

Hoja de ejercicios 1

Alejandro Perez Barreiro

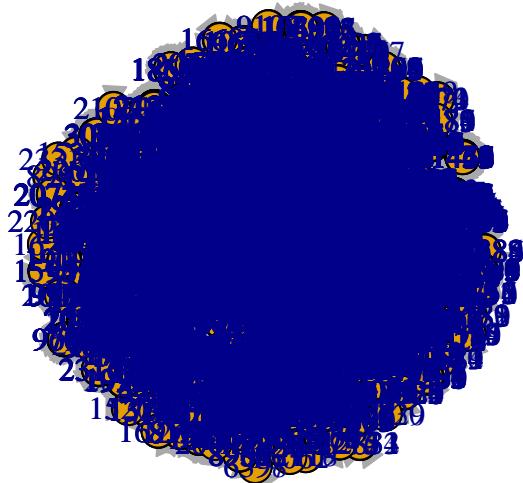
13 de febrero de 2017

```
library(igraph)
```

Pregunta 1

```
setwd("C:/Users/MSI/Desktop/Master en Data Science/Analisis de grafos y redes sociales/Evaluacion/Practica 1")
data <- read.table("out.ego-twitter", sep = " ")
vertid <- c(1:max(data[,1:2]))
g <- graph.data.frame(data,directed=TRUE, vertices = vertid)
```

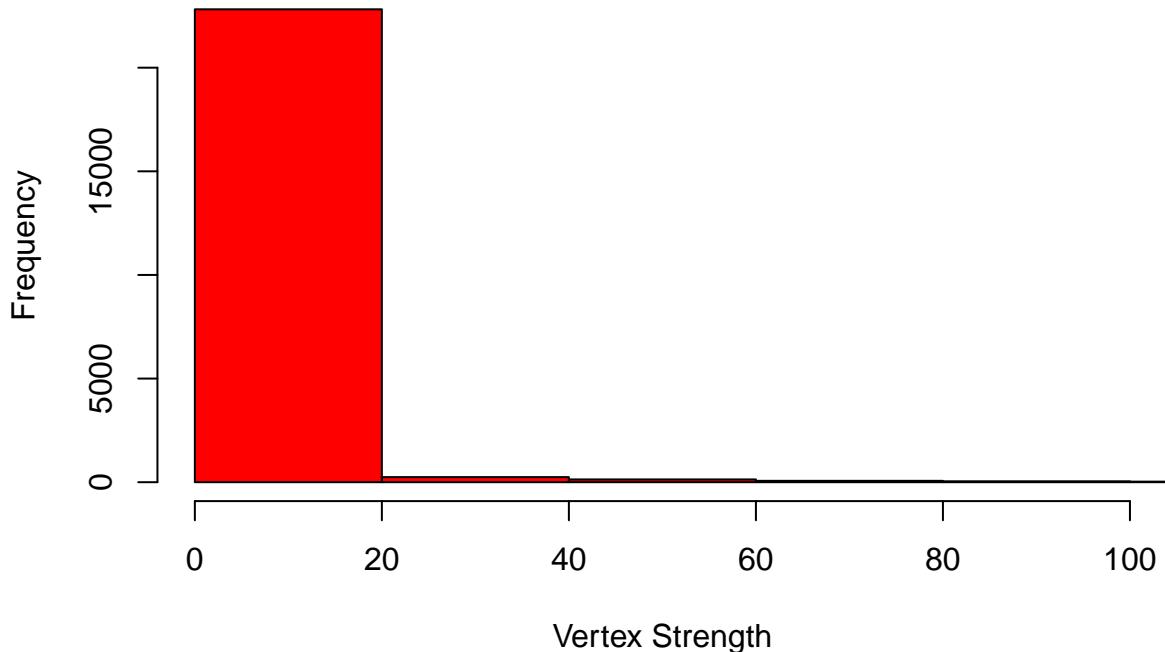
Pregunta 2



¿Es el grafo conexo? FALSE ¿Se trata de un grafo dirigido? TRUE ¿Es un grafo ponderado? FALSE Por tanto se trata de un grafo no conexo que se divide en 72 subgrafos, dirigido y no ponderado. La densidad del grafo es del 0.006061%, por lo que se trata de un grafo disperso. En este caso, la matriz de adyacencia no es eficiente, puesto que contiene una gran cantidad de huecos. Por lo tanto, es recomendable utilizar la lista de adyacencia, que es mas compacta. En el caso de la visualizacion no se puede apreciar nada a simple vista.

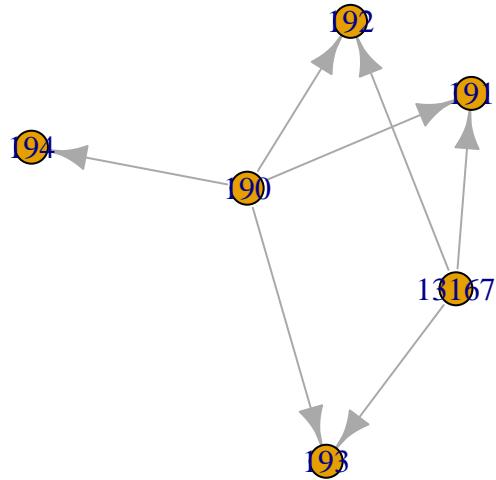
Pregunta 3

En este caso, al ser el grafo un grafo dirigido, cada vertice va a tener un grado de entrada y un grado de salida. El grado de entrada indicara que el numero de seguidores que tiene ese usuario en Twitter mientras que el grado de salida indicara el numero de personas a las que sigue dicho usuario. Siendo la suma de ambos grados el grado total.



