Exploring the Dynamics between Social Media Use, Real-Life Socializing, Avoidance, and Mood

Emily Campos Sindermann, Margarida Da Silva, Teresa Gehrig, Mohammad Hamdan

2021-12-21

Contents

1 Introduction

In this project we wanted to investigate the relationship between social media use, real-life socializing and mood. To this end, two of out team members Margarida and Mohammad collected data on the following variables:

(1) ***Social media use***: included both active and passive online behaviors (Trifiro & Gerson, 2019): active social media use comprised online behaviors that facilitate "direct exchanges" among users (e.g. liking, commenting, sending messages, and otherwise engaging with other users); passive social media use comprised monitoring of others without direct engagement (e.g. checking if someone has responded to a message). Subjects had to indicate approximately how many minutes they have engaged in online social activity since the last prompt.

(2) ***Real-life socializing***(hereinafter: socializing): Defined as any kind of voluntary interaction with other people for leisure. This includes e.g., talking to friends during class, but does not include ordering coffee from a barista. Subjects had to indicate approximately how many minutes they have engaged in offline social activity since the last prompt.

(3) ***Mood***

(a) ***Positive mood***: Included moods such as energetic, enthusiastic and content (Fisher et al., 2017)

(b) ***Negative mood***: Including moods such as angry and afraid (Fisher et al., 2017).

(3) ***Behavioral avoidance***: refraining from, or escaping from, an action, person or thing (Ottenbreit & Dobson, 2004)

(a) ***Avoiding people***

(b) ***Avoiding activities***

(c) ***Procrastination***

The subjects filled in a Qualtrics survey four times a day (1 p.m., 4 p.m., 7 p.m., 10 p.m.), starting Nov 17, 2021 and ending on Dec 12, 2021 (a total of 26 days). This yielded 104 time-points for each subject, which we used to estimate and compare their personalized networks.

2 Data Import

```
data <- read_csv("Network Analysis final data.csv")
data <- data[-c(1:2), ]
data <- data %>%
  select("date_time" = RecordedDate,
         "person" = Q3,
         "social_media" = Q6,
         "socializing" = Q7,
         "avoid_people" = Q8,
         "avoid_activities" = Q9,
         "procrastinating" = Q10,
         "positive" = Q1,
         "negative" = Q5,
         "submit" = Q4) %>%
    filter(submit == 1) %>%
    add_column("day" = rep(c(1:26), each = 8), #counting days
               "time" = rep(c(1:4), 52),
               "conc" = c(1:208))%>%
    select(-submit) %>%
    relocate(conc, day, time, #reorganizing columns
           social_media, socializing,
           avoid_people, avoid_activities, procrastinating, positive, negative,
           person, date_time)

data[, 4:(ncol(data)-1)] <- sapply(data[, 4:(ncol(data)-1)], as.numeric)

#separate datasets
margarida <- filter(data, person == 2)
mohammad <- filter(data, person == 1)
data$person <- ifelse(data$person == 2, "Margarida", "Mohammad")
```

3 Descriptive Statistics

```
ggplot(data, mapping = aes(y = socializing))+
  geom_boxplot(fill = c("paleturquoise1", "lightseagreen"))+
  scale_x_discrete(NULL)+
  facet_wrap(~person)+
  labs(y = "Time (in minutes)")
```

