Assignment 2

Network Analysis 2017

Data-analysis

In R, run the following code:

```
install.packages("psych")
```

```
library("psych")
data("bfi")
bfiData <- bfi[,1:25]
```

The data frame `bfiData` contains the questions of the `bfi` (Big Five Inventory) data contained in the *psych* package. More information on this dataset can be obtained by running:

```
?bfi
```

The questions are designed to measure five central personality traits: Agreeableness, Conscientiousness, Extraversion, Neuroticism, and Opennness. The following table gives the item descriptions:

Item label	Item description	Trait
A1	Am indifferent to the feelings of others	Agreeableness
A2	Inquire about others' well-being	Agreeableness
A3	Know how to comfort others	Agreeableness
A4	Love children	Agreeableness
A5	Make people feel at ease	Agreeableness
C1	Am exacting in my work	Conscientiousness
C2	Continue until everything is perfect	Conscientiousness
C3	Do things according to a plan	Conscientiousness
C4	Do things in a half-way manner	Conscientiousness
C5	Waste my time	Conscientiousness
E1	Don't talk a lot	Extraversion
E2	Find it difficult to approach others	Extraversion
E3	Know how to captivate people	Extraversion
E4	Make friends easily	Extraversion
E5	Take charge	Extraversion
N1	Get angry easily	Neuroticism
N2	Get irritated easily	Neuroticism
N3	Have frequent mood swings	Neuroticism
N4	Often feel blue	Neuroticism
N5	Panic easily	Neuroticism
O1	Am full of ideas	Opennness
O2	Avoid difficult reading material	Opennness
O3	Carry the conversation to a higher level	Opennness
O4	Spend time reflecting on things	Opennness
O5	Will not probe deeply into a subject	Opennness

We can compute a polychoric correlation matrix based on this data as follows:

```
library("qgraph")
corMat <- cor_auto(bfiData)
```

Next we can use *qgraph* to compute a partial correlation network:

```
qgraph(corMat, graph = "pcor", layout = "spring", cut = 0)
```

We can use the *bootnet* function `estimateNetwork` to automate this process:

```
library("bootnet")
Result_pcor <- estimateNetwork(bfiData, default = "pcor")
plot(Result_pcor, layout = "spring", cut = 0)
```

Exercise 1 (1 point) Obtain the weights matrices from qgraph and bootnet by applying the `getWmat` function to output of both. Confirm that the results are identical (tip: the operator == tests if values in R are equal).
Solution:

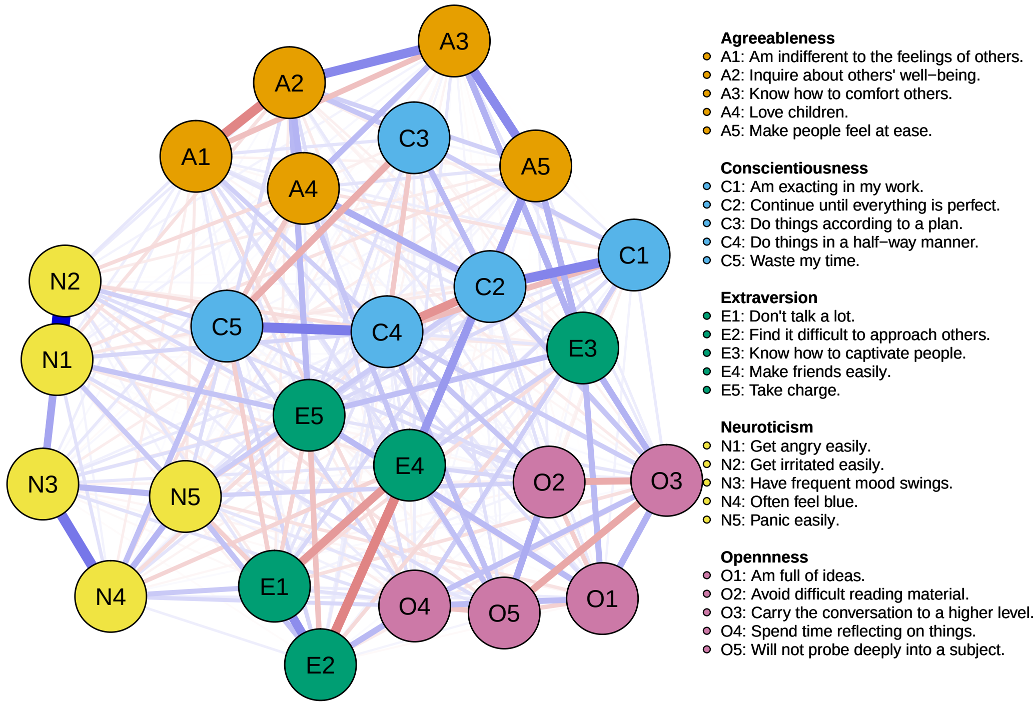
```
g1 <- qgraph(corMat, graph = "pcor", layout = "spring", cut = 0)
w1 <- getWmat(g1)
g2 <- estimateNetwork(bfiData, default = "pcor")
w2 <- getWmat(g2)

all(w1==w2)
# Returns TRUE
```

