

Session 7: Introduction to igraph

16-12-2022

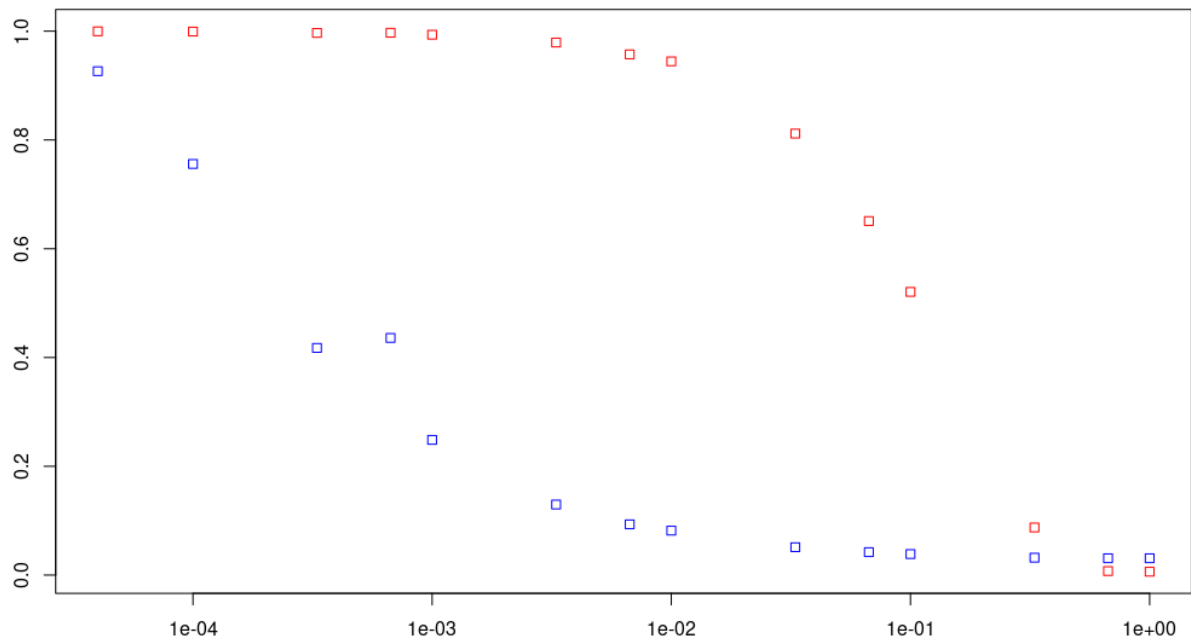
He Chen

Daniel Muñoz

Grupo 12

Tarea 1:

Para generar el gráfico visto en teoría, en el eje x, los valores de p van de 0.0001 a 1 en potencias de 10 y el eje de las y lo hemos normalizado para poder ver el coeficiente de clustering y el average shortest path en función del valor de p ($C(p)$ y $L(p)$ respectivamente). Para normalizar los valores entre 0 y 1, lo que hemos hecho, ha sido dividir los valores obtenidos con el valor obtenido cuando $p = 0$, es decir $C(p)/C(0)$ y $L(p)/L(0)$, tal y como indicaba en el enunciado del laboratorio. Finalmente, hemos obtenido los siguientes valores generando un grafo de 2000 vértices y 8000 aristas.



eje X: la probabilidad de rewiring
eje Y representa el coeficiente de clustering(rojo) y average shortest path(azul) de los grafos generados con dichas probabilidades.

