

Session 6: Introducing to *igraph*

Q1 2022-23



Integrantes:

Pol Pérez Castillo

Daniel Ruiz Jiménez

Fecha de entrega: 23-12-2022

Introducción

En esta última práctica vamos a familiarizarnos con el paquete software *igraph* con el objetivo de analizar redes. Hemos decidido usar Python para esta práctica en lugar de R, ya que nos parece un lenguaje más sencillo.

Task 1

Para la primera tarea se nos ha pedido reproducir un grafo a partir de una gráfica, y se nos ha comentado que se trata del modelo Watts-Strogatz. En el eje de las x tendremos la probabilidad (con escala logarítmica) de reconexión de aristas (rewiring), y en el eje de las y tenemos, en cuadrados amarillos, el coeficiente de clustering del modelo, y en círculos rojos, el average shortest path, expresado en valores normalizados (entre 0 y 1).

Una vez ejecutado el programa, hemos ido ajustando los valores de la función Watts_Strogatz, del paquete *igraph*, y hemos ido cambiando el tamaño del grafo y el número de vecinos, y nuestro intento con más éxito fue con $\text{size} = 6000$ y $\text{n}^\circ\text{vecinos} = 4$ (este 4 sacado de las transparencias de la asignatura). Obtenemos el gráfico siguiente:

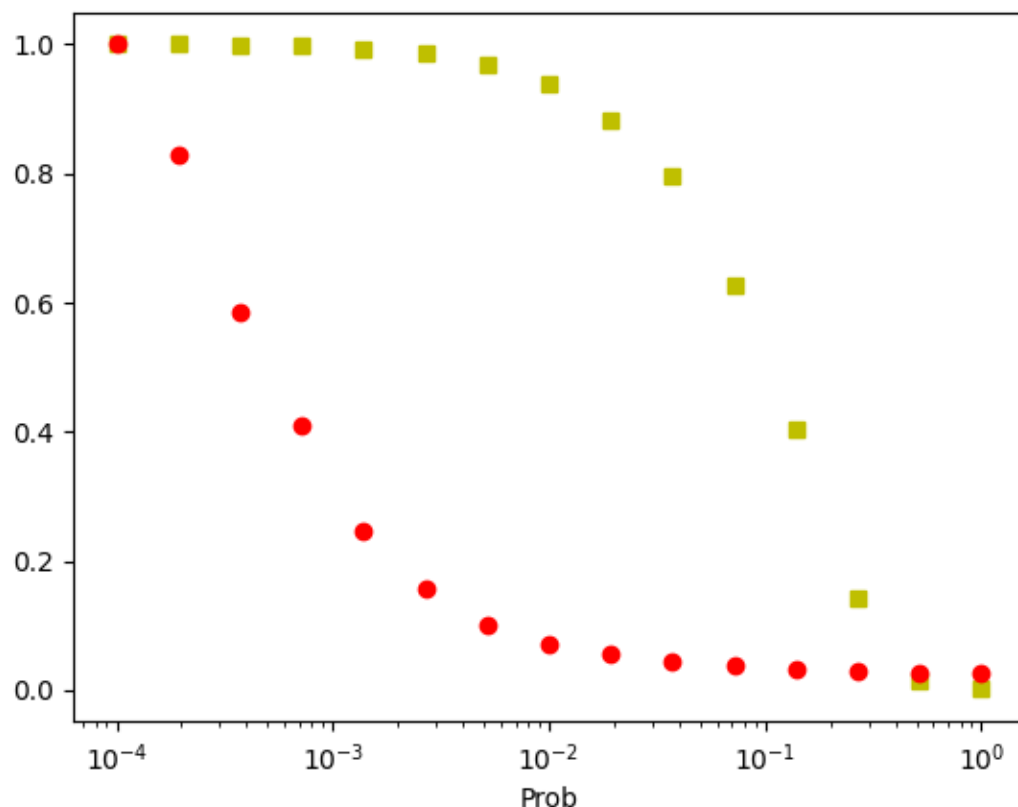


Imagen 1: Gráfico del modelo Watts-Strogatz

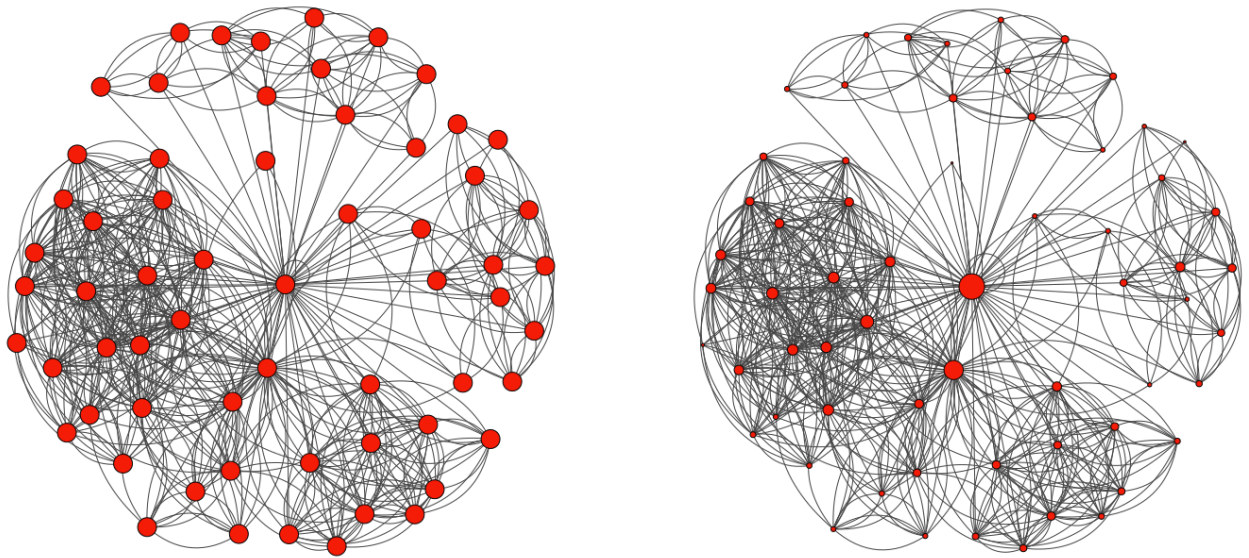
Observamos, tal y como explica dicho modelo, al aumentar la probabilidad de rewiring, observamos un decremento rápido del average shortest path y un decremento lento del coeficiente de clustering.

Task 2

En el segundo ejercicio se nos pide describir diversos aspectos de una red cuyas aristas se nos proporcionan en el fichero *edges.txt*.

Gracias al paquete *igraph*, la tarea se reduce simplemente a imprimir y a analizar varios valores de la clase Graph. Más en concreto, hemos obtenido los datos siguientes:

El grafo tiene 602 aristas y 62 nodos. Tiene diámetro 2 y transitividad 0,52. Su degree distribution es el de una distribución normal de media 19,42 y desviación típica 12,53.



Imágenes 2 y 3: Distribución de nodos de la red y su pageRank

Abajo tenemos la distribución de grados de los nodos y el grado de cada nodo:

```
[ 3,  4): ** (2)
[ 4,  5): (0)
[ 5,  6): (0)
[ 6,  7): (0)
[ 7,  8): ***** (8)
[ 8,  9): (0)
[ 9, 10): ***** (5)
[10, 11): (0)
[11, 12): ***** (4)
[12, 13): (0)
[13, 14): ***** (7)
[14, 15): (0)
[15, 16): ***** (7)
[16, 17): (0)
[17, 18): *** (3)
[18, 19): (0)
[19, 20): ** (2)
[20, 21): (0)
[21, 22): *** (3)
[22, 23): (0)
```

