

Homework 1

Introduction to Network Analysis

Aleksey Shadrinov (Group 2, asshadrinov@edu.hse.ru)

2022-11-06

Contents

prepare packages and load data	2
first examination of data	2
basic network statistics	4
density	5
triads	5
case of directed network	5
transitivity	6
reciprocity	6

prepare packages and load data

install packages (uncomment if needed):

```
# install.packages("network")
# install.packages("sna")
```

load data `trade.Rdata`:

```
suppressPackageStartupMessages(library(network))
suppressPackageStartupMessages(library(sna))
load("trade.Rdata")
```

check what have been loaded:

```
ls()

## [1] "crude"      "diplomacy"  "food"      "manufacture" "minerals"
## [6] "trade.all"  "trade.att"
```

first examination of data

for further analysis I've selected network **"manufacture"**. let's convert it to matrix and check its dimensions.

```
manufacture = as.matrix(manufacture)
dim(manufacture)
```

```
## [1] 24 24
```

in addition, we have loaded attributes of the vertices in `trade.att` object. its dimensions are:

```
dim(trade.att)
```

```
## [1] 24 4
```

we may output several cells of each matrix to check:

```
manufacture[1:6,1:6]
```

```
##           ALGERIA ARGENTINA BRAZIL CHINA CZECHOSLOVAKIA ECUADOR
## ALGERIA           0         0     0    1             1       0
## ARGENTINA         1         0     1    1             0       1
## BRAZIL            1         1     0    1             1       1
## CHINA             1         1     1    0             1       0
## CZECHOSLOVAKIA    1         1     1    1             0       1
## ECUADOR           0         0     1    0             0       0
```

```
head(trade.att)
```

```
##           POP_GROWTH GNP SCHOOLS ENERGY
## ALGERIA           3.3 3.0     33     814
## ARGENTINA         1.6 0.3     56    2161
## BRAZIL            2.1 5.3     32    1101
## CHINA             1.5 NA      43     618
## CZECHOSLOVAKIA    0.7 NA      44    6847
## ECUADOR           3.4 4.7     40     692
```

```
all(colnames(manufacture) == colnames(trade.all))
```

```
## [1] TRUE
```

as we see, both matrices have the same column names, and `trade.att` contains `POP_GROWTH`, `GNP`, `SCHOOLS` and `ENERGY` columns. so, we can safely apply attributes to the network.

basic network statistics

first step is to create a network:

```
manufacture.net <- as.network(  
  x = manufacture,  
  directed = FALSE,  
  matrix.type = "adjacency"  
)
```

here `directed` parameter equals `FALSE`, however it is not correct, since the network is directed. to prove that, we should compare adjacency matrix with its transposition. in our case, these matrices differ, so network is directed. but for simplicity we ignore that.

```
all(t(manufacture) == manufacture) # is network undirected?
```

```
## [1] FALSE
```

let's look at the network parameters:

```
manufacture.net
```

```
## Network attributes:  
## vertices = 24  
## directed = FALSE  
## hyper = FALSE  
## loops = FALSE  
## multiple = FALSE  
## bipartite = FALSE  
## total edges= 195  
## missing edges= 0  
## non-missing edges= 195  
##  
## Vertex attribute names:  
## vertex.names  
##  
## No edge attributes
```

as we observe, the network has 24 vertices and 195 edges, no loops and duplicating edges.

now let's count other statistics. `## dyads`

```
# number of dyads:  
network.dyadcount(manufacture.net)
```

```
## [1] 276
```

```
# types of dyads  
dyad.census(manufacture.net)
```

```
## Mut Asym Null  
## [1,] 195 0 81
```

number of dyads is just the number of all possible pairs of the vertices (potential connection) and could be in simple case calculated as

$$\frac{n \cdot (n - 1)}{2}$$

however, `census` is more informative: there are 195 mutual dyads and 81 null ones, so 81 pairs (about 30%) of vertices are not connected directly.

density

```
# number of ties we already know from above, but this is another way to get it
network.edgcount(manufacture.net)
```

```
## [1] 195
```

```
# density of the network
network.density(manufacture.net)
```

```
## [1] 0.7065217
```

network density describes the portion of the potential connections in a network that are actual connections. our network is highly dense, which means that there are many connections in the network and international trade of manufactured goods is very active.

triads

```
# types of triads
triad.census(manufacture.net)
```

```
##      003 012 102 021D 021U 021C 111D 111U 030T 030C 201 120D 120U 120C 210 300
## [1,] 174   0 275    0    0    0    0    0    0    0 710    0    0    0    0 865
```

in our undirected network we have only few types of triads, specifically 003, 102, 201, 300. as last two types are in total 78% of all triads, we may conclude again that connectivity is high.