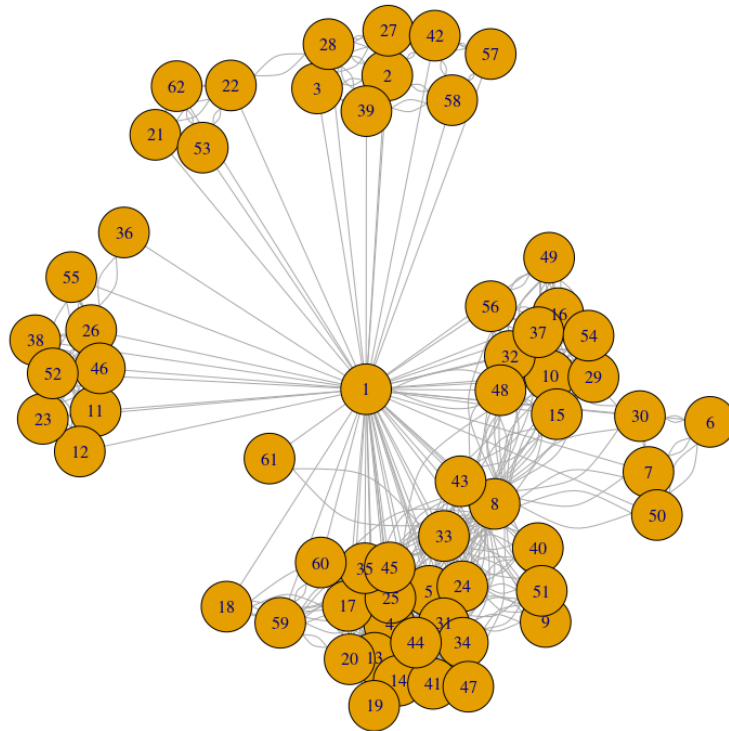
Podemos observar que el average shortest path sigue una distribución logarítmica. Empieza en 1 y va disminuyendo rápidamente hasta mantenerse en 0.02 aproximadamente cuando la probabilidad es 0.01 o más.

En cuanto al coeficiente de clustering, vemos que se mantiene hasta llegar a una probabilidad de 0.01, cuando empieza a caer en picado.

Tarea 2:

Describe the network a little. How many edges and nodes does it have? What is its diameter? And transitivity? And degree distribution? Does it look like a random network? Visualize the network with node sizes proportional to their pagerank.

Visualización de la red:

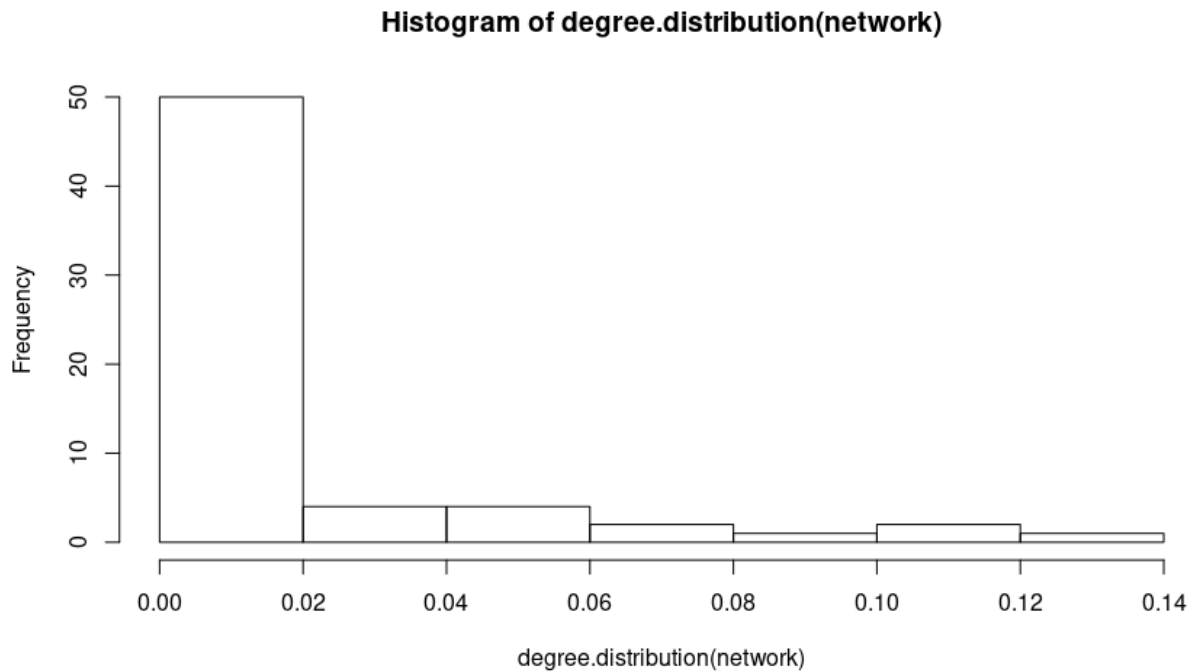


Num aristas	Num nodos	Diámetro	Transitividad
602	62	2	0.52

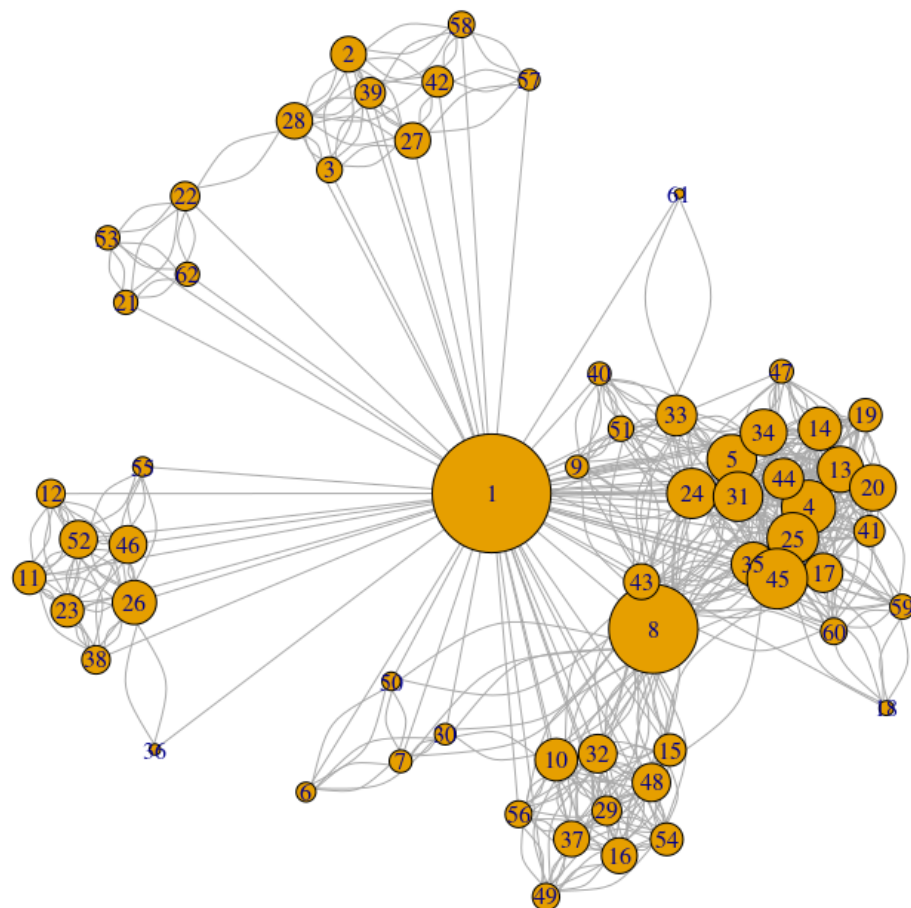
Grado de nodos: [63 13 9 39 35 7 9 59 13 25 13 11 33 31 19 21 27 7 23 33 7 9 13 35 37 17 13 13 17 9 35 23 25 33 31 3 21 11 11 13 21 11 23 29 43 15 15 23 15 7 15 15 7 19 7 15 7 9 15 17 3 7]

Como podemos observar, tenemos 602 aristas y 62 nodos y a pesar de no ser un grafo completo ya que tendría $(62 * 61)/2 = 1891$ aristas, pero solo tiene 602, podemos ver que el diámetro es relativamente bajo. Esto es debido al fenómeno de mundo pequeño visto en teoría.

