Network Science Homework 4

Arghya Kannadaguli (ak5357)

2025-02-28

Assignment DDescription: Select a real-world network, compute key centrality measures for nodes, select the most important node based on each centrality, and discuss what structural features make them important.

```
library(igraph)
library(tidyverse)
library(ggplot2)
library(network)
library(sna)
```

Zebra Network Dataset

“This undirected network contains interactions between 28 Grévy’s zebras (Equus grevyi) in Kenya. A node represents a zebra and an edge between two zebras shows that there was an interaction between them during the study.”

Data Source: The Konnect Project (http://konect.cc/networks/moreno_zebra/)

Data Import

Find a real-world network.

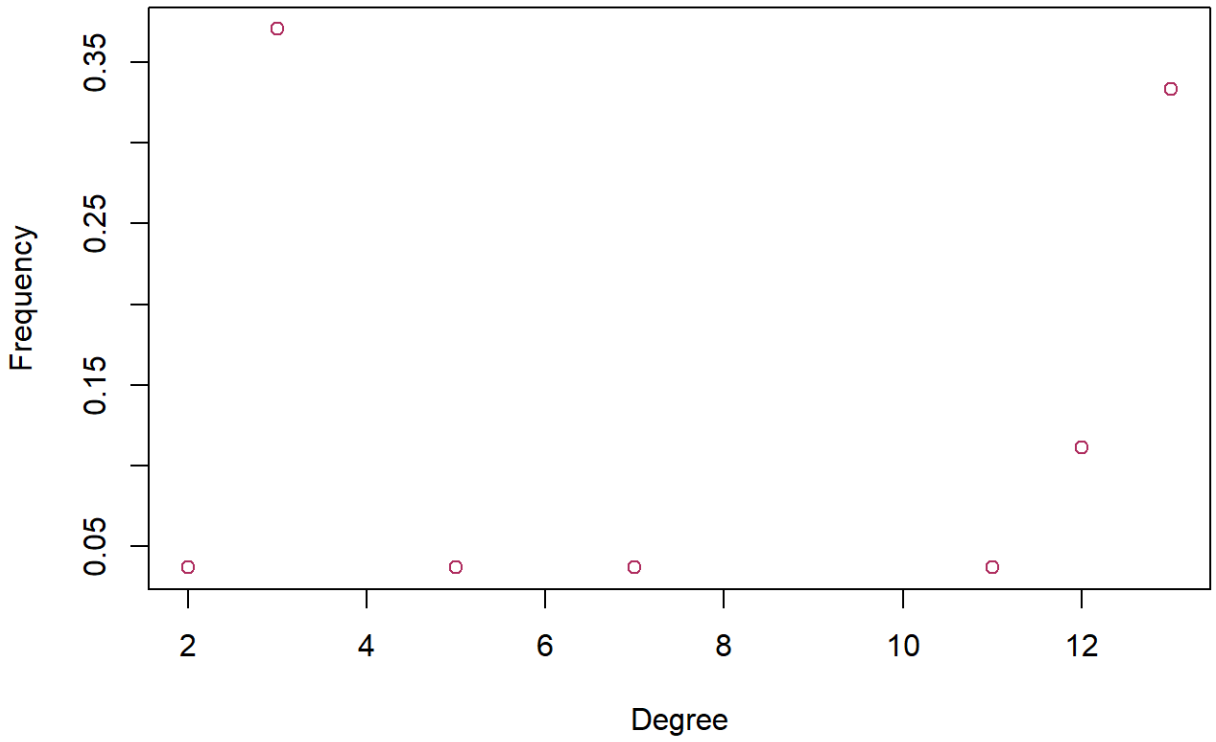
Here we are pulling in the physicians network described above.

```
zebra_edgelist =
  readr::read_delim("data/moreno_zebra/out.moreno_zebra_zebra", skip = 1) |>
  janitor::remove_empty(which = "cols") |>
  janitor::clean_names()

zebra_graph =
  as.matrix(zebra_edgelist) |>
  graph_from_edgelist()

rm(zebra_edgelist)
```

