Social Network Analysis

Suwon Jung

4/27/2020

Goal:

Generating and analyzing three social networks (also known as graphs or sociograms) based on three different measures

Data Source:

Representing Classroom Social Structure. Melbourne: Victoria Institute of Secondary Education, M. Vickers and S. Chan, (1981)

Available from the Index of Complex Networks (ICON (https://icon.colorado.edu/#!/))

The data were collected by Vickers & Chan from 29 seventh grade students in a school in Victoria, Australia. Students were asked to nominate their classmates on a number of relations including the following three “layers”:

1. Who do you get on with in the class?
2. Who are your best friends in the class?
3. Who would you prefer to work with?

Data Wrangling

Manipulate each of the data sets so that it is suitable for building a social network using iGraph.

```
#Load packages
library(dplyr)
```

```
##
## Attaching package: 'dplyr'
```

```
## The following objects are masked from 'package:stats':
##
##     filter, lag
```

```
## The following objects are masked from 'package:base':
##
##     intersect, setdiff, setequal, union
```

```
library(tidyr)

#import data
bestfriend <- read.csv("best.friends.csv")
getonwith <- read.csv("get.on.with.csv")
workwith <- read.csv("work.with.csv")

# best friend
bestfriend <- bestfriend[,-1]
  #edge
edge_bf <- count(bestfriend, from, to)
names(edge_bf) <- c("from", "to", "count")
  #vertex
vertex_bf <- bestfriend[,c(1,3)]
vertex_bf <- unique(vertex_bf)
names(vertex_bf) <- c("id", "gender")

# get on with
getonwith <- getonwith[,-1]
  #edge
edge_gow <- count(getonwith, from, to)
names(edge_gow) <- c("from", "to", "count")
  #vertex
vertex_gow <- getonwith[,c(1,3)]
vertex_gow <- unique(vertex_gow)
names(vertex_gow) <- c("id", "gender")

# work with
workwith <- workwith[,-1]
  #edge
edge_ww <- count(workwith, from, to)
names(edge_ww) <- c("from", "to", "count")
  #vertex
vertex_ww <- workwith[,c(1,3)]
vertex_ww  <- unique(vertex_ww )
names(vertex_ww ) <- c("id", "gender")
```

Visualize the Networks

Create a graph for each of the data sets, are the graphs directed or undirected? Visualize each of the graphs you have created and color the nodes according to gender.

```
#load package
#install.packages('igraph')
library(igraph)
```

```
## Warning: package 'igraph' was built under R version 3.6.3
```

```
##
## Attaching package: 'igraph'
```

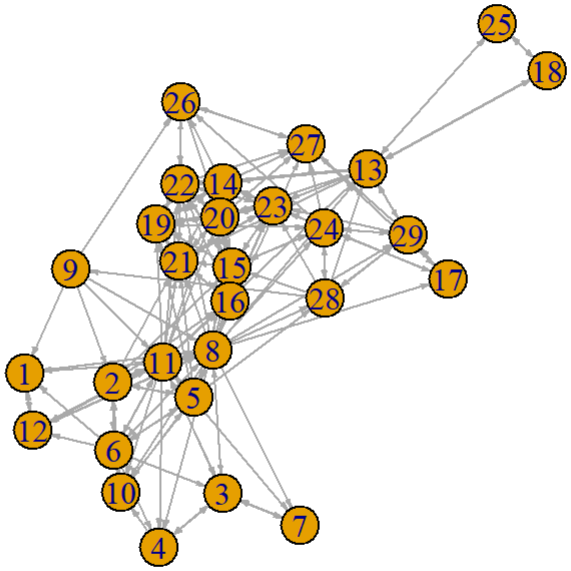
```
## The following object is masked from 'package:tidyr':
##
##     crossing
```

```
## The following objects are masked from 'package:dplyr':
##
##     as_data_frame, groups, union
```

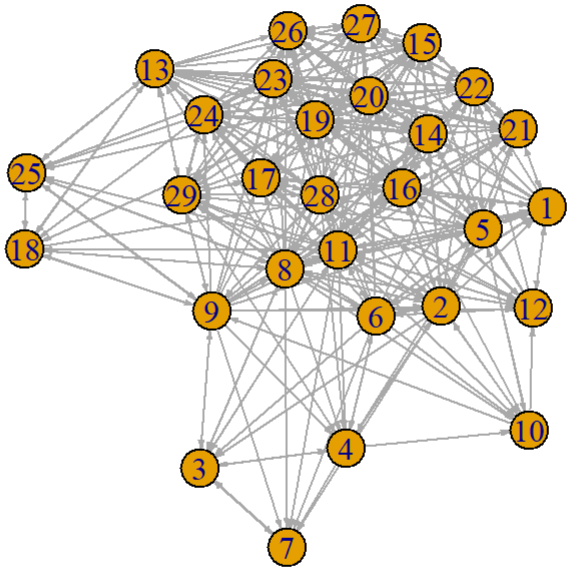
```
## The following objects are masked from 'package:stats':
##
##     decompose, spectrum
```

```
## The following object is masked from 'package:base':
##
##     union
```