Se observa como la mayoria de los vertices no tienen un grado superior a 20, lo que significa que en esta red la mayoria de los usuarios siguen y son seguidos por poca gente.

Pregunta 4



Se ve claramente como el subgrafo inducido por esos vértices no es fuertemente conexo ya que desde el vértice 194 no hay ningún camino hasta el 190. Este subgrafo sería débilmente conexo puesto que si lo suponemos no dirigido entonces existe un camino entre todos los pares de vértices.

Pregunta 5

El usuario con mayor número de seguidores es el 36 con 57 seguidores. Y el usuario que más gente sigue es el 11824 que sigue a 238 usuarios.

Pregunta 6

El usuario 1305 tiene 4 seguidores, que son: 1301, 7331, 13200, 16843

Pregunta 7

El usuario 1373 sigue a 57 usuarios, que son: 501, 537, 920, 1872, 1880, 2048, 2050, 2051, 2055, 2089, 2092, 3687, 3816, 5289, 6279, 6280, 6281, 6291, 6292, 6294, 6297, 6303, 6304, 6305, 6307, 6312, 6319, 6320, 6321, 6322, 6325, 8968, 10635, 16309, 16310, 16311, 16312, 16313, 16314, 16315, 16316, 16317, 16318, 16319, 16320, 16321, 16322, 16323, 16324, 16325, 16326, 16327, 16328, 16329, 16330, 16331, 16332

Pregunta 8

La proporcion de seguidores/seguidos es 0.1111111 que es una proporcion muy baja, por tanto no es un usuario influyente.

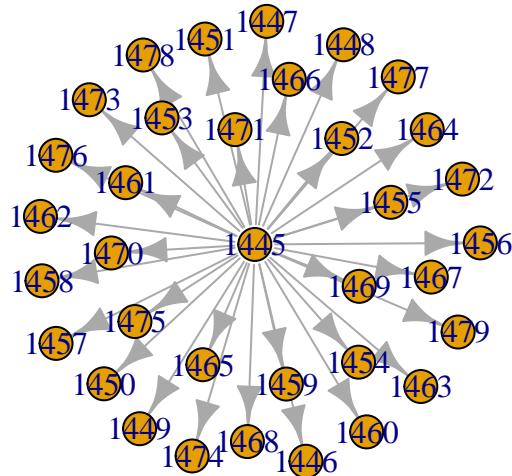
Pregunta 9

```
im <- infomap.community(g)
sizes(im)
```

```
## Community sizes
##   1    2    3    4    5    6    7    8    9    10   11   12
## 22283 74  52  40  35  34  33  31  31  30  30  28
## 13   14  15  16  17  18  19  20  21  22  23  24
## 28   27  26  24  22  21  20  17  18  17  15  15
## 25   26  27  28  29  30  31  32  33  34  35  36
## 15   15  14  14  14  14  14  13  12  12  12  10
## 37   38  39  40  41  42  43  44  45  46  47  48
## 11   11  10  10  10  10  9   9   9   8   8   7
## 49   50  51  52  53  54  55  56  57  58  59  60
## 7    7   6   6   6   6   6   3   5   5   5   5
## 61   62  63  64  65  66  67  68  69  70  71  72
## 4    4   4   3   3   3   3   3   2   2   2   2
## 73   74  75  76  77  78  79  80  81  82  83  84
## 2    2   2   2   2   2   2   2   2   2   2   2
## 85   86  87  88  89  90  91  92  93  94  95
## 2    2   2   2   2   2   2   2   2   2   2
```

El grafo esta formado por 95 comunidades donde hay una principal con 22283 nodos.

Pregunta 10



Representando el subgrafo de esa comunidad se observa claramente que el usuario menos influyente es el 1445 ya que sigue a 34 de 35 usuarios y a el le siguen 0

Pregunta 11

```
indices <- which(im$membership[vertid]==1)
subgraph3 <- induced.subgraph(g, indices)
```

A)

Para responder a esta pregunta hay que mirar la centralidad de los vertices con la medida closeness centrality, el mas central sera el mas cercano al resto.

```
sort(closeness(subgraph3), decreasing = TRUE) [1]
```

```
##          7045
## 2.35702e-09
```

Asi se obtiene que el vertice mas cercano al resto es el 7045 con una closeness centrality de 2.35702

B)

El controlador de mayor flujo de informacion sera aquel con mayor betweenness centrality

```
sort(betweenness(subgraph3), decreasing = TRUE) [1]
```

```
##      1368  
## 61233.13
```

Por tanto el vertice que controla un mayor flujo de informacion es el 1368