Compute Centrality Measures

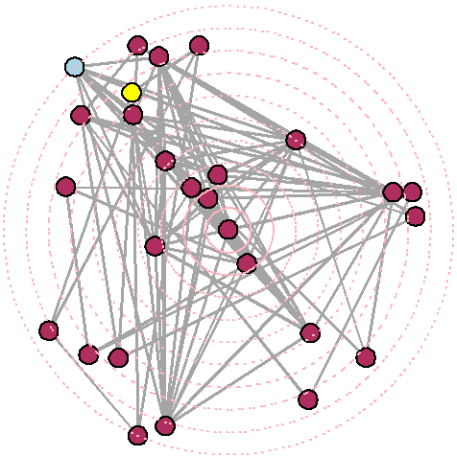
Plot of Degree Distribution

```
d_zebra = igraph::degree(zebra_graph)
dd_zebra = degree_distribution(zebra_graph)

d = 1:max(d_zebra)-1
ind = (dd_zebra!=0)

plot(d[ind], dd_zebra[ind],
     col = "maroon",
     xlab = "Degree",
     ylab = "Frequency",
     main = "Degree Distribution")
```



Plot of Degree Centrality

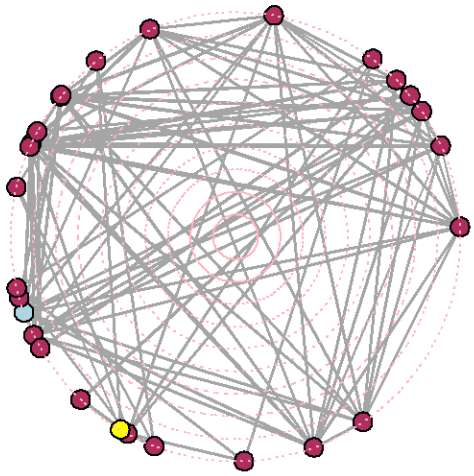
```
A = as_adjacency_matrix(zebra_graph, sparse = FALSE)
g = network::as.network.matrix(A)

sna::gplot.target(
  g,
  degree(g, gmode = "graph"),
  main = "Degree",
  circ.lab = FALSE,
  circ.col="lightpink",
  usearrows=FALSE,
  vertex.col=c("lightblue",rep("maroon",25),"yellow"),
  edge.col="darkgray"
  )
```



Plot of Closeness

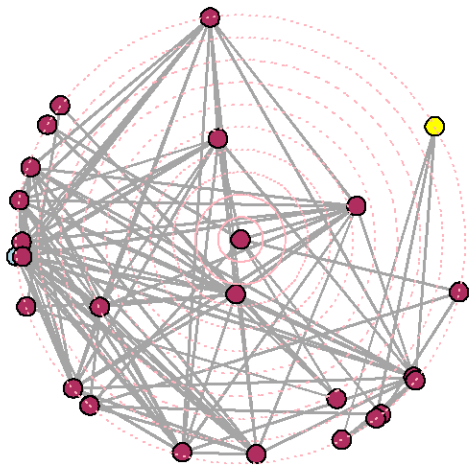
```
sna::gplot.target(
  g,
  closeness(g, gmode = "graph"),
  main = "Closeness",
  circ.lab = FALSE,
  circ.col="lightpink",
  usearrows=FALSE,
  vertex.col=c("lightblue",rep("maroon",25),"yellow"),
  edge.col="darkgray"
  )
```

Closeness



Plot of Betweenness

```
sna::gplot.target(
  g,
  betweenness(g, gmode = "graph"),
  main = "Betweenness",
  circ.lab = FALSE,
  circ.col="lightpink",
  usearrows=FALSE,
  vertex.col=c("lightblue",rep("maroon",25),"yellow"),
  edge.col="darkgray"
  )
```



Plot of Eigenvector

The following plotting attempt yields an error, because of NAs in the graph.