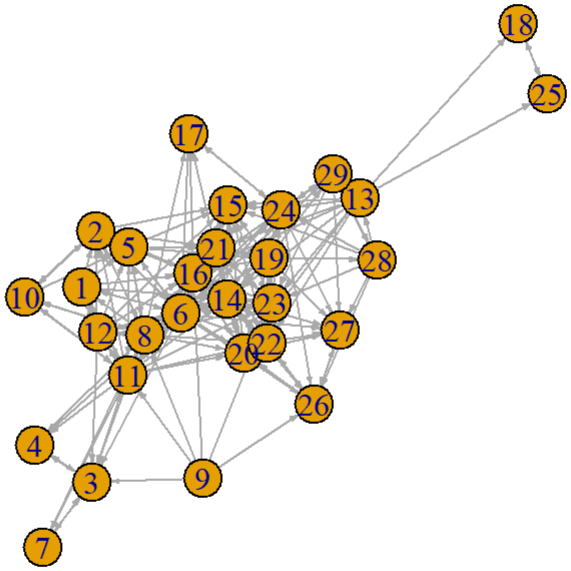
```
#graph of best
g_bf <- graph.data.frame(edge_bf, directed=TRUE, vertices=vertex_bf)
#graph of get on with
g_gow <- graph.data.frame(edge_gow, directed=TRUE, vertices=vertex_gow)
#graph of work with
g_ww <- graph.data.frame(edge_ww, directed=TRUE, vertices=vertex_ww)

#plot the graphs
plot_bf<-plot(g_bf, layout = layout.fruchterman.reingold,
             vertex.color = vertex_bf$gender.from,
             vertex.size=15, edge.arrow.size=0.2, edge.width=edge_bf$count)
```




```
plot_gow<-plot(g_gow, layout = layout.fruchterman.reingold,
             vertex.color = vertex_gow$gender.from,
             vertex.size=15, edge.arrow.size=0.2, edge.width=edge_gow$count)
```




```
plot_ww<-plot(g_ww, layout = layout.fruchterman.reingold,
              vertex.color = vertex_ww$gender.from,
              vertex.size=15, edge.arrow.size=0.2, edge.width=edge_ww$count)
```




```
pdf("SocialNetworkAnalysis_JUNG.pdf", paper = 'letter', onefile=TRUE)
```

Centrality Measures

Who in the class has the highest degree centrality for each measure?

```
## Degree Centrality

bf_degree <- degree(g_bf)
gow_degree <- degree(g_gow)
ww_degree <- degree(g_ww)

which(bf_degree==max(bf_degree))
```

```
## 8
## 8
```

```
  #ID 8 has the highest degree centrality in best friend network
which(gow_degree==max(gow_degree))
```

```
## 11
## 11
```

```
  #ID 11 has the highest degree centrality in get on wth netwrok
which(ww_degree==max(ww_degree))
```

```
## 6
## 6
```

```
  #ID 6 has the higherst degree centrality in work with network
```

Who in the class has the highest closeness centrality?

```
## Betweenness Centrality

bf_close <- closeness(g_bf)
gow_close <- closeness(g_gow)
ww_close <- closeness(g_ww)
```

```
## Warning in closeness(g_ww): At centrality.c:2784 :closeness centrality is
## not well-defined for disconnected graphs
```

```
which(bf_close==max(bf_close))
```

```
## 8
## 8
```

```
  #ID 8 has the highest closeness cen trality in best friend network
which(gow_close==max(gow_close))
```

```
##  8 11
##  8 11
```

```
  #ID 8 and 11 have the highest closeness centrality in get on with network
which(ww_close==max(ww_close))
```

```
## 9
## 9
```

```
  #ID 9 has the highest closeness cen trality in work with network
```

Simple structures

Count the number of dyads and the number and type of triads using the following commands.

```
dyad_census(g_bf)
```

```
## $mut
## [1] 55
##
## $asym
## [1] 71
##
## $null
## [1] 280
```

```
dyad_census(g_gow)
```

```
## $mut
## [1] 121
##
## $asym
## [1] 119
##
## $null
## [1] 166
```

```
dyad_census(g_ww)
```

```
## $mut
## [1] 46
##
## $asym
## [1] 106
##
## $null
## [1] 254
```

