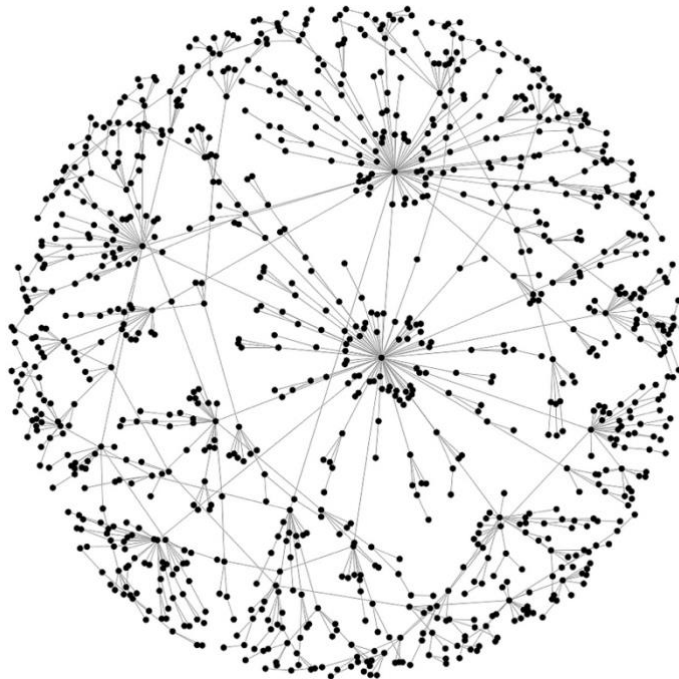


ANÁLISIS DE REDES SOCIALES

PRÁCTICA 2

REDES ALEATORIAS Y REDES LIBRES DE ESCALA



GRUPO 7

Juan Luis Romero

Carlos Piña

Alberto Fernández-Baíllo

Tabla de contenido

INTRODUCCION	2
PROPIEDADES DE UNA RED	2
ANÁLISIS PARA REDES ALEATORIAS	2
ANÁLISIS DE REDES LIBRES DE ESCALA	2
ANÁLISIS DE LAS REDES CON DISTINTOS MODELOS.....	3
DISTRIBUCION DE GRADOS PARA CADA UNA DE LAS REDES DE CADA MODELO	4
RED ALEATORIA.....	4
RED LIBRE DE ESCALA.....	8
GRAFOS OBTENIDOS CON GEPHI	10
ANALISIS DE LAS REDES	13
ESTUDIO DE LA RELACION ENTRE LAS REDES ESTUDIADAS Y LA PRÁCTICA ANTERIOR.....	14

INTRODUCCION

- En este documento vamos a realizar un estudio sobre las redes de Erdős-Renyi y de Barabasi-Albert, además de un estudio sobre el grafo realizado en la práctica anterior sobre las interacciones entre personajes de Los Simpsons.

PROPIEDADES DE UNA RED

ANÁLISIS PARA REDES ALEATORIAS

- **Hubs:** En una red aleatoria no existen Hubs, ya que es una red tan aleatoria que la poca interconectividad que pueda existir en la red no permite que existan nodos con grados altos.
- **Densidad de la red:** Cuanto mayor sea el número de nodos menor será la conectividad en las redes críticas, subcríticas y supercríticas. Al contrario, en las redes de conectadas a mayores nodos, mayor densidad.
- **Coeficiente de clustering:** En las etapas críticas y subcríticas el valor será nulo, debido a la aleatoriedad de la red. En la etapa supercrítica disminuye el valor a medida que aumentamos el número de nodos. En la red conectada el coeficiente puede llegar a establecerse.
- **Distancia media:** En la fase subcrítica, la distancia disminuye con el número de nodos. En la fase crítica, la distancia disminuye hasta llegar a 5000 nodos. En la fase supercrítica, a mayor número de nodos, mayor distancia. En la fase conectada, la distancia se mantiene mas o menos estable para todos los nodos.

ANÁLISIS DE REDES LIBRES DE ESCALA

- **Hubs:** En estas redes si existen Hubs, y cuanto más grande sea la red mayor será el número en que estos aparecen. Tienen la característica “*rich get richer*”, lo que quiere decir que cuantas más aristas se unen a un nodo, más probabilidad tiene éste de que se le unan más aristas nuevas.
- **Densidad de la red:** Cuanto mayor sea el tamaño de la red menor será el valor, porque la red crece de manera exponencial.
- **Coeficiente de clustering:** Cuanto mayor sea el tamaño de la red, menor será el valor, debido a que los nodos estarán menos interconectados.
- **Distancia media:** Aumenta ligeramente cuanto más grande sea la red ya que los nodos se alejan entre sí.

ANÁLISIS DE LAS REDES CON DISTINTOS MODELOS

- **Hubs:** Podemos comprobar la diferencia entre ambos tipos de redes ya que en el modelo de red aleatoria no existen hubs, ya que en este tipo de algoritmo las aristas se unen basándose únicamente en una probabilidad, sin importar a qué parejas de nodos conecten, mientras que en el modelo de red libre de escala si influye cuántas aristas conecten a un nodo.

La principal diferencia es que en las redes aleatorias, a priori, no podemos saber el grado medio con tanta exactitud (a pesar de las estimaciones del modelo teórico) mientras que en las redes libres de escala si podemos controlar el número de enlaces de cada nodo nuevo, obteniendo así un gran salto, en lo que al grado de nodo se refiere, entre las distintas etapas de redes aleatorias, y siendo un salto más suave en las libres de escala dependiendo del número de enlaces de cada nuevo nodo.

- **Densidad de red:** Observamos que, en general, la densidad disminuye al aumentar el tamaño, esto es lógico ya que los nodos de las redes tienden a separarse. Ambas redes suelen tener valores parecidos excepto en la etapa conectada de las redes aleatorias, donde la densidad es significativamente mayor, mientras que en las redes libres de escala la densidad depende del número de enlaces de los nuevos nodos.
- **Coeficiente de Clustering:** En los dos tipos de redes el coeficiente de agrupamiento se mantiene bastante bajo ya que hay pocas aristas que interconecten los nodos, sobre todo en las primeras etapas de las redes aleatorias. La excepción es la etapa conectada, en la cual la probabilidad de que sus nodos se unan es bastante más alta.
- **Distancia media:** Como hemos comentado anteriormente, ambos tipos de redes tienden a separarse según crecen, lo que conlleva que se formen caminos más largos y aumente la distancia media, concordando así con la densidad de la red y el coeficiente de clustering.