Error message:

```
Error in elen[elen < minlen] <- (outer(radii, radii, "+")/sqrt(2))[elen <  : NAs are not allowed in subscripted assignments
```

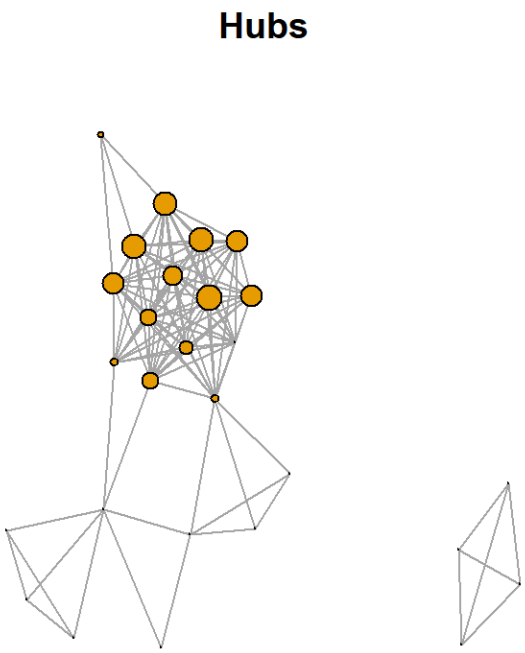
```
# sna::gplot.target(
#   g,
#   evcent(g, gmode = "graph"),
#   main = "Eigenvector",
#   circ.lab = FALSE,
#   circ.col="lightpink",
#   usearrows=FALSE,
#   vertex.col=c("lightblue",rep("maroon",25),"yellow"),
#   edge.col="darkgray"
#   )
```

Select Important Nodes

Hubs

Next, I will compute the hubs and authorities within the Zebra network. Based on the visualizatoin below, it looks like there are two components within this network. Within the larger component, there is a group of zebras who have a lot of interactions as well as some peripheral zebras who do not have as many interactions between each other or with the higher-interactions group. In the second, smaller component, the zebras all interact with one another, but perhaps less frequently than do those in the main component's high-interactions group. You can see that the structural features that lead to certain nodes having higher hub scores appear to be the degree of the nodes.

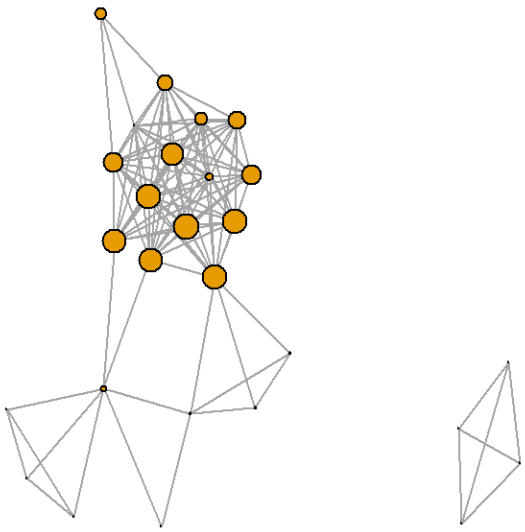
```
l = layout_with_kk(zebra_graph)
plot(zebra_graph,
     layout=l,
     main="Hubs",
     vertex.label="",
     vertex.size=10*sqrt(hub_score(zebra_graph)$vector),
     edge.arrow.size = 0)
```



Authorities

It appears the zebraswith the highest authority scores are also within the high-interactions group in the main component. However, these scores are slightly different from the hub scores. For example, the zebra that connects the high-interaction group to the smaller group to the bottom left has a higher authority score than hub score. This is likely because it is connected to high-degree zebras in the high-interactions group. The smaller component's zebras have similarly low authority scores as their hub scores.

```
plot(zebra_graph,
     layout=l,
     main="Hubs",
     vertex.label="",
     vertex.size=10*sqrt(authority_score(zebra_graph)$vector),
     edge.arrow.size = 0)
```



K-Core decomposition

```
cores = coreness(zebra_graph)
A = as_adjacency_matrix(zebra_graph, sparse=FALSE)
g = network::as.network.matrix(A)
sna::gplot.target(g,
                  cores,
                  main="K-Core Decomposition",
                  circ.lab = FALSE,
                  circ.col = "lightpink",
                  usearrows = FALSE,
                  vertex.col=cores,
                  edge.col="darkgrey")
```

K-Core Decomposition