluto del indicador de centralidad, sino su valor relativo, que se define así:

$$\text{Closeness relativo del nodo } i : CR_i = \frac{n_i}{\sum_{j \neq i} d(i, j)} = n_i \cdot C_i$$

Luego basta con invertir el vector de Closeness y hacer la media para obtener la distancia media. > ¡Ojo! En nuestra Red Original hay nodos aislados, cuyo closeness es cero, y al invertirlos nos dan infinito. Luego la media nos quedaría infinito. Estos nodos se contabilizan a la hora de hacer la media como si contribuyeran con una **distanzia 0** lo cual es incoherente, pero es la forma en la que las medidas tomadas por ambos caminos confluyen (y como se define la distancia media de los grafos que no son completamente conexos).

2. Empleando la funcionalidad de **Pajek**: El segundo camino es bastante más corto, basta con usar la funcionalidad **Network > Create Vector > Distribution of Distances\*** y mirar la *Report Window* en la que está la distancia promedio de la red.

Tras comprobar que ambos métodos arrojaban los mismos resultados he obtenido los siguientes valores para cada red:

- Original:  $\langle d \rangle = 2'776 \mid d_{max} = 6 \mid \frac{\log(N)}{\log(\langle k \rangle)} = 2'346$
- Aleatoria:  $\langle d \rangle = 2'624 \mid d_{max} = 5 \mid \frac{\log(N)}{\log(\langle k \rangle)} = 2'367$
- Libre de Escala:  $\langle d \rangle = 2'798 \mid d_{max} = 7 \mid \frac{\log(N)}{\log(\langle k \rangle)} = 2'382$
- de Pequeño Mundo:  $\langle d \rangle = 2'988 \mid d_{max} = 5 \mid \frac{\log(N)}{\log(\langle k \rangle)} = 2'283$

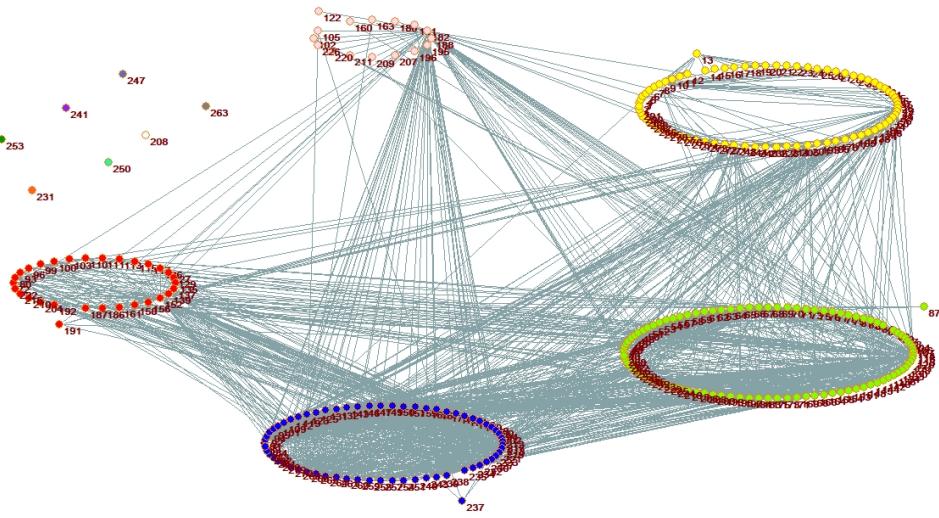
De los resultados hay que destacar que las redes que hemos catalogado como libres de escala presentan distancias máximas mayores a las de las redes catalogadas como aleatorias, esto tiene sentido ya que estas redes presentan varios nodos con un número de aristas pequeño, aunque suelen estar conectados a nodos con gran cantidad de aristas. En general a la vista de los resultados no podemos descartar que las redes sean de pequeño mundo, de hecho a mi juicio todas las redes son redes de pequeño mundo ya que se aproxima bastante bien su distancia promedio mediante la **fórmula de los pequeños mundos** y en todas ellas las distancias máximas son similares a  $\log(N)$ .