DISTRIBUCION DE GRADOS PARA CADA UNA DE LAS REDES DE CADA MODELO

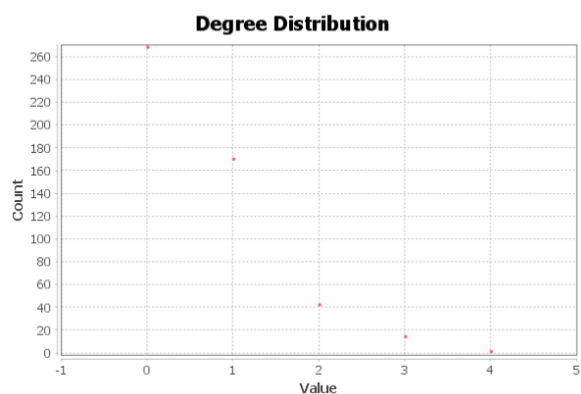
RED ALEATORIA

ETAPA SUBCRITICA

500 nodos

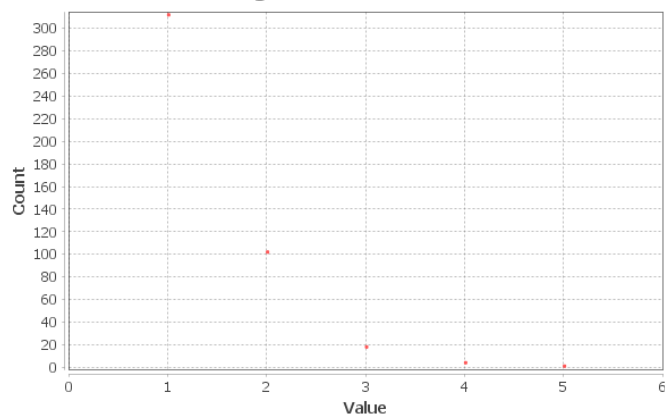
Results:

Average Degree: 0,620



1000 nodos

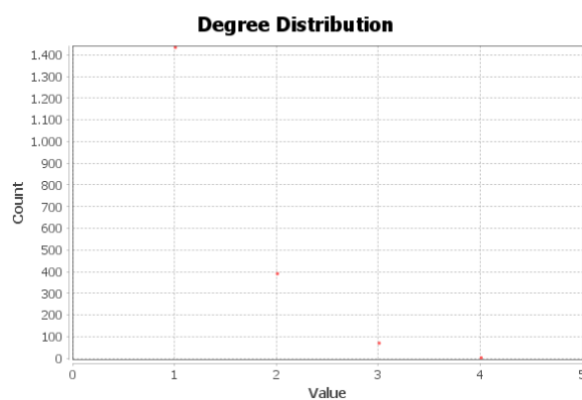
Degree Distribution



5000 nodos

Results:

Average Degree: 1,295



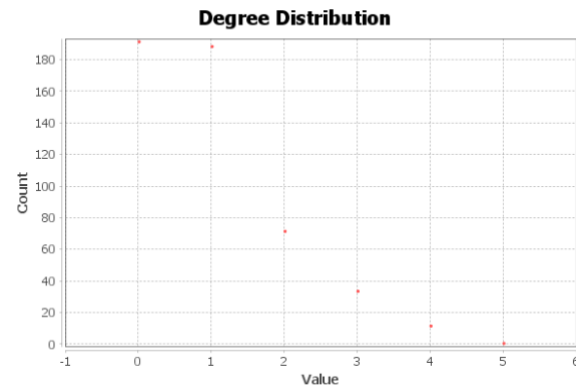
Debido a la baja probabilidad que caracteriza a esta etapa, podemos comprobar que la mayoría de nodos tienen un grado bajo.

ETAPA CRITICA

500 nodos

Results:

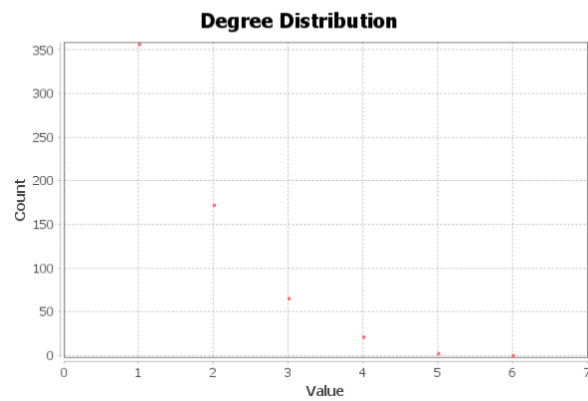
Average Degree: 0,976



1000 nodos

Results:

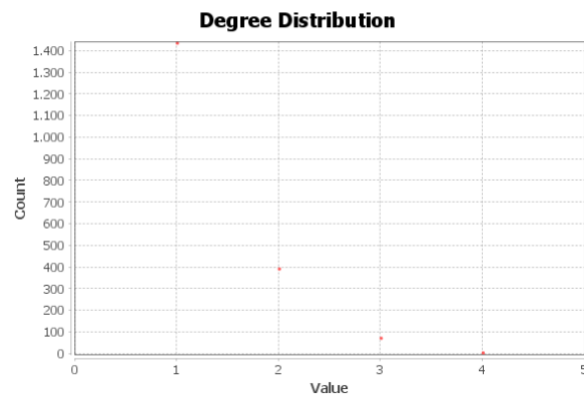
Average Degree: 1,624



5000 nodos

Results:

Average Degree: 1,295



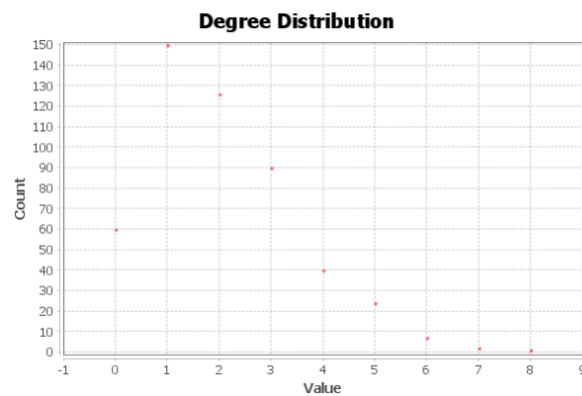
Igual que pasaba en la etapa subcrítica, la mayor parte de los nodos tienen un bajo grado debido a su baja probabilidad, que apenas varía respecto a la etapa anterior.