case of directed network

now we should repeat the procedure with directed network, which our network in fact is.

```
manufacture.directed <- as.network(manufacture, directed = TRUE)
manufacture.directed
```

```
## Network attributes:
##   vertices = 24
##   directed = TRUE
##   hyper = FALSE
##   loops = FALSE
##   multiple = FALSE
##   bipartite = FALSE
##   total edges= 310
##   missing edges= 0
##   non-missing edges= 310
##
## Vertex attribute names:
##   vertex.names
##
## No edge attributes
```

```
# number of dyads:
network.dyadcount(manufacture.directed)
```

```
## [1] 552
```

```
# types of dyads
dyad.census(manufacture.directed)
```

```
##      Mut Asym Null
```

```
## [1,] 115 80 81
# density of the network
network.density(manufacture.directed)

## [1] 0.5615942
# types of triads
triad.census(manufacture.directed)

##      003 012 102 021D 021U 021C 111D 111U 030T 030C 201 120D 120U 120C 210 300
## [1,] 174 164 111 171 7 19 17 381 18 0 115 12 230 20 226 359
```

in case of directed network amount of possible connections is $n(n-1)$. there are 310 total edges, so density is around 0.56, which is also a significant value.

number of mutual connections is higher than assymetrical, so more countries exchange goods rather than trade in *one direction*.

considering triads, there are only 18 transitive ones from 2024, and most frequent type is still 300 (all three countries are mutually connected). also there are many 021D triads, which represent the distribution from one country to two others.

transitivity

the transitivity is an important property of a network, which describes whether $a \rightarrow b \rightarrow c \Rightarrow a \rightarrow c$. usually in social networks this parameter is high, and in other networks not so. for random network transitivity coefficient is $m / \frac{n(n-1)}{2}$

```
gtrans(manufacture.net, measure="weak") # transitivity
```

```
## [1] 0.785174
```

our network's transitivity coefficient is 0.78, which is a high value, as it should be in social networks. for comparison, in a random network with $m = 195$ edges $T = \frac{195}{12 \cdot 23} = 0.7$, so expectations are quite correct.

reciprocity

the reciprocity of a directed network reflects the proportion of edges that are symmetrical. that is, the proportion of outgoing edges that also have an incoming edge. it is commonly used to determine how inter-connected directed networks are.

```
grecip(manufacture.directed, measure = 'dyadic')
```

```
##      Mut
## 0.7101449
```

```
grecip(manufacture.directed, measure = 'dyadic.nonnull')
```

```
##      Mut
## 0.5897436
```

```
grecip(manufacture.directed, measure = 'edgewise')
```

```
##      Mut
## 0.7419355
```

```
grecip(manufacture.directed, measure = 'edgewise.lrr')
```

```
##      Mut
## 0.2784828
```

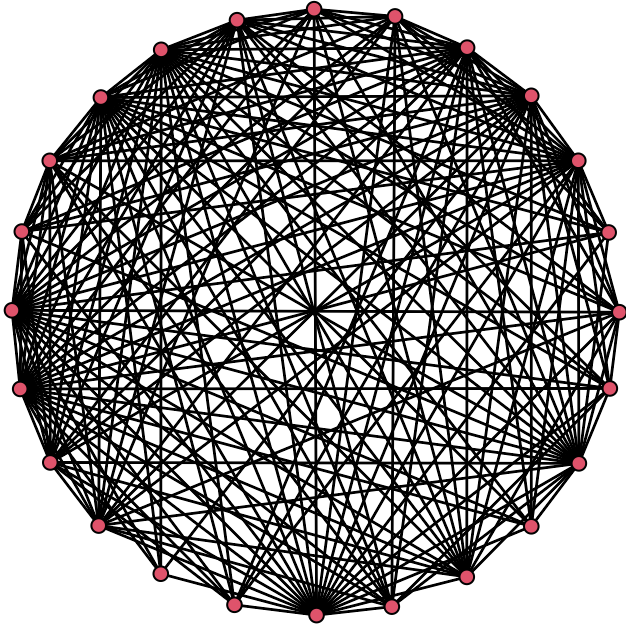
```
grecip(manufacture.directed, measure = 'correlation')
```

```
## [1] 0.411357
```

```
?gtrans
```

```
par(mar=c(1,1,1,1))
```

```
plot(manufacture.net, mode = 'circle')
```