## 6 - Para la red USAir97 realizar un clustering por el método de Louvain e interpretar la modularidad alcanzada. Saca de forma separada cada una de las redes de cada cluster y la red donde todos los cluster se representan como un punto (1.5 puntos).

Sobre este ejercicio no está claro cual es el camino adecuado: Si usamos una resolución de 1, que corresponde con el método de Louvain estándar obtenemos una solución con varios clústeres y un modularidad intermedia. Si reducimos la resolución vamos obteniendo cada vez menos clústeres, con una modularidad mayor, ya que en esencia lo que está ocurriendo es que se está forzando la generación de clústeres aislados, quedando los aeropuertos que no presentan ninguna arista, y el resto como un único clúster.

La forma que yo he considerado óptima es con resolución unidad, que aplica el algoritmo de Louvain estándar. Se generan varios clústeres de más de un miembro, de forma que tiene sentido hablar de relevancia de los clústeres, ya que de otra forma un clúster sería esencialmente toda la red, y el resto individuos sueltos.

- La modularidad alcanzada por este método es: **0.229** que es un valor intermedio, no se han logrado islas fuertemente intraconectadas y débilmente interconectadas, pero si se han encontrado grupos fuertemente intraconectados. Por supuesto sigue existiendo la desviación ocasionada por los nodos sin conexiones (estos nodos en realidad no pertenecen a la red, y no deberían formar parte de ella, no obstante creo que esto debiera haberse facilitado con el fichero original). La disposición de mis clústeres es la siguiente (debe tenerse en cuenta que la numeración de vértices para valores superiores a 249 es  $x + 41$  en la original eg: mi 250 es el 291 original):

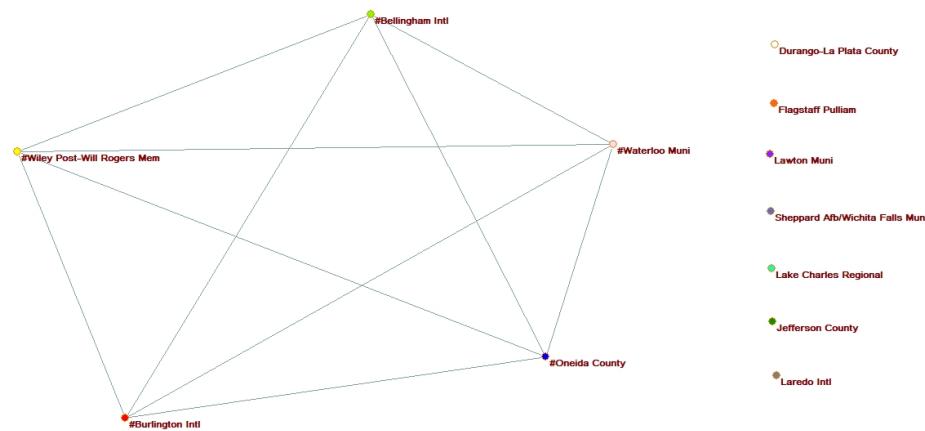


### **Rojo: C-3, Azul: C4, Rosa: C-5, Verde: C-2 y Amarillo: C-1**

Los clusters de un solo miembro son los aeropuertos: Durango-La Plata County, Flagstaff Pulliam, Lawton Muni, Sheppard Afb/Wichita Falls Mun, Lake Charles Regional, Jefferson County y Laredo Intl.