Distribución de los grados:

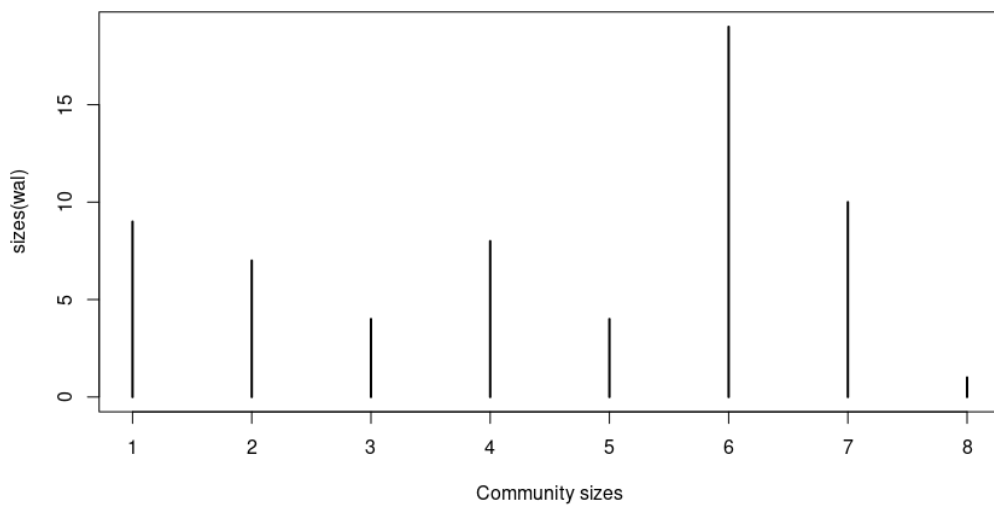


Podemos hacernos una idea de que la red no es una red aleatoria, ya que la distribución de los grados en una red aleatoria sigue una distribución normal y, como se puede observar, esta distribución no se adapta a una distribución normal, sino que se adapta a una exponencial.



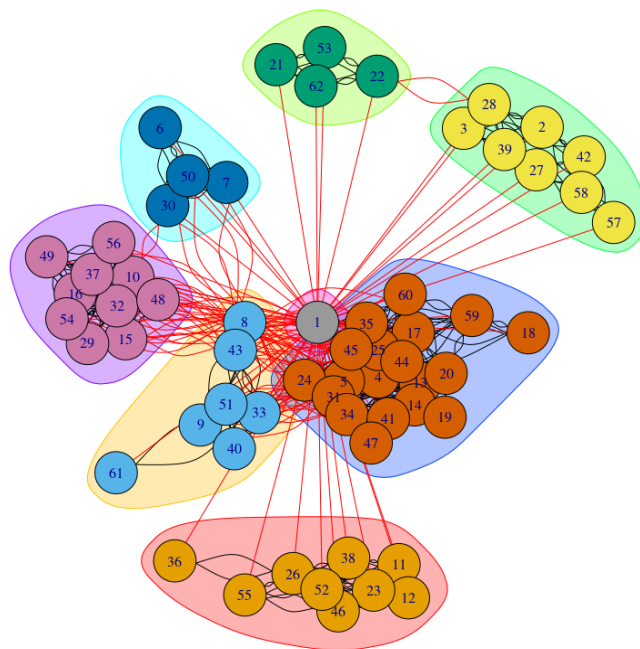
network with node size proportional to pagerank

En el gráfico anterior de pagerank podemos ver que hay un nodo central que hace de 'puente' con todos los demás nodos, también tiene un grado mucho más grande que otros nodos. Esta propiedad la podemos ver por ejemplo en el modelo de Barabasi-Albert, donde la probabilidad de que un nodo nuevo sea conexo con otro es proporcional a su grado, el resultado es que tenemos pocos nodos con grados muy elevados y los otros nodos con grados bajos y uniformes. Por lo tanto, este modelo de grafo, como podíamos intuir en la distribución de grados, no es aleatorio, sino que sigue el modelo de Barabasi-Albert.



Now, use a community detection algorithm of your choice from the list provided. How many nodes does the largest community found contain? Plot the histogram of community sizes. Plot the graph with its communities.

Se puede ver que el algoritmo ha separado el grafo en 8 comunidades diferentes con tamaños desde 1 a 18 nodos por comunidad, siendo 18 el mayor número de nodos que llega a tener una comunidad.



En este plot, podemos ver cómo están separadas las comunidades. Es fácil de ver que el factor clave está en las conexiones que hay entre los nodos. Los conjuntos de nodos con mayores conexiones internas forman comunidades. En particular, hay una comunidad formada con el nodo que tiene mayor pagerank (nodo 1), este nodo llega a formar por sí sola una comunidad, ya que está conectada con todos los otros nodos.