We can load for each node the item description and factor the item is aimed to measure as follows:

```
Names <- scan("http://sachaepskamp.com/files/BFIitems.txt",
              what = "character", sep = "\n")
Traits <- rep(c(
  'Agreeableness',
  'Conscientiousness',
  'Extraversion',
  'Neuroticism',
  'Opennness'
),each=5)
```

These can be used to plot a legend next to the graph. In combination, we can make the graph friendly to colorblind viewers using the `theme` option:

```
plot(Result_pcor,
     layout = "spring",
     cut = 0,
     theme = "colorblind",
     groups = Traits,
     nodeNames = Names,
     legend.cex = 0.4)
```



In `estimateNetwork`, the `fun` argument can be used to specify a custom function estimating the network from data. To aid the user, several default functions have been built in. For example, `default = "pcor"` specified a function that estimates a partial correlation networks (in the help file this function is called `bootnet_pcor`).

Exercise 2 (1 point) Use the `default` argument in `estimateNetwork` to estimate a partial correlation network using glasso and EBIC model selection.
Solution:

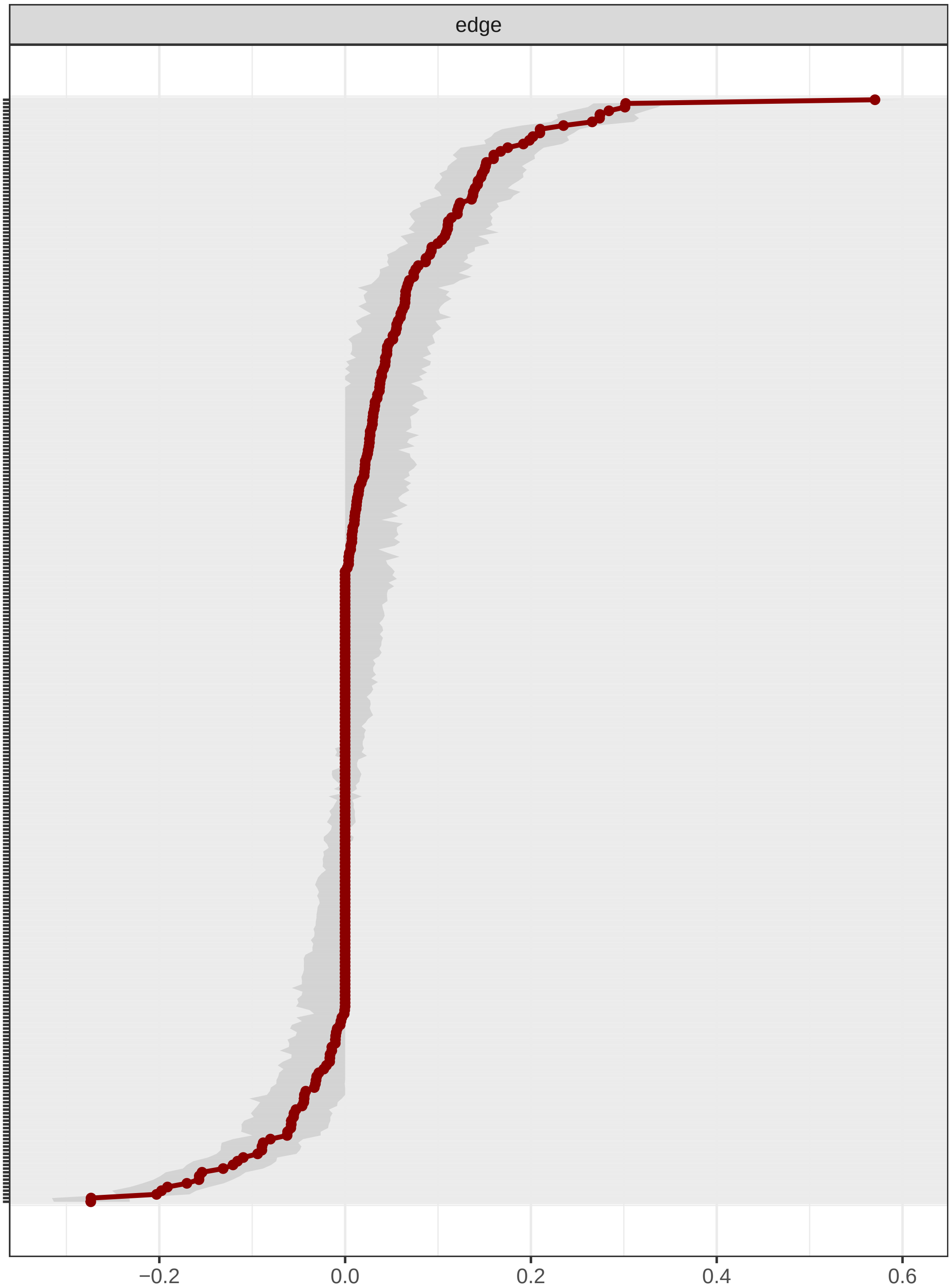
```
Res_glasso <- estimateNetwork(bfiData, default = "EBICglasso")
```

Exercise 3 (2 points) Perform a *non-parametric* bootstrap on the EBICglasso network and plot the confidence intervals of the edge-weights.

Solution:

```
Boot1 <- bootnet(Res_glasso,nBoots = 1000, nCores = 8)
```

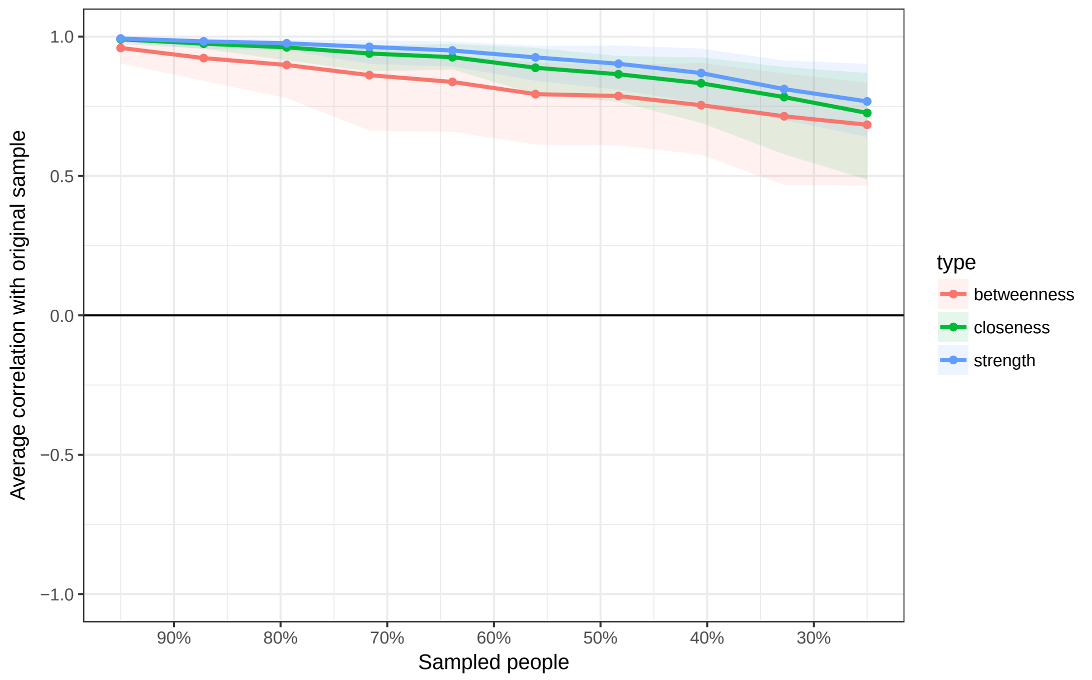
```
plot(Boot1, labels = FALSE, order = "sample")
```