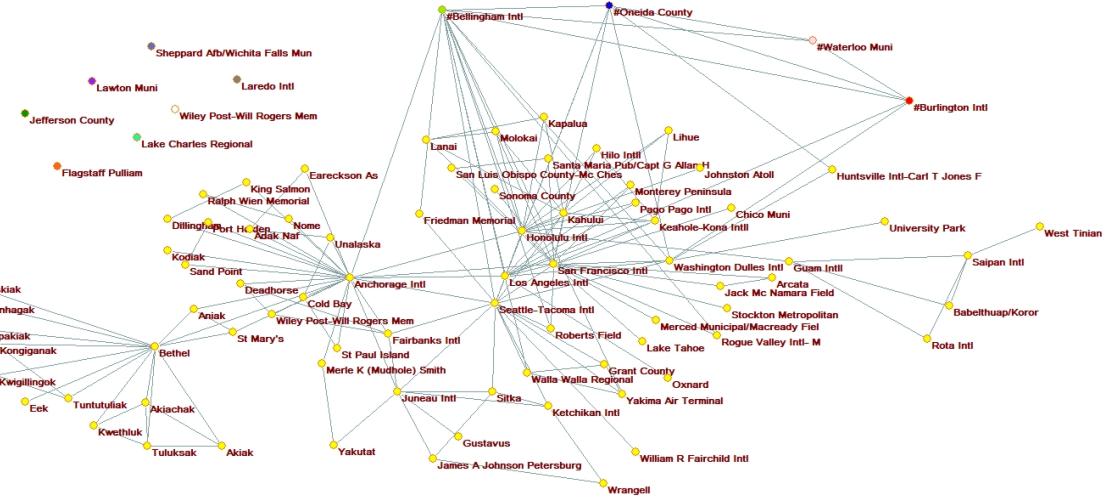
No sé como hacer lo que se pide en el enunciado ya que en **Pajek**, cuando hacemos en la barra de herramientas **Operations > Network + Partition > Shrink Network** únicamente nos permite elegir una red para que no aparezca contraída (o ninguna). Como no conozco la forma de agrupar un único clúster lo que voy a hacer es representar todos agrupados (es un gráfico muy similar al mostrado más arriba) y además representar cada clúster expandido con los demás agrupados. Con estos gráficos yo creo que se recoge bastante bien mucha información: cómo de intraconectado está nuestro clúster, cuál es su tamaño y como de interconectado está, que creo que es la información que se pedía que representásemos.

*Quiero dejar claro que si sé como representar cada clúster aislado, pero entre representarlo aislado y representarlo con el resto de clústeres contraídos no había gran diferencia, por eso me he decantado por esta representación.*

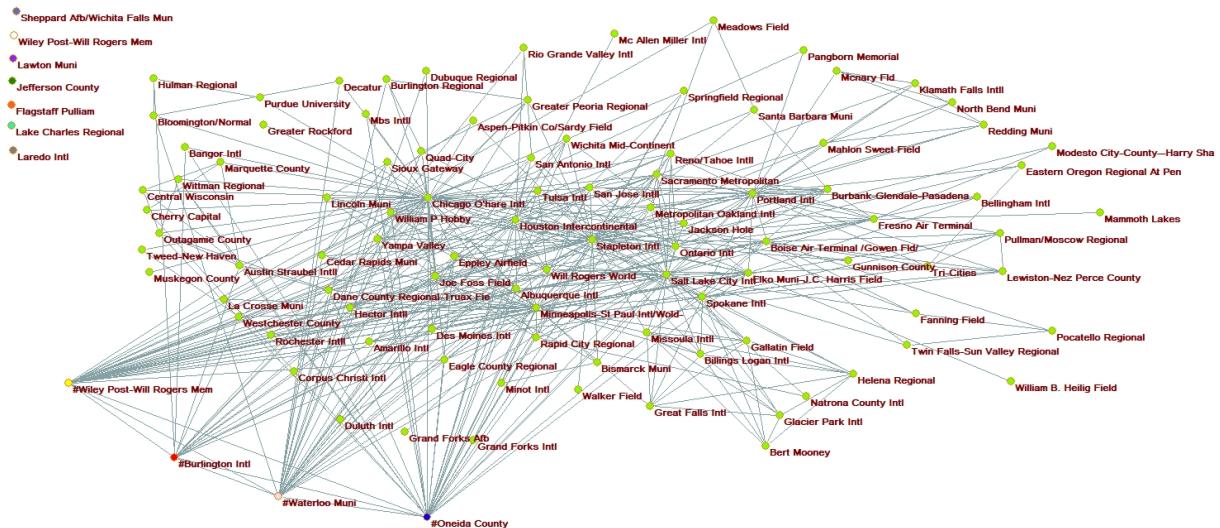


Los gráficos son los siguientes:

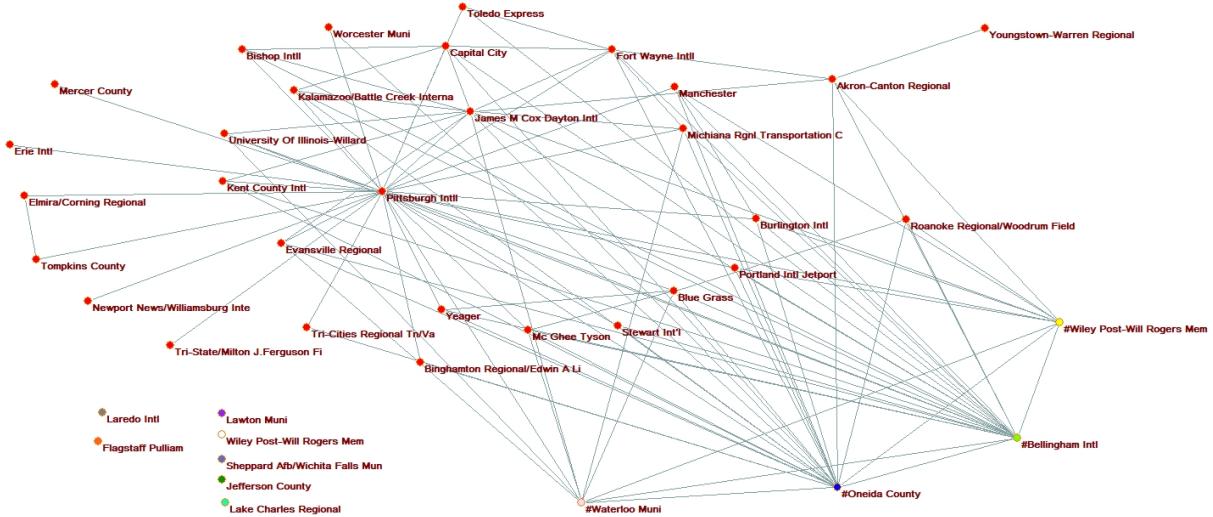
- Clúster 1 (C-1):



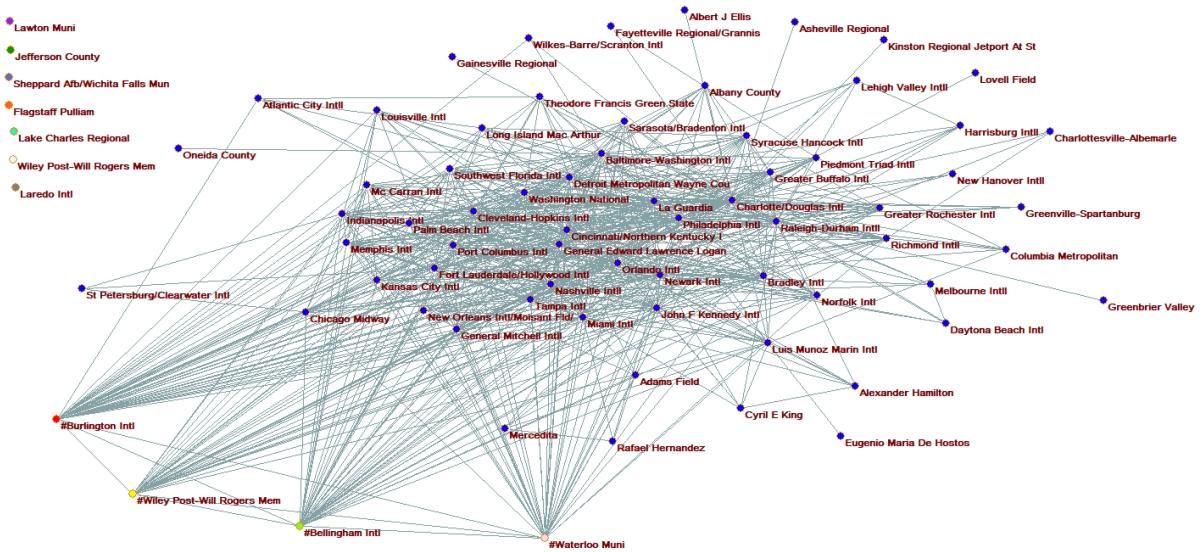
- Clúster 2 (C-2):