Documentation (http://igraph.org/r/doc/dyad_census.html)

```
triad_census(g_bf)
```

```
##  [1] 1297  791  746  118   26   75  123  214   34    2   39   59   38    3
## [15]   43   46
```

```
triad_census(g_gow)
```

```
##  [1] 329 510 632 186  55  88 235 432  97   8 215 193 120  75 287 192
```

```
triad_census(g_ww)
```

```
##  [1] 1022  999  528  187  116  132  108  177   88    0   18  102   81   19
## [15]   40   37
```

Documentation (http://igraph.org/r/doc/triad_census.html)

Cliques

Answer the following questions using the clique functions (http://igraph.org/r/doc/cliques.html)

What is the size of the largest clique(s) in each of the three networks?

```
clique_num(g_bf)
```

```
## Warning in clique_num(g_bf): At cliques.c:1087 :directionality of edges is
## ignored for directed graphs
```

```
## [1] 7
```

```
  #the size of the largest cluques in best friend netwrok is 7
clique_num(g_gow)
```

```
## Warning in clique_num(g_gow): At cliques.c:1087 :directionality of edges is
## ignored for directed graphs
```

```
## [1] 12
```

```
  #the size of the largest cluques in best friend netwrok is 12
clique_num(g_ww)
```

```
## Warning in clique_num(g_ww): At cliques.c:1087 :directionality of edges is
## ignored for directed graphs
```

```
## [1] 9
```

```
  #the size of the largest cluques in best friend netwrok is 9
```

Which nodes/vertices are in the largest cliques for the three networks? Is there much overlap?

```
largest.cliques(g_bf)
```

```
## Warning in largest.cliques(g_bf): At cliques.c:1087 :directionality of
## edges is ignored for directed graphs
```

```
## [[1]]
## + 7/29 vertices, named, from 8f3f792:
## [1] 23 14 15 19 20 21 22
##
## [[2]]
## + 7/29 vertices, named, from 8f3f792:
## [1] 23 14 15 16 20 21 22
##
## [[3]]
## + 7/29 vertices, named, from 8f3f792:
## [1] 8  15 16 21 14 20 22
##
## [[4]]
## + 7/29 vertices, named, from 8f3f792:
## [1] 8  11 22 14 20 21 16
##
## [[5]]
## + 7/29 vertices, named, from 8f3f792:
## [1] 8  11 22 14 20 21 13
```

```
largest.cliques(g_gow)
```

```
## Warning in largest.cliques(g_gow): At cliques.c:1087 :directionality of
## edges is ignored for directed graphs
```

```
## [[1]]
## + 12/29 vertices, named, from 8f3f793:
##  [1] 8  11 27 15 13 24 23 20 14 19 22 26
##
## [[2]]
## + 12/29 vertices, named, from 8f3f793:
##  [1] 8  11 27 15 13 24 23 20 14 19 22 21
##
## [[3]]
## + 12/29 vertices, named, from 8f3f793:
##  [1] 8  11 22 14 16 19 15 20 26 13 23 24
##
## [[4]]
## + 12/29 vertices, named, from 8f3f793:
##  [1] 8  11 22 14 16 19 15 20 21 23 13 24
```

```
largest.cliques(g_ww)
```

```
## Warning in largest.cliques(g_ww): At cliques.c:1087 :directionality of
## edges is ignored for directed graphs
```

```
## [[1]]
## + 9/29 vertices, named, from 8f3f793:
## [1] 6  15 16 14 19 23 20 21 22
##
## [[2]]
## + 9/29 vertices, named, from 8f3f793:
## [1] 6  8  16 14 19 20 21 22 23
##
## [[3]]
## + 9/29 vertices, named, from 8f3f793:
## [1] 6  8  16 14 11 20 21 22 23
```

How many **maximal cliques** are there in each of the networks?

```
count_max_cliques(g_bf)
```

```
## Warning in count_max_cliques(g_bf): At maximal_cliques_template.h:203 :Edge
## directions are ignored for maximal clique calculation
```

```
## [1] 35
```

```
  #largest clique is 35 in best friend network
count_max_cliques(g_gow)
```

```
## Warning in count_max_cliques(g_gow): At maximal_cliques_template.h:
## 203 :Edge directions are ignored for maximal clique calculation
```

```
## [1] 64
```

```
  #largest clique is 64 in get on with network
count_max_cliques(g_ww)
```

```
## Warning in count_max_cliques(g_ww): At maximal_cliques_template.h:203 :Edge
## directions are ignored for maximal clique calculation
```

```
## [1] 36
```

```
  #largest clique is 36 in work with network
```

Components & Cutpoints

Find the cutpoints (articulation points) for each of the three networks you generated. What does this tell you about the graphs? Does what you find match a visual exploration of the networks?

Putting it all together

Write a narrative description of the social ties among members of this 7th grade class using the network, centrality and structural metrics you have developed. Can you provide any recommendations to the teacher of this class based on your analysis? What other information would you want to know? Can you remeber being in seventh grade, does this refelct your experience?