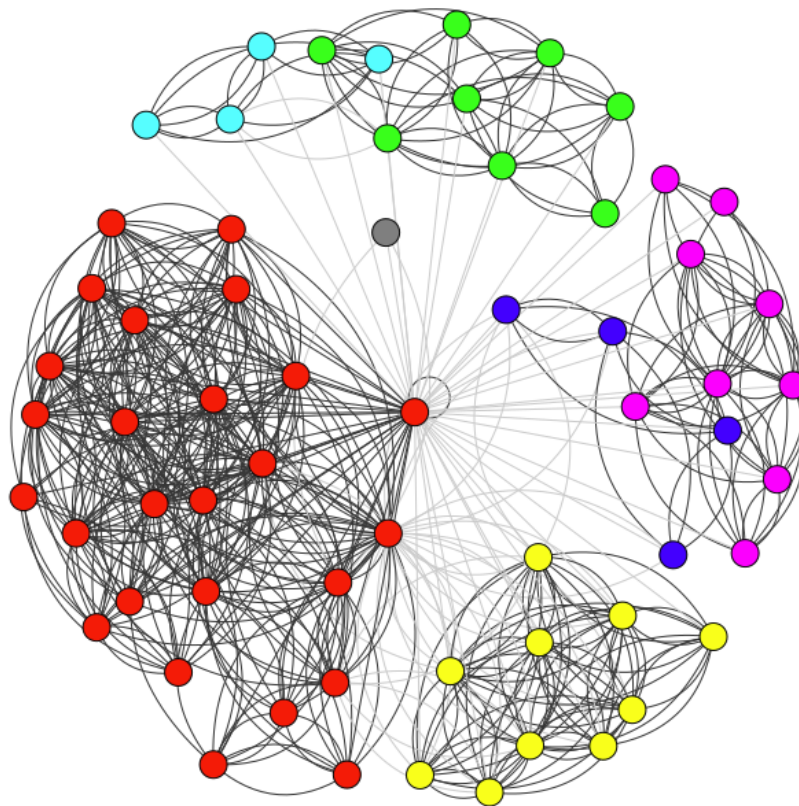
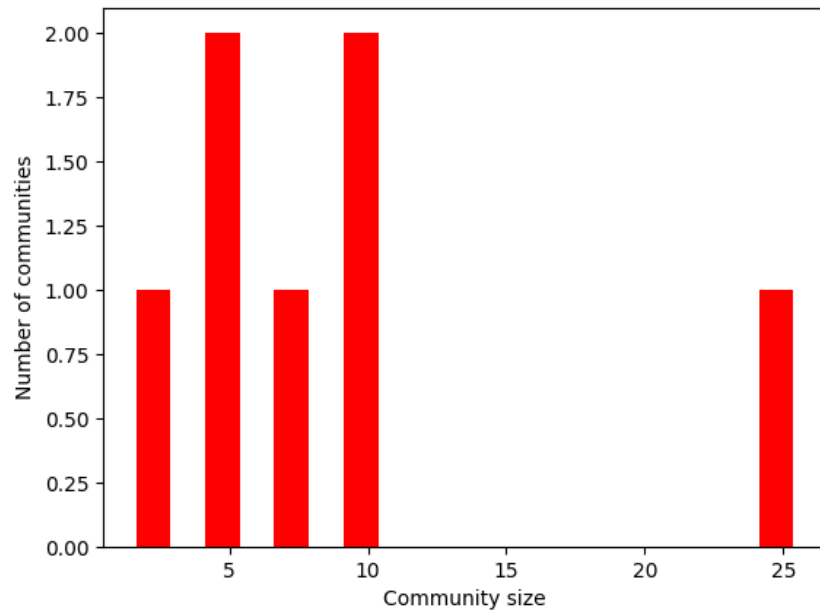
```

[23, 24): **** (4)
[24, 25): (0)
[25, 26): ** (2)
[26, 27): (0)
[27, 28): * (1)
[28, 29): (0)
[29, 30): * (1)
[30, 31): (0)
[31, 32): ** (2)
[32, 33): (0)
[33, 34): *** (3)
[34, 35): (0)
[35, 36): *** (3)
[36, 37): (0)
[37, 38): * (1)
[38, 39): (0)
[39, 40): * (1)
[40, 41): (0)
[41, 42): (0)
[42, 43): (0)
[43, 44): * (1)
[44, 45): (0)
[45, 46): (0)
[46, 47): (0)
[47, 48): (0)
[48, 49): (0)
[49, 50): (0)
[50, 51): (0)
[51, 52): (0)
[52, 53): (0)
[53, 54): (0)
[54, 55): (0)
[55, 56): (0)
[56, 57): (0)
[57, 58): (0)
[58, 59): (0)
[59, 60): * (1)
[60, 61): (0)
[61, 62): (0)
[62, 63): (0)
[63, 64): * (1)
degree: [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, 1
5, 7, 9, 15, 17, 3, 7]

```

Al observar sobretodo el grafo del pageRank, intuimos que no se trata de una red aleatoria, ya que existe un nodo central (hub) cuyo peso es el mayor de todos, que está conectado al resto de nodos y a sí mismo.

El siguiente apartado trataba de usar un algoritmo de detección de comunidades para observar cómo se agrupan los nodos de la red según su distancia. El histograma nos muestra lo siguiente: La comunidad más grande es la de tamaño 26, y viendo la imagen generada más adelante con el modelo de Kamada-Kawai, se trata de la comunidad de los nodos rojos. En total hay 7 clusters.



Imágenes 4 y 5: Histograma de la red y división por communities