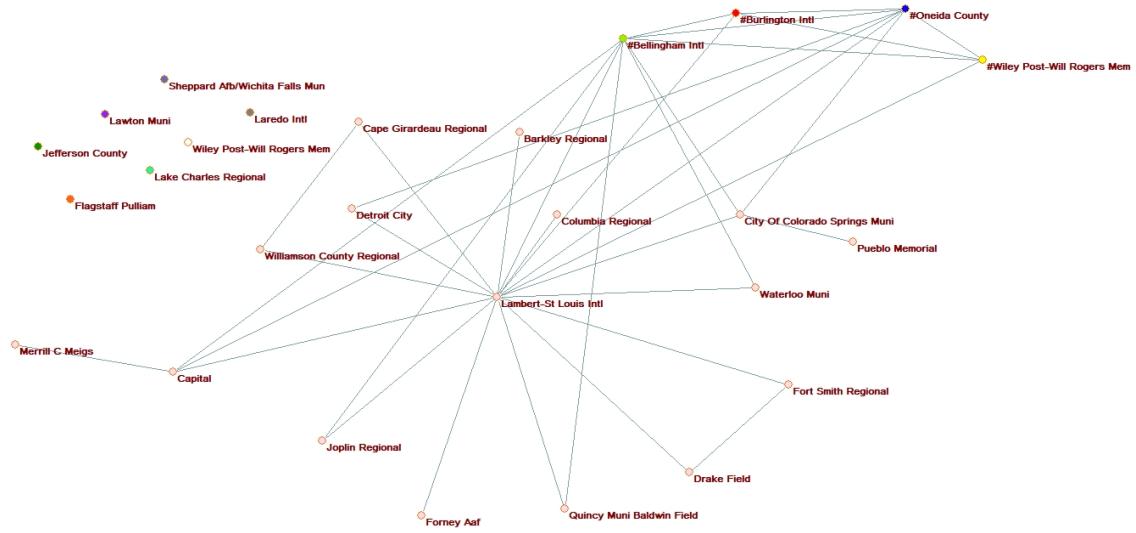
- Clúster 3 (C-3):



- Clúster 4 (C-4):



- Clúster 5 (C-5):



**7 - Ver la importancia de cada cluster a los que pertenece alguno de los tres nodos seleccionados y la capacidad de cada uno de ellos para intermediar (1.5 puntos).**

Los nodos que se me ha pedido que estudie son:

116. Klamath Falls Intl
117. Erie Intl
118. Chicago O'hare Intl

El primero y el último pertenecen al segundo clúster mientras que Enrie Intl pertenece al tercero.

Si atendemos a la red generada por los clústeres, veremos que todos los clústeres están comunicados entre sí, de forma que su relevancia la medimos mediante el criterio de centralidad *Closeness* y nos aparece en todos con el mismo valor e igual a 1 (si eliminamos los nodos aislados), si aplicamos la fórmula definida anteriormente:

$$\text{Closeness del nodo } i : C_i = \frac{1}{\sum_{j \neq i} d(i, j)} \rightarrow \text{Closeness relativa del nodo } i : n_i \cdot C_i$$

la distancia entre todos los nodos es 1, luego ratificamos que el valor de *Closeness* es el *Closeness* relativo pero además vemos que no tiene sentido analizar de esta forma la relevancia de los clusters ya que todos ellos están conectados entre sí, luego todos tienen *Closeness* igual a 1 (si ignoramos los nodos aislados) y todos tienen *Betweenness* igual a 0 ya que no intermedian en ninguna relación. Supongo que este proceso tiene sentido realizarlo habiendo contraído un único clúster, manteniendo el resto dispersos, sin embargo, esto no he encontrado como puede realizarse, y no me termina de parecer riguroso comparar clústeres con nodos, lo suyo es comparar entidades del mismo orden.