ETAPA SUPERCRITICA

500 nodos

Results:

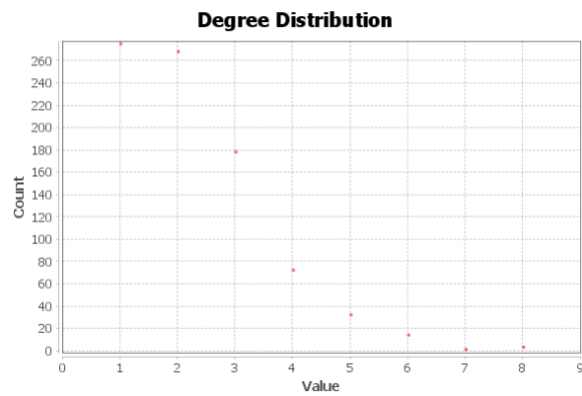
Average Degree: 2,032



1000 nodos

Results:

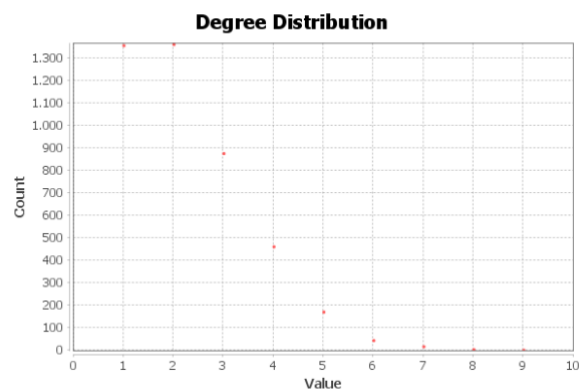
Average Degree: 2,284



5000 nodos

Results:

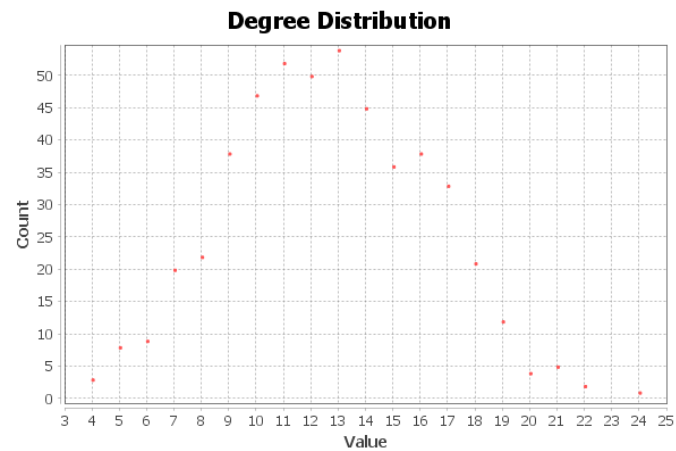
Average Degree: 2,290



Respecto a la etapa supercrítica, podemos observar que el grado de los nodos comienza a alejarse de los valores más bajos, aunque seguimos teniendo pocos nodos con un grado alto.

ETAPA CONECTADA

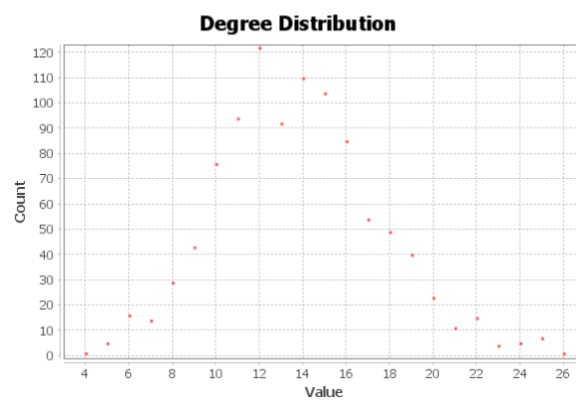
500 nodos



1000 nodos

Results:

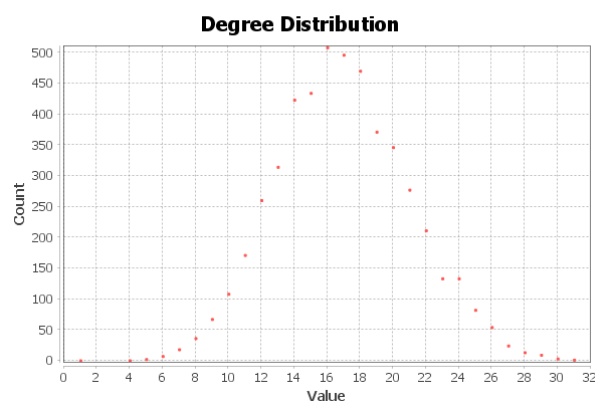
Average Degree: 13,750



5000 nodos

Results:

Average Degree: 16,940



Dado que en la etapa conectada la probabilidad de unión crece notoriamente, el grado de los nodos se distribuye en torno al grado medio de la red, teniendo pocos nodos tanto de grado bajo como alto.

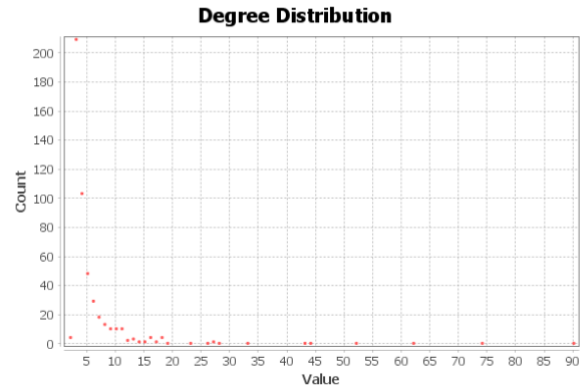
RED LIBRE DE ESCALA

CON 3 ENLACES

500 nodos

Results:

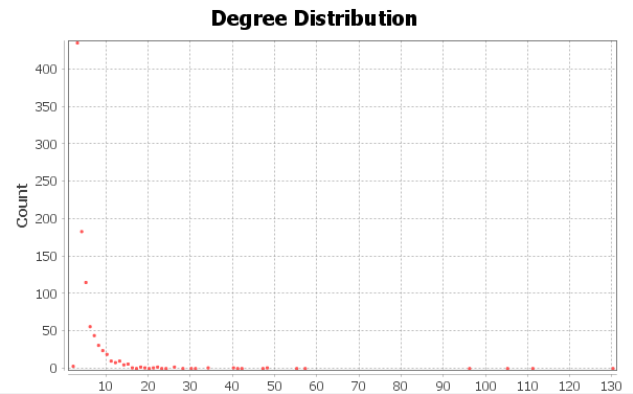
Average Degree: 5,908



1000 nodos

Results:

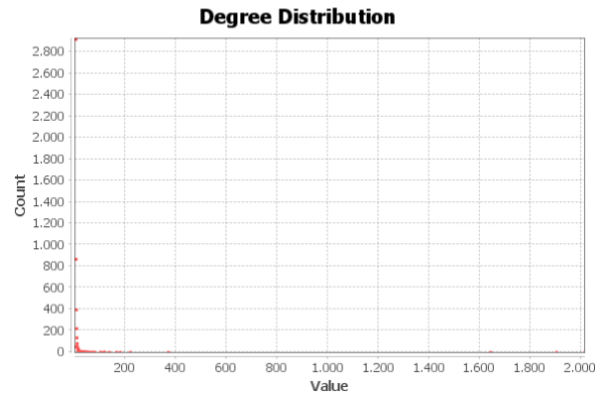
Average Degree: 5,942



5000 nodos

Results:

Average Degree: 5,958

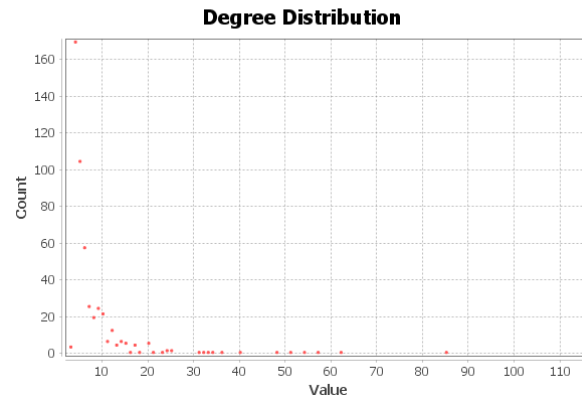


CON 4 ENLACES

500 nodos

Results:

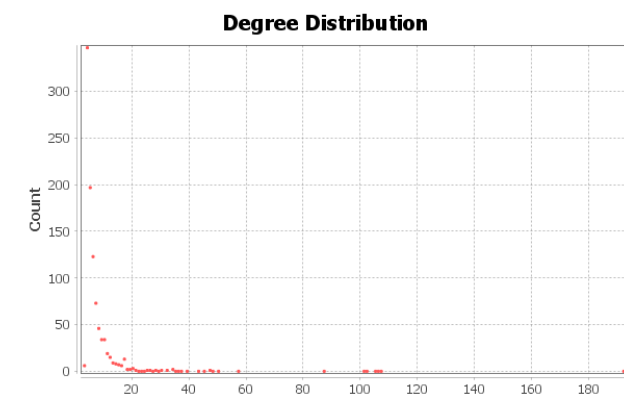
Average Degree: 7,796



1000 nodos

Results:

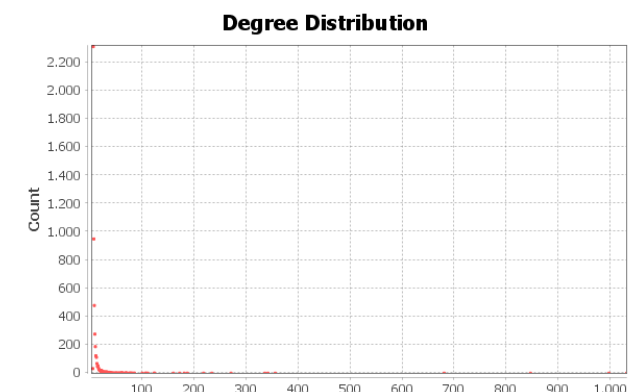
Average Degree: 7,870



5000 nodos

Results:

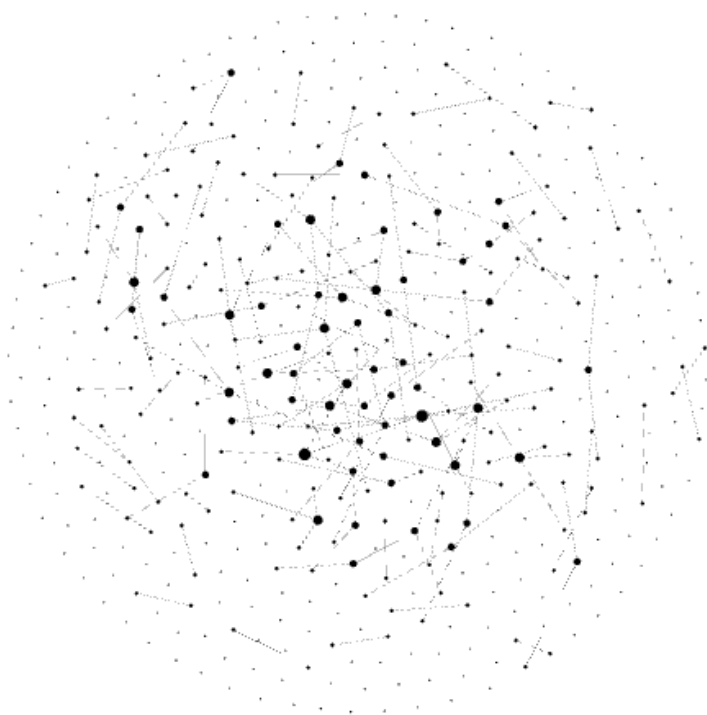
Average Degree: 7,960



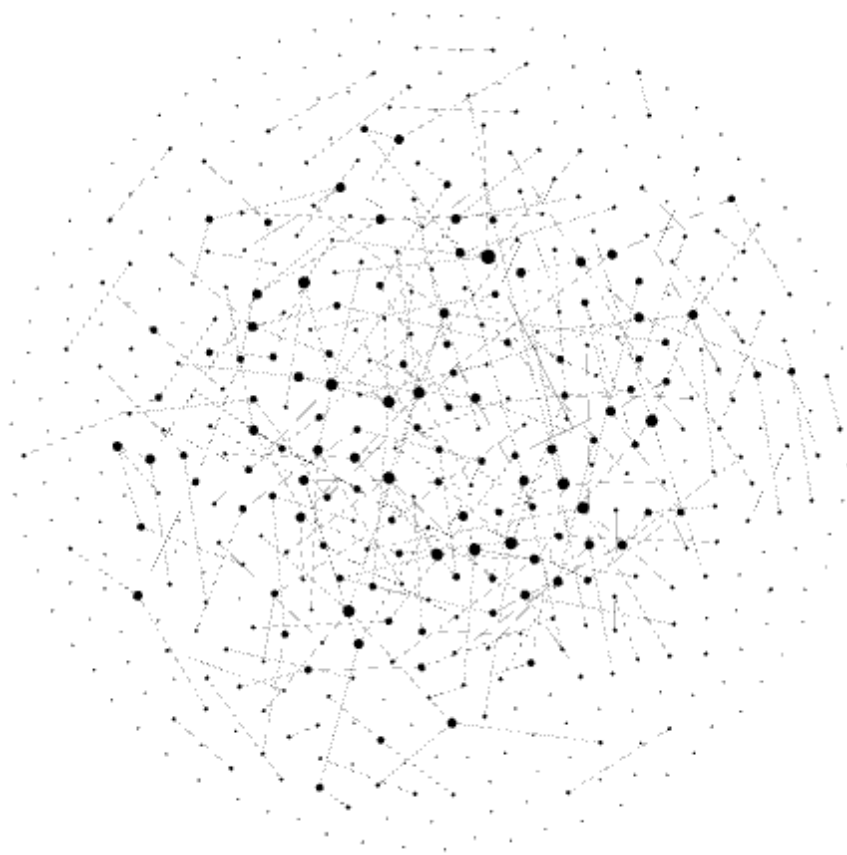
Común a todas las redes libres de escala generadas, el grado de la mayoría de los nodos gira en torno al grado medio, obteniendo curvas mucho más pronunciadas y definidas cuanto mayor es el número de nodos. Podemos observar una minoría de nodos con un alto grado, es decir, los hubs de la red.

