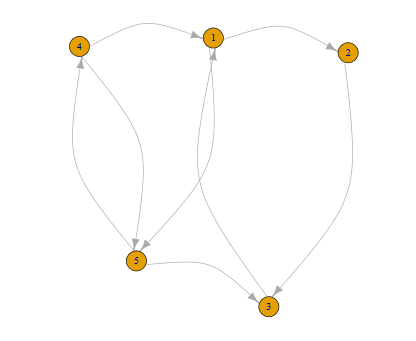
1. Consider the following network



a. What is the adjacency matrix of this graph?

b. What is the edge-list of this graph?

c. If we want to use igraph and tkplot to plot this graph, what is the R code?

Answer:

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | 1 | 2 | 3 | 4 | 5 |
| 1 | 0 | 1 | 0 | 0 | 1 |
| 2 | 0 | 0 | 1 | 0 | 0 |
| 3 | 1 | 0 | 0 | 0 | 0 |
| 4 | 1 | 0 | 0 | 0 | 1 |
| 5 | 0 | 0 | 1 | 1 | 0 |

b.

|  |  |
| --- | --- |
| 1 | 2 |
| 1 | 5 |
| 2 | 3 |
| 3 | 1 |
| 4 | 1 |
| 4 | 5 |
| 5 | 3 |
| 5 | 4 |

C.

install.packages("igraph");

library(igraph);

el\_d <- matrix( c(1,2,1,5,2,3,3,1,4,1,4,5,5,3,5,4 ),8,2,byrow=TRUE)

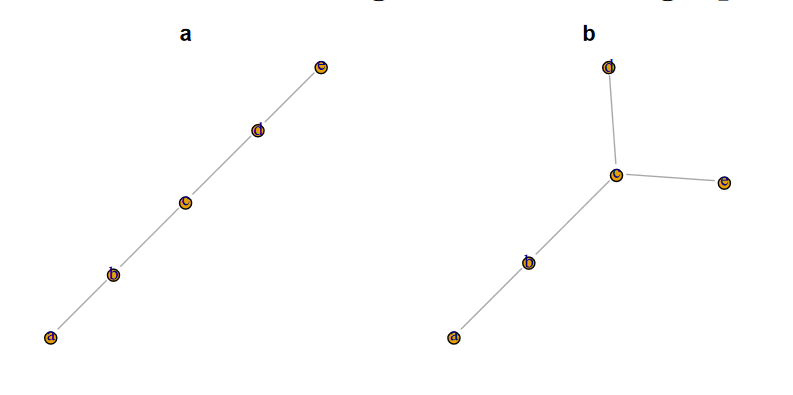
g\_el\_d <- graph.edgelist(el\_d,directed=TRUE)

tkplot(g\_el\_d)

2. Consider the following 2 undirected graphs:

a. which nodes you believe to be most central and why?

b. If we want to use igraph and tkplot to plot the graphs below, what is the R code?



Answer:

a. the plot is b,c,d ; the second plot is c, because the highest degree

b.

//first plot

install.packages("igraph");

library(igraph);

el\_d <- matrix( c('a','b','b','c','c','d','d','e' ),4,2,byrow=TRUE)

g\_el\_d <- graph.edgelist(el\_d,directed=FALSE)

tkplot(g\_el\_d)

//first plot

install.packages("igraph");

library(igraph);

el\_d <- matrix( c('a','b','b','c','c','d','c','e' ),4,2,byrow=TRUE)

g\_el\_d <- graph.edgelist(el\_d,directed=FALSE)

tkplot(g\_el\_d)

3. Visit <https://twitter.com/callutheransom>,( if you don’t have a Twitter account, create one). Analyze the followers and the followings of School of Management. Based on the description of each user, identify five groups of followers and five groups of followings. What are the main characteristics of each group?

Answer:

library(twitteR)

library(httpuv)

consumer\_key <- '4uIzbufjdqjxqc2pA2KTssKXZ'

consumer\_secret <- 'O40Ec4NYi3vb0kd5xQw2CxFeAfPsjnVm2jf4lH5uc2xQlPV9PI'

access\_token <- '1014453924667932672-TnssrO7NSr07LcUqC0ZRQY80XspeM4'

access\_secret <- 'cs6zX02SOaWMsD6fpEMEvUhpPpAKF8McYCgJ3w1enwa7u'

setup\_twitter\_oauth(consumer\_key, consumer\_secret, access\_token, access\_secret)

twitterUser <- getUser("CalLutheranSOM")

followers <- twitterUser$$getFollowers() # my followers

followers\_df <- twListToDF(followers)

mat = matrix(followers\_df$description)

library(tm)

corpus <- (VectorSource(mat))

corpus <- Corpus(corpus)

summary(corpus)

twtrTermDocMatrix <- TermDocumentMatrix(corpus)

which(apply(twtrTermDocMatrix,1,sum)>=30)

twtrTermDocMatrix2 <- removeSparseTerms(twtrTermDocMatrix, sparse = 0.97)

tweet\_matrix <- as.matrix(twtrTermDocMatrix2)

distMatrix <- dist(scale(tweet\_matrix))

fit <- hclust(distMatrix,method="single")

cutree(fit, k = 1:5)

plot(fit)

Follower

1. business
2. Ventura
3. Education
4. Marketing
5. community

following---very difficult to cluster by description, there is no significant classifier

1. student
2. politics or social
3. business
4. education
5. Community