De esta forma podemos definir los valores de mi red de clústeres para todos los clústeres con más de un nodo son:

- *Closeness*: 0.4167 (considerando los nodos aislados)
- *Betweenness*: 0

No debemos olvidar que **la red de clústeres es una red cíclica.**

**8 - Busca los cinco nodos más importantes en la red y los cinco nodos con una interpretación más diferente en función del criterio de centralidad usado (1.5 puntos).**

Para la realización de este ejercicio he manejado los dos criterios de centralidad vistos en clase:

- *Closeness*: Ya lo he manipulado en varias ocasiones, nos da una idea del nodo que tiene a menor distancia al resto de nodos de la red, es decir, nos indica que nodo se encuentra más próximo al resto de nodos de la red. En nuestro caso sería el aeropuerto desde el que es más sencillo llegar a cualquier otro. Es el criterio de relevancia que tendría en consideración un viajero.
- *Betweenness*: Nos dice cuantos de los caminos más cortos entre nodos pasan por cada nodo. En nuestro caso sería el aeropuerto que presenta mayor tráfico aéreo. Es el criterio de relevancia que tendría en consideración una empresa a la hora de elegir en que aeropuerto montar una tienda.

Para determinar los nodos más relevantes de la red he elegido aquellos cuyo promedio *Closeness - Betweenness* era mayor, obteniendo: Chicago O'hare Intl, Anchorage Intl, San Francisco Intl, Seattle-Tacoma Intl y Lambert-St Louis Intl.

Para los nodos con una interpretación más diferente al principio pensé en considerar la diferencia en valor absoluto entre las posiciones que ostentaban empleando las medidas de centralidad como criterios de ordenación, sin embargo, noté que 122 nodos presentan *Betweenness* igual a 0, lo que implicaba que su posición variaba mucho, requiriendo algún criterio adicional. Por ello me decidi a escalar los valores de las 2 medidas de centralidad y extraer la diferencia en valor absoluto. De esta forma aquellos con una diferencia mayor pueden considerarse más difíciles de interpretar. Y he obtenido: Mc Caran Intl, General Edward Lawrence Logan, Cleveland-Hopkins Intl, Cincinnati/Northern Kentucky I y Kansas City Intl. Todos ellos presentando un valor elevado de *Closeness* pero un valor bajo de *Betweenness*.

**9 - Indicar la medida de centralidad que nos parece mejor en este caso y la motivación de esta opinión en esta red (0.5 puntos).**

Como ya se ha indicado en el ejercicio anterior cada medida de centralidad obedece a unos criterios, en función de lo que vayamos buscando elegiremos una u otra. En mi caso, como estoy más cerca de la posición de un viajero que de la del propietario de una empresa, creo que es mejor emplear el criterio *Closeness*. Desde mi posición lo más relevante me parece conocer desde que aeropuertos tengo más cerca cualquier otro punto de la red, suponiendo que viajase a los EEUU. Además el *Betweenness* no nos permite discriminar entre 122 nodos (casi el 50%), ya que vale para todos ellos 0, luego incluso atendiendo a criterios más puramente funcionales *Closeness* es la medida de centralidad que realmente nos puede permitir establecer una jerarquía sobre toda la red.

**10 - Para los tres nodos indicados para cada uno de vosotros, mira su centralidad en general en la red (de forma comparativa con el resto de la red), en su cluster y para intermediar entre los distintos cluster (1.5 puntos).**

Para determinar la capacidad de intermediación de cada nodo con el resto de clústeres he empleado las redes *shrinked* en las que el clúster que contiene a los nodos aparece disperso y el resto contraídos. He calculado la medida de centralidad *Betweenness* en estas redes y mirado su valor para los nodos seleccionados.