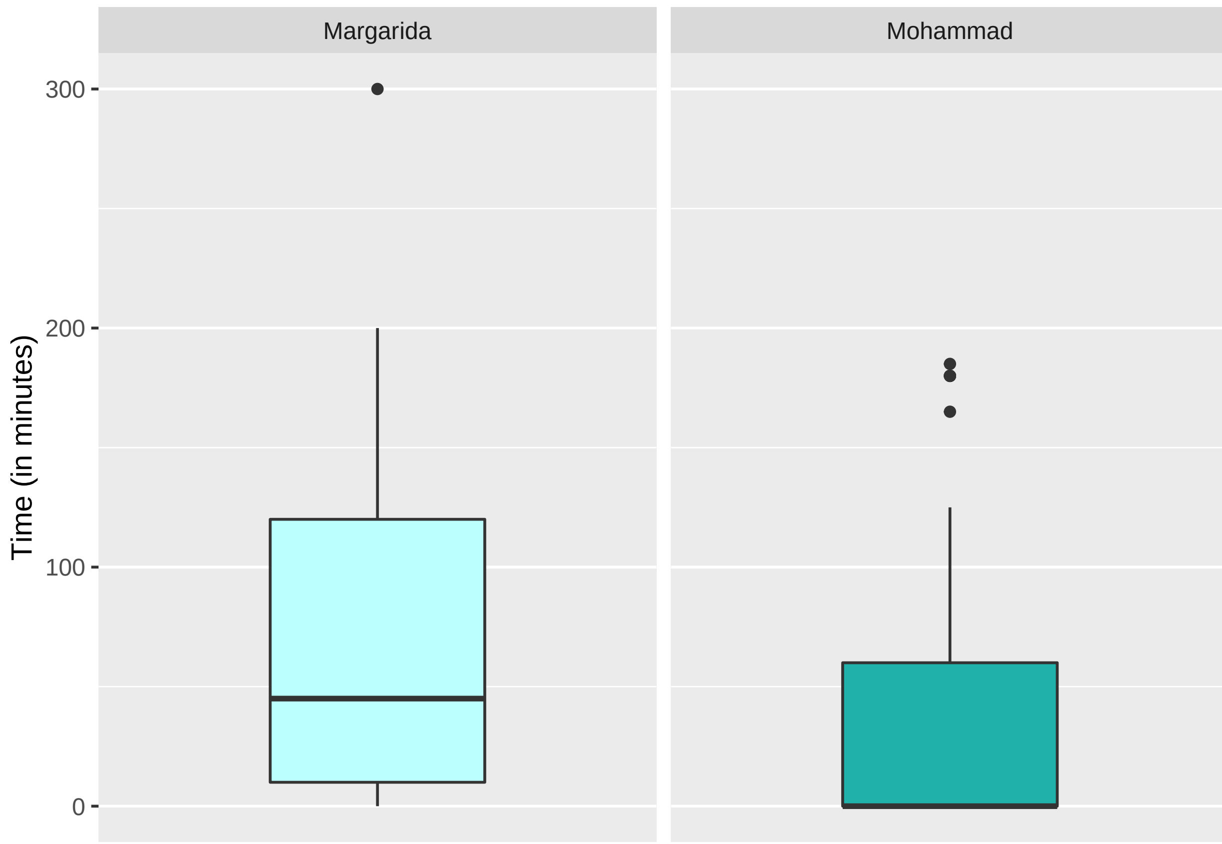

Figure 1: Socializing

```
ggplot(data, mapping = aes(y = social_media))+
  geom_boxplot(fill = c("paleturquoise1", "lightseagreen"))+
  facet_wrap(~person)+
  scale_x_discrete(NULL)+
  labs(y = "Time (in minutes)")
```

4 Stationarity

Before we start the analysis of our time-series data, we first check if the assumption of stationarity holds. To do so, we regress each variable on time. If the summary indicates time to be a significant predictor, we assume that the assumption of stationarity is violated. To give an idea of how we assess stationarity, we summarize the linear model for one variable and plot the time-series. We choose not to report this process for every variable.