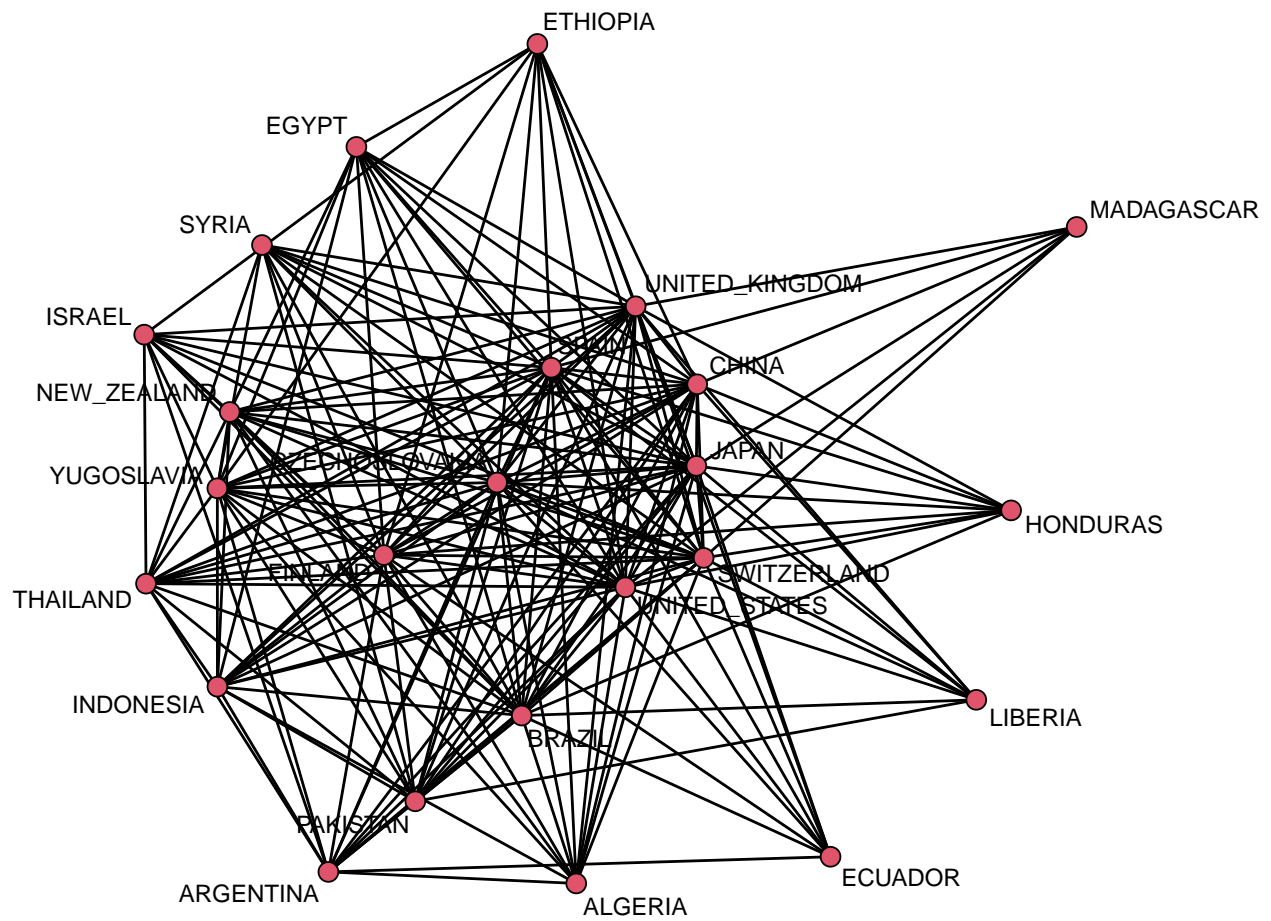
```
plot(manufacture.net, mode = 'fruchtermanreingold', main="first", displaylabels=TRUE, label.cex=0.8, label.dist=0.1)
```

```
## Warning in plot.window(...): "label.dist" is not a graphical parameter
```

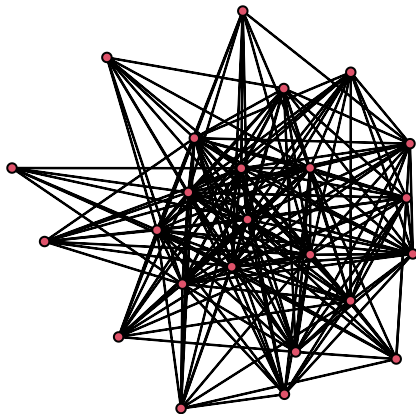
```
## Warning in plot.xy(xy, type, ...): "label.dist" is not a graphical parameter
```

```
## Warning in title(...): "label.dist" is not a graphical parameter
```

first



```
plot(manufacture.net, mode = 'kamadakawai')
```



```
# install.packages('sna')
```

```
?network.layout
```



```
par = list(c(10000))
names(par) = c("niter")
par
```

```
## $niter
## [1] 10000
```

```
head(trade.att)
```

```
##           POP_GROWTH GNP SCHOOLS ENERGY
## ALGERIA           3.3 3.0      33      814
## ARGENTINA         1.6 0.3      56     2161
## BRAZIL            2.1 5.3      32     1101
## CHINA             1.5 NA       43      618
## CZECHOSLOVAKIA    0.7 NA       44     6847
## ECUADOR           3.4 4.7      40      692
```

```
set.vertex.attribute(manufacture.net, attrname="POP_GROWTH", value=trade.att[,1])
set.vertex.attribute(manufacture.net, attrname="GNP", value=trade.att[,2])
set.vertex.attribute(manufacture.net, attrname="SCHOOLS", value=trade.att[,3])
set.vertex.attribute(manufacture.net, attrname="ENERGY", value=trade.att[,4])
```

```
manufacture.net
```

```
## Network attributes:
##   vertices = 24
##   directed = FALSE
##   hyper = FALSE
##   loops = FALSE
##   multiple = FALSE
##   bipartite = FALSE
##   total edges= 195
##     missing edges= 0
##     non-missing edges= 195
##
## Vertex attribute names:
##   ENERGY GNP POP_GROWTH SCHOOLS vertex.names
##
## No edge attributes
```

```
is network not directed?
```