Para la comparativa lo que he hecho ha sido mirar la posición que ocupaban en el *ranking* mis nodos, si establecía el ranking según *closeness*, *betweenness*, *media* (de *closeness* y *betweenness*) y *diferencia* (en términos absolutos entre *closeness* y *betweenness* escalados), para los nodos de toda la red, y para los nodos de la red asociada al clúster y para la red *shrinked* del clúster únicamente he mirado el *ranking* según *betweenness*. Los resultados los he recogido en la siguiente tabla:

Y con ellos ya podemos hacer el análisis que se nos pide:

Table 1: Posiciones en los Rankings

	Todos los nodos				Nodos del clúster				Shrinked bet
	clo	bet	med	dif	clo	bet	med	dif	
Chicago O'hare Intl	1	1	1	1	1	1	1	1	1
Klamath Falls Intll	244	126	245	58	81	60	82	16	50
Erie Intl	202	277	205	98	21	21	21	13	24

- **Chicago O'hare Intl:** Es claramente el aeropuerto más relevante, Esto ya lo vimos en el [ejercicio 8](#), aquí lo corroboramos. Es el aeropuerto más relevante segun *closeness* y *betweenness*, luego también según la media y la diferencia, y por supuesto también según *betweenness* en la red *shrinked*.
- **Klamath Falls Intll:** Este aeropuerto es bastante menos relevante. Si lo comparamos con todos los nodos se sitúa en la posición 244º según *closeness*, pero mejora bastante según *betweenness*, subiendo hasta la 126º. Esto implica que se trata de un aeropuerto relevante en cuanto a tráfico aéreo, pero que no está especialmente próximo a los demás, lo que nos lleva a pensar en un nodo de abundantes conexiones pero con nodos poco relevantes. Si miramos la media, su posición es aún peor que la peor según los dos criterios anteriores y en cuanto a la diferencia de posiciones lo llevan hasta la posición 98º. Dentro de su clúster no cambia de naturaleza. Pertenece al segundo clúster, que cuenta con 96 nodos y apenas hay 15 que estén en peor lugar. Se trata pues de un nodo de los que más alejado está del resto de la red, con una capacidad de intermediar promedio, ambas lo convierten en un nodo de difícil interpretabilidad (según las medidas de centralidad). En cuanto a su capacidad de intermediar con otros clústers ocupa una posición también promedio, mejorando un poco su capacidad con respecto a la medida intraclúster.
- **Erie Intl:** Este es un nodo poco relevante, se mire por donde se mire. Si atendemos a su *closeness* se sitúa en la posición 202º y si miramos *betweenness* es también muy baja, (aparece 277º, sabemos que entre esta y la 169º no hay diferencia), luego es un nodo alejado y con ninguna capacidad de intermediación. La media es consecuencia de las dos posiciones ya comentadas y su interpretabilidad, al tener posiciones bajas (bajas = malas) según ambos criterios se sitúa en la posición 98º, luego podemos seguir considerándolo un nodo sencillo de interpretar. Dentro de su clúster, que es el tercero y cuenta con 30 nodos, se sitúa en torno al percentil 67º. Comparte la coherencia posicional según *closeness* y *betweenness*, en la misma posición según *closeness*, *betweenness* y la media. Esto lo sitúa en la posición 13º en cuanto a interpretabilidad (que podría ser la 6º, ya que en ese tramo hay varios nodos con el mismo valor), siendo de los nodos más sencillos de interpretar de su clúster. Finalmente en cuanto a su capacidad de intermediación en la red *shrinked* es bastante intermedia, situándose en la posición 24º de 41 nodos.

## Comentarios

Considero que si se va a poner un ejercicio muy dependiente del manejo de un software tan específico (Pajek no ha sido usado en ningún otro módulo) se debe incorporar mayor información sobre el software, o al menos fuentes donde encontrarla (yo he encontrado algunos de cierta utilidad como [este](#) o [este](#) pero me han sido insuficientes). En otro caso lo que se está evaluando es nuestro manejo del software, o nuestra memoria (lo capaces que seamos de recordar lo que ocurrió en clase) y no nuestros conocimientos sobre el campo sobre el que trata el módulo.

Aún así haciendo el ejercicio me he familiarizado con los conceptos fundamentales del análisis de redes, aunque haya consumido mucho más tiempo el familiarizarme con Pajek.