GRAFOS OBTENIDOS CON GEPHI

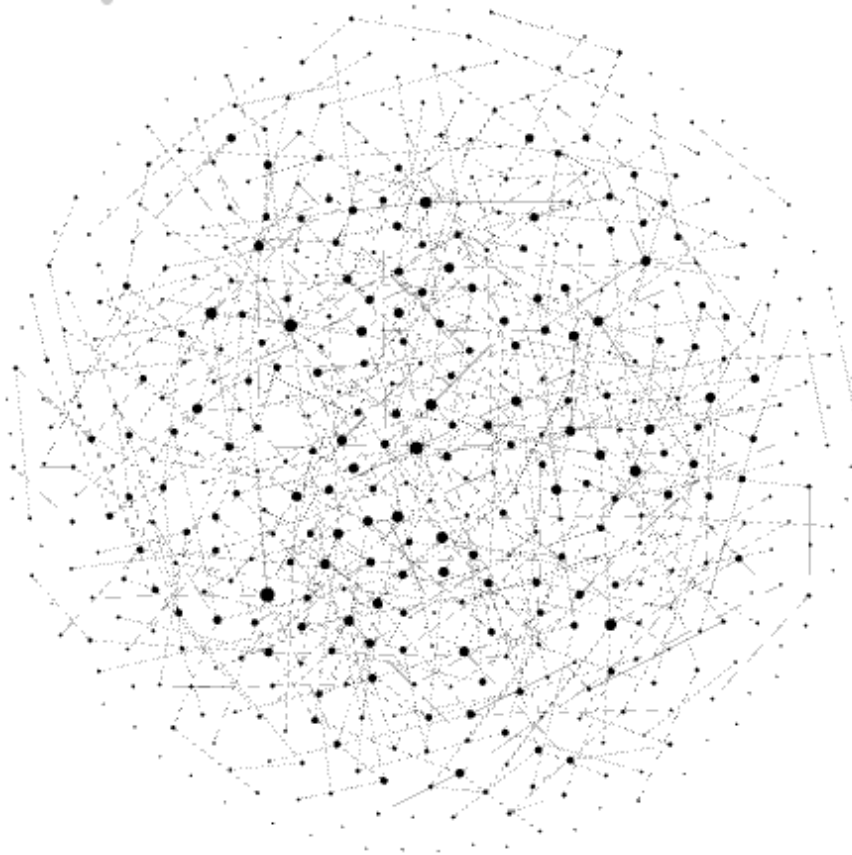
- RED DE 500 NODOS ETAPA SUBCRITICA



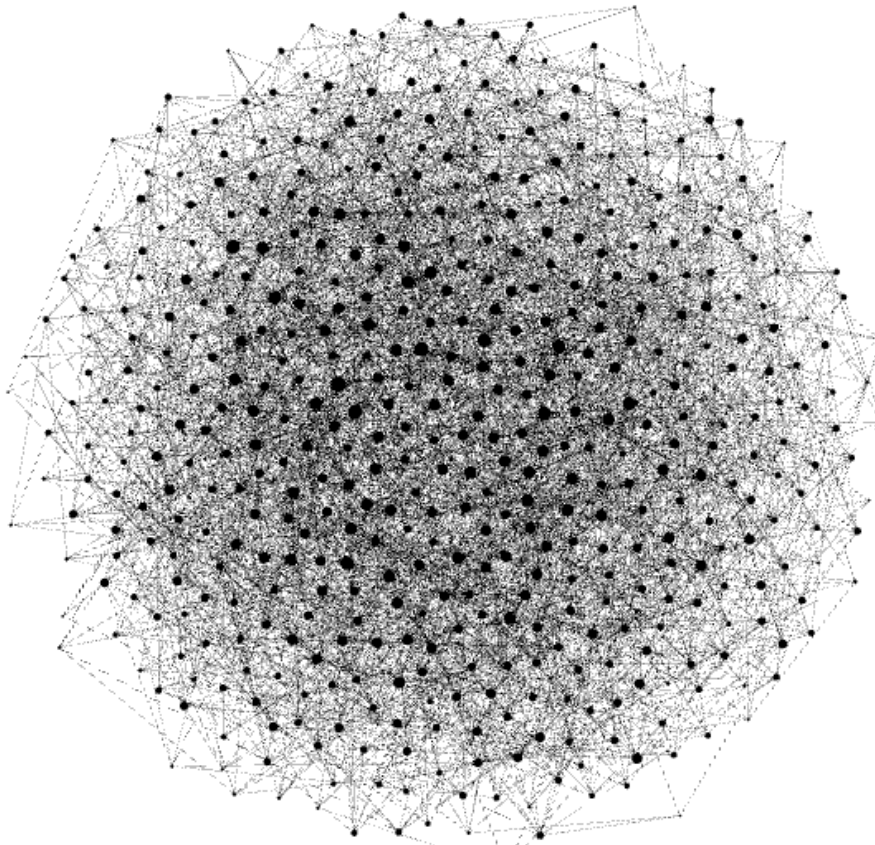
- RED DE 500 NODOS ETAPA CRITICA



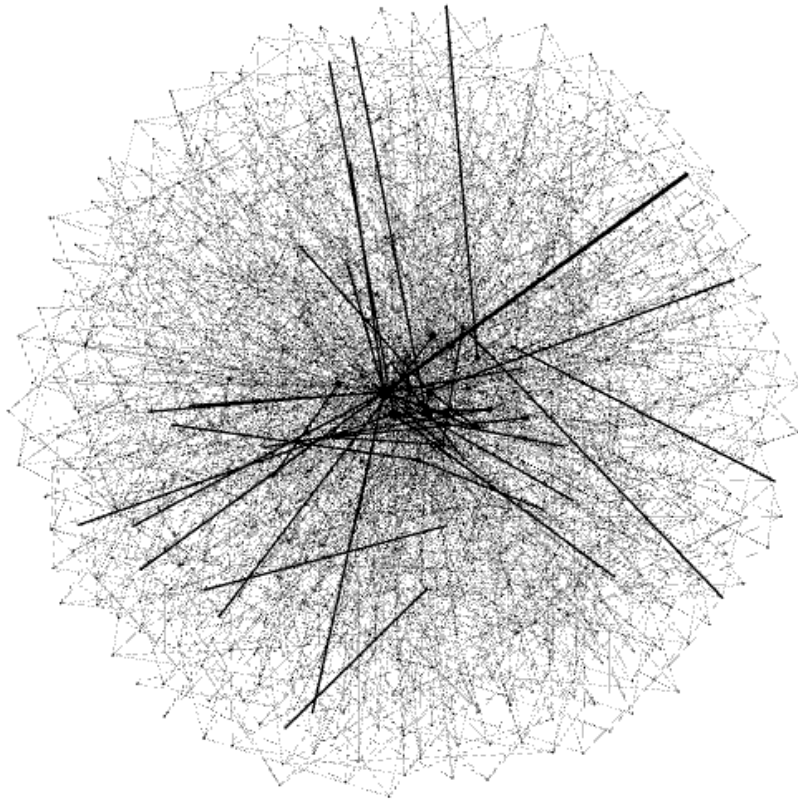
- RED DE 500 NODOS SUPERCRTICA



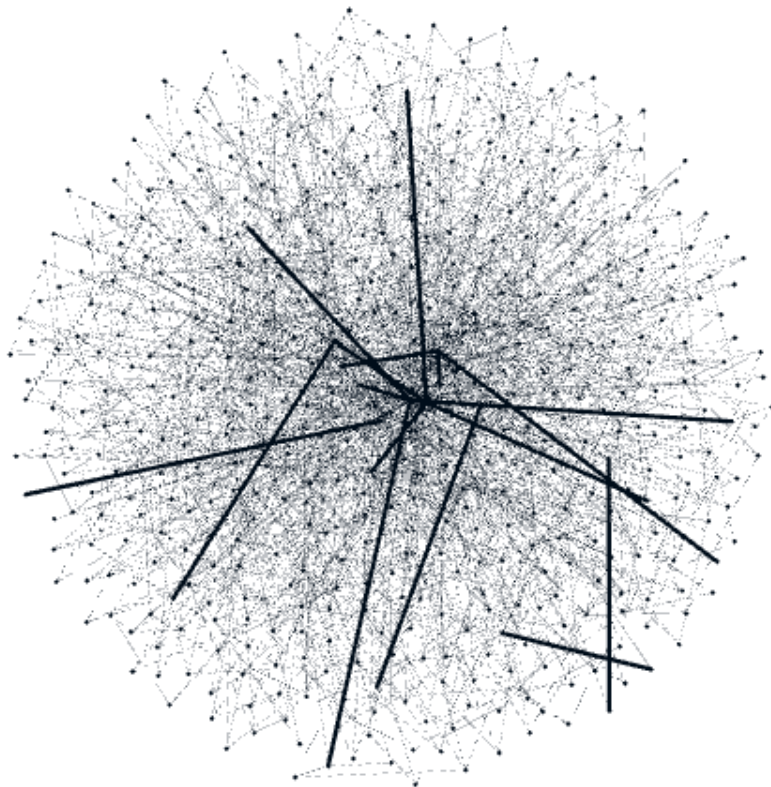
- RED DE 500 NODOS CONECTADA



- RED DE BARABASI DE 500 NODOS CON 4 ENLACES



- RED DE BARABASI DE 500 NODOS CON 3 ENLACES



ANALISIS DE LAS REDES

1. COMPARACION ENTRE GEPHI Y DATOS TEORICOS EN REDES ALEATORIAS

- **Hubs:** Como podemos observar, analizando las redes vemos que la diferencia entre la cantidad de nodos es existente, pero no lo suficiente como para considerarlos Hubs de la red.
- **Densidad de la red:** Es prácticamente igual tanto en los datos calculados de manera teórica, como los datos calculados mediante Gephi
- **Distancia media de las redes:** Podemos destacar que varía mucho en las etapas crítica y subcrítica ya que de manera teórica obtenemos resultados negativos. Sin embargo, los resultados de las redes supercríticas y conectadas son prácticamente iguales.
- **Valores medios de clustering:** En todos los tipos de redes son muy próximos a ceros. Hay resultados obtenidos con Gephi que son completamente nulos, datos que teóricamente no se obtienen.