Exercise 4 (2 points) Perform a *case-drop* bootstrap on the EBICglasso network and plot the stability of centrality indices.
Solution:

```
Boot2 <- bootnet(Res_glasso,nBoots = 1000, nCores = 8, type = "case")
```

```r
plot(Boot2)
```



Exercise 5 (1 point) Give the *CS*-coefficient of the three centrality indices, and explain how this measure can be interpreted.

Solution:

```r
corStability(Boot2)
```

```
## === Correlation Stability Analysis ===
##
## Sampling levels tested:
##    nPerson Drop%   n
## 1      700  75.0  99
## 2      918  67.2 105
## 3     1136  59.4 108
## 4     1353  51.7 120
## 5     1571  43.9 101
## 6     1789  36.1  82
## 7     2007  28.3  83
## 8     2224  20.6 122
## 9     2442  12.8 105
## 10    2660   5.0  75
##
## Maximum drop proportions to retain correlation of 0.7 in at least 95% of the samples:
##
## betweenness: 0.361
##   - For more accuracy, run bootnet(..., caseMin = 0.283, caseMax = 0.439)
##
## closeness: 0.594
##   - For more accuracy, run bootnet(..., caseMin = 0.517, caseMax = 0.672)
##
## strength: 0.672
##   - For more accuracy, run bootnet(..., caseMin = 0.594, caseMax = 0.75)
##
## Accuracy can also be increased by increasing both 'nBoots' and 'caseN'.
```

The measure gives the maximum proportion of cases that can be dropped to retain with 95% probability a correlation of 0.7 with the original centrality indices. ▪

Part 3: SEM re-analysis

Doosje, Loseman, and Bos (2013) analyzed radicalization of Islamic youth in the Netherlands using a large-scale structural equation model (SEM), which can be drawn as a directed causal network:

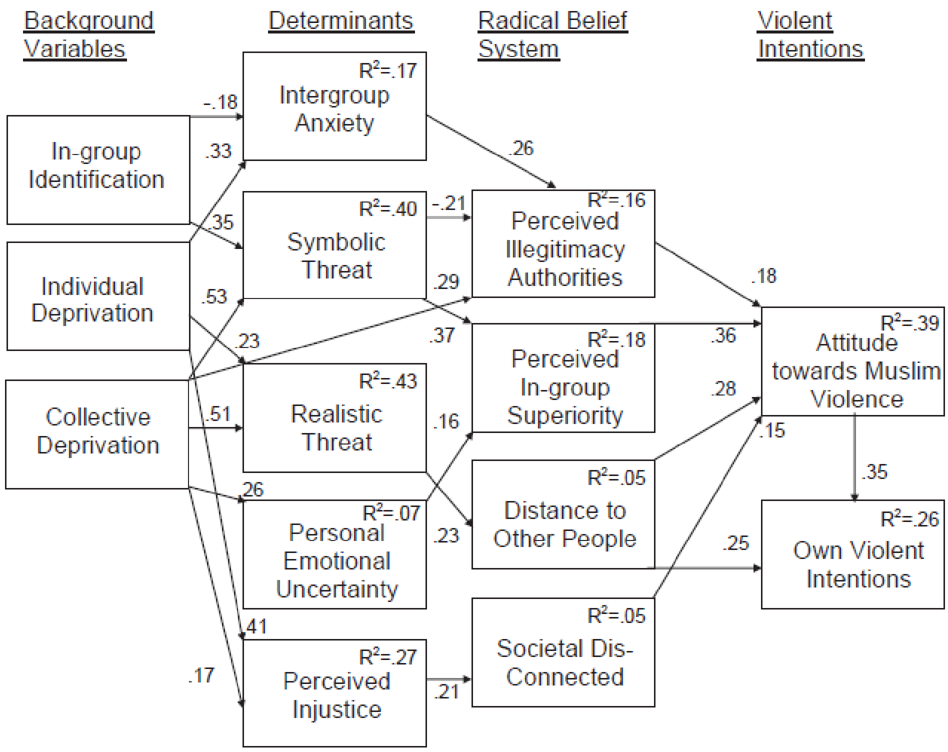


Fig. 1. Final structural equation model. All paths are significant. $R^2 = \%$ variance explained.

As with many SEM papers, Doosje et al. (2013) reported the correlation matrix and sample size ($N = 131$) to reproduce their analyses. We can load the correlation matrix in R as follows:

```
source("http://sachaepskamp.com/files/DoosjeData.R")
View(corMat)

## Error in View(corMat):  X11 dataentry cannot be loaded
```

The `corMat` object now contains the correlation matrix. A SEM analysis as shown above can be used to test a confirmatory model, as is done by Doosje et al. (2013). Suppose however we had no theory and want to exploratory find a good fitting model. SEM is less useful for exploratory model search, as there are many equivalent models possible that fit just as well. For this reason, undirected networks offer a powerfull tool in gaining exploratory insight in which variables might interact.

Exercise 6 (3 points)
Estimate a Gaussian graphical model using LASSO regularization and EBIC model selection on the data from Doosje et al. (2013). Note that you do not have the raw data, so you can not use `estimateNetwork` and need to use the underlying estimation function from the *qgraph* package (`EBICglasso`). Set the EBIC tuning parameter γ to zero. Compare your estimated network to the SEM model reported. Are there edges in your network that are not included in the model shown by Doosje et al. (2013)?

Solution:

```
# Optional: let's replicate the layour somewhat of the original SEM model:
Layout <- matrix(NA,14,2)

# First column:
Layout[1:3,1] <- 0
Layout[1:3,2] <- seq(0.9,0.5,length=3)
```

```
# Second column:
Layout[4:8,1] <- 1
Layout[4:8,2] <- seq(1,0,length=5)

# Third column:
Layout[9:12,1] <- 2
Layout[9:12,2] <- seq(0.75,0.1,length=4)

# Fourth column:
Layout[13:14,1] <- 3
Layout[13:14,2] <- seq(0.5,0.2,length=2)

# Estimate the network (with some extra optional arguments):
qgraph(corMat, graph = "glasso", sampleSize = 131, tuning = 0,
       layout = Layout, labels = colnames(corMat),
       shape = "rectangle", vsize = 15, vsize2 = 8,
       theme = "colorblind")
```

Comparing the models:

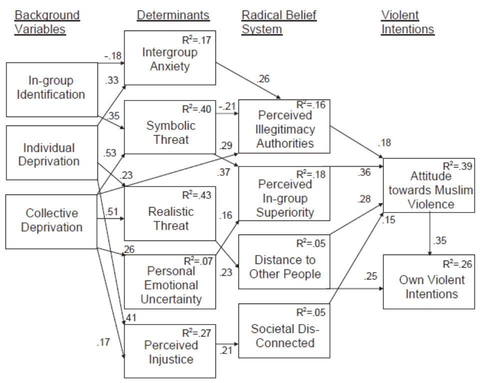
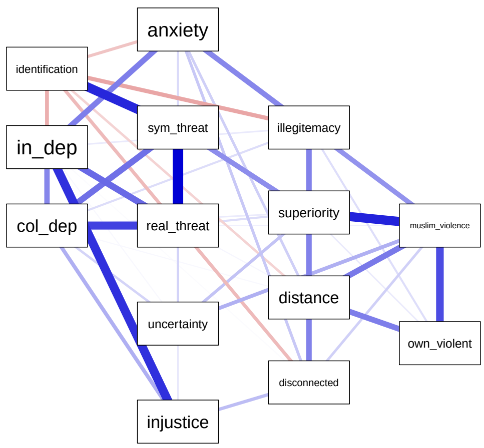


Fig. 1. Final structural equation model. All paths are significant. R^2 = % variance explained.



Some things that can be seen: most links are the same. Some links emerged between items of the same group, but these were actually also added in the SEM model as correlations and not shown (you did not have to read the paper so it is fine if you did not catch this). The most notable node is in-group identification, which features noteworthy *negative* connections to perceived illegitimacy, distance to other people and societal disconnected (possibly such negative connections can emerge due to the influence of a collider structure). Perceived illegitimacy is still directly linked to collective deprivation but now also to individual deprivation. The dependent variable, own violent intent, is now directly linked to perceived illegitimacy and perceived in-group superiority. Overall, many of the connections of the SEM model are retrieved but also more connections show up.

Optional, we could also see how these results differ with $\gamma = 0.5$:

```
# Estimate the network (with some extra optional arguments):
qgraph(corMat, graph = "glasso", sampleSize = 131, tuning = 0.5,
       layout = Layout, labels = colnames(corMat),
       shape = "rectangle", vsize = 15, vsize2 = 8,
       theme = "colorblind", title = "gamma = 0.5")
```

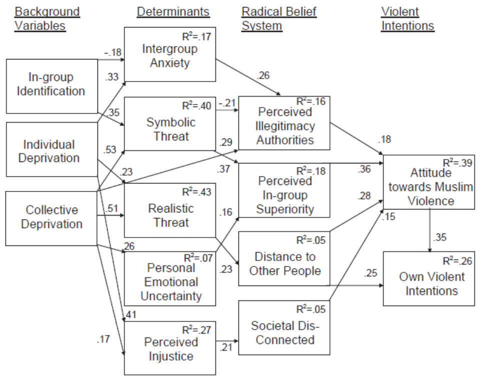
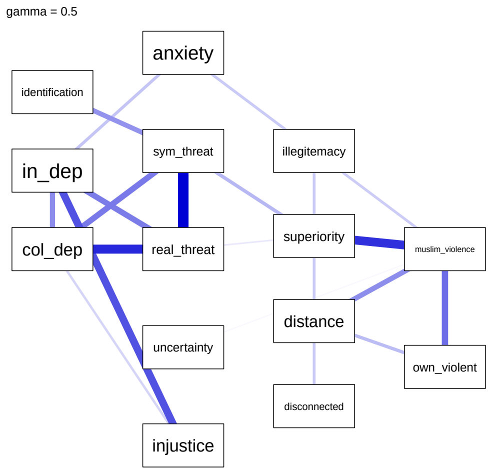


Fig. 1. Final structural equation model. All paths are significant. R^2 = % variance explained.



We obtain a much sparser model. Interestingly, realistic threat is linked to perceived in-group superiority in our model but not in the original SEM model.

References

Doosje, B., Loseman, A., & Bos, K. (2013). Determinants of radicalization of islamic youth in the netherlands: Personal uncertainty, perceived injustice, and perceived group threat. *Journal of Social Issues*, *69*(3), 586–604.

Fried, E. I., Bockting, C., Arjadi, R., Borsboom, D., Amshoff, M., Cramer, O. J., … Stroebe, M. (2015). From loss to loneliness: The relationship between bereavement and depressive symptoms. *Journal of abnormal psychology*, *124*(2), 256–265.