2. COMPARACION ENTRE GEPHI Y DATOS TEORICOS EN REDES DE LIBRE ESCALA

- **Hubs:** En este tipo de redes si existe el Hub de red, ya que las diferencias de los nodos obtenidos de ambas maneras son grandes, sobre todo en las redes de 5000 nodos tanto de 3 enlaces como de 4 enlaces.
- **Distancia media:** Observamos claramente como los resultados obtenidos de la distancia media son el doble que los obtenidos con los de Gephi, con lo cual difieren en todos.
- **Densidad:** En este caso la densidad obtenida de forma teórica como mediante Gephi es exactamente la misma.
- **Distancia media de clustering:** Observamos que las redes con 3 y 4 enlaces de 500 y 1000 nodos son resultando muy similares. Sin embargo, en la red de 5000 nodos varía bastante.

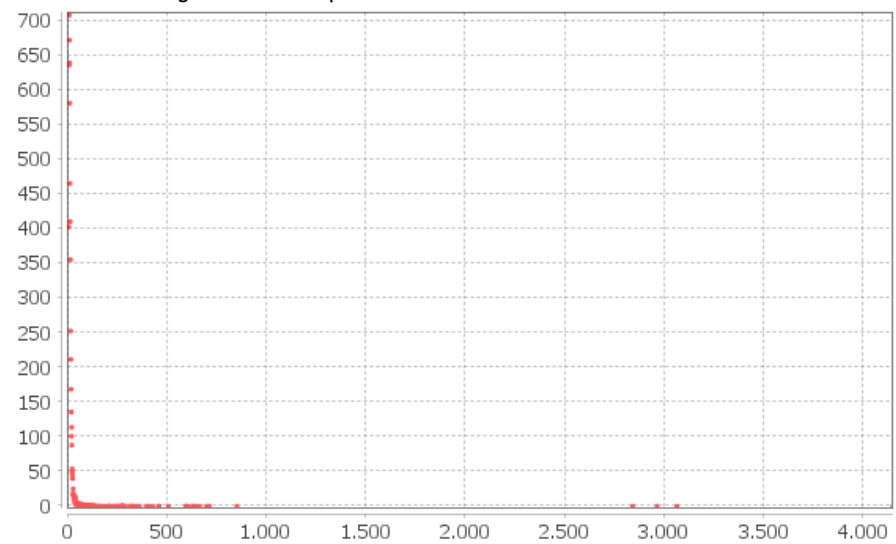
ESTUDIO DE LA RELACION ENTRE LAS REDES ESTUDIADAS Y LA PRÁCTICA ANTERIOR

LOS SIMPSONS

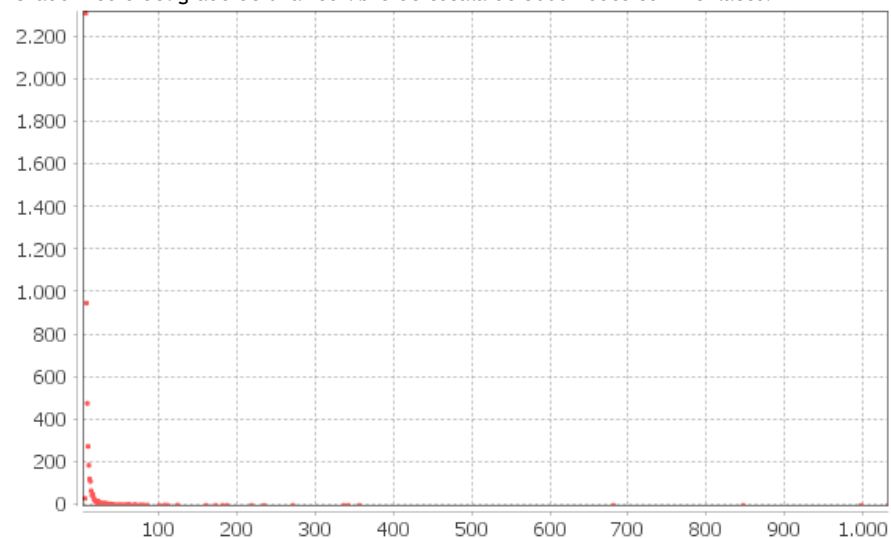
N	m	L		Density		Largest Hub		Avg. Distance		Average Clustering Coefficient	
		Teórico	Gephi	Teórico	Gephi	Teórico	Gephi	Teórico	Gephi	Teórico	Gephi
6566	6	39403	39418	0,0018282	0,002	486,18515	290	6,56168274	2,869	0,00267	0,905

Tras analizar los resultados de la practica 1 y todos los tipos de redes tanto aleatorias como libre de escala, nos hemos percatado de que la distribución de los grados de los nodos de la practica 1 es muy similar a una red libre de escala:

Grado medio del grafo de los simpson:

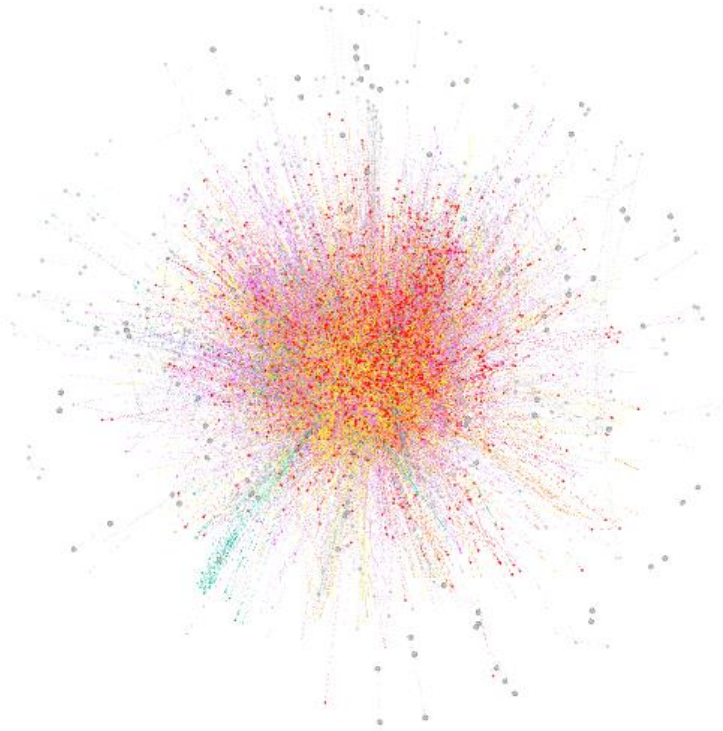


Grado medio del grado de una red libre de escala de 5000 nodos con 4 enlaces:

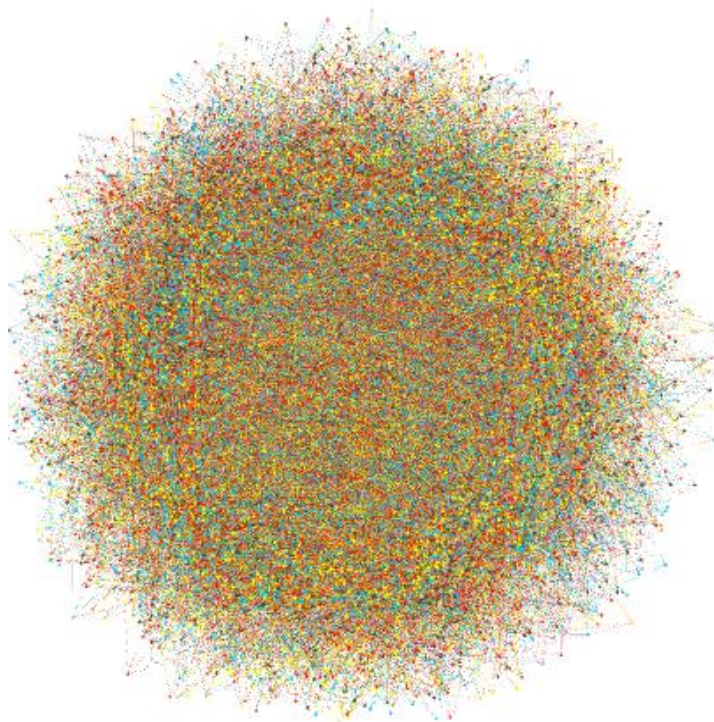


Vamos a analizar a continuación los datos obtenidos en Gephi de cada grafo.

- La red de los Simpson está formada por 6566 nodos y 41608 aristas.
- Para analizar el grafo de la red libre de escala usaremos 6566 nodos y una relación de 6 aristas, para que se asemeje al de los Simpson.
- Grafo de los Simpson:



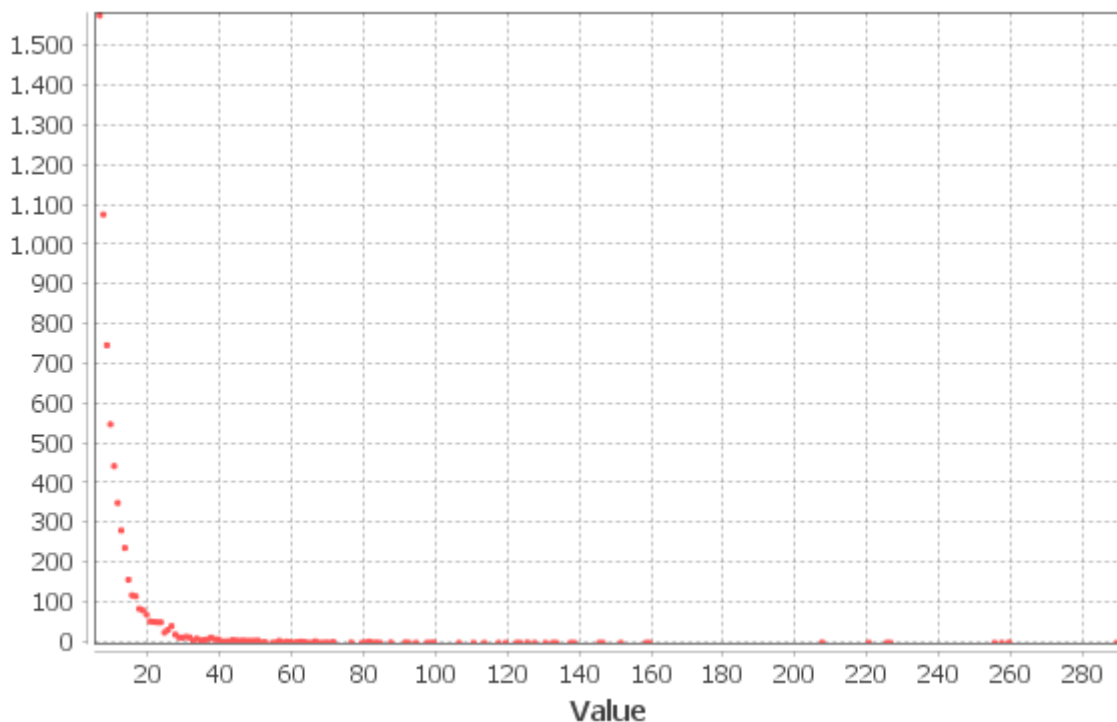
- Grafo de red libre de escala:



Podemos observar que los grafos son muy parecidos, vamos a comparar ahora los datos obtenidos:

	Grafo Simpson	Grado Red Libre de Escala
Grado Medio	6,337	5,997
Media Clustering	0,905	0,011
Densidad	0,002	0,002
Diámetro de la red	7	5
Longitud media	2,869	3,364
Modularidad	0,228	0,256

- Observando los resultados nos damos cuenta de que el Grado Medio, densidad y modularidad de ambos grafos es similar.
- Con los datos de la Media de Clustering nos percatamos de que el grafo de los Simpson tiene un valor muy alto, lo que indica que la red está muy conectada entre sí, mientras que, la red libre de escala prácticamente no tiene relación entre los nodos, y esto se debe a que la generación de enlaces es aleatoria.
- Gráfica del grado medio en la red libre de escala con 6566 nodos y 6 enlaces
- La longitud media de camino es más pequeña en el grafo de los Simpson debido a que los nodos están más relaciones que en la red libre de escala.



- Tras analizar todos estos datos podemos concluir que la red de los Simpson no responde a ninguno de los dos tipos de redes que estamos estudiando a pesar de que tenga cierta similitud con Barabasi-Albert. La única coincidencia que podemos destacar es la existencia en ambos grafos de Hubs, pero la interconectividad que hay entre estos nos hace pensar que el grafo de los Simpson responde a un modelo de red similar a una red libre de escala.
- Descartamos por completo que pueda pertenecer a una red de Erdős-Renyi, porque claramente los Simpson son los Hubs de la red y en una red aleatoria no hay Hubs, por lo tanto, no podríamos asegurar la relación entre personajes secundarios y personajes primarios.