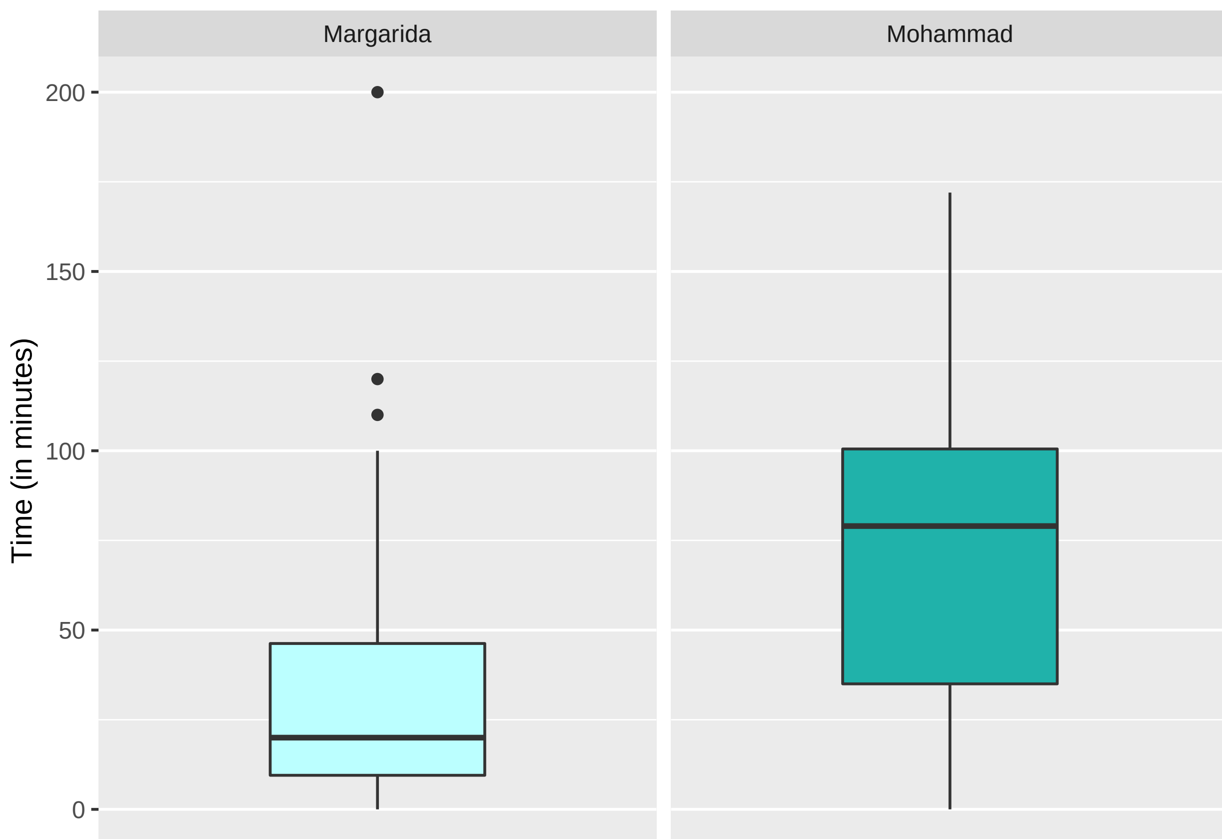

Figure 2: Social media use

For one of our subjects (Margarida), we found **socializing** to be non-stationary (see Fig. 3).

```
#Margarida
mar_stationarity <- data.frame(
  "social_media" = summary(lm(social_media~conc, data=margarida))$coef["conc",4],
  "socializing" = summary(lm(socializing~conc, data = margarida))$coef["conc",4],
  "avoid_people" = summary(lm(avoid_people~conc, data = margarida))
  $coef["conc",4],
  "avoid_activities" = summary(lm(avoid_activities ~ conc, data = margarida))
  $coef["conc",4],
  "procrastinating" = summary(lm(procrastinating~conc, data = margarida))
  $coef["conc",4],
  "positive_mood" = summary(lm(positive~conc, data = margarida))$coef["conc",4],
  "negative_mood" = summary(lm(negative ~ conc, data = margarida))
  $coef["conc",4]
)

mar_stationarity %>%
  kable(caption = "Margarida - p-values of Stationarity Check") %>%
  kable_styling(latex_options = c("striped", "HOLD_position", "scale_down"))
```

Table 1: Margarida - p-values of Stationarity Check

social_media	socializing	avoid_people	avoid_activities	procrastinating	positive_mood	negative_mood
0.0285338	0.0011638	0.0041077	0.0058276	0.0595329	0.4008084	0.2141861

```
#add short table description
```

```
#plot socializing (time as significant predictor!)
plot(ts(margarida$socializing, start = 1, end = 26, frequency = 4),
     ylab = "Time (in minutes)",
     xlab = "Day")
```

```
#we see a slight trend here
```

```
#De-Trending significant vars
margarida$socializing[!is.na(margarida$socializing)] <-
  residuals(lm(socializing~conc, data = margarida))

margarida$avoid_activities[!is.na(margarida$avoid_activities)] <-
  residuals(lm(avoid_activities~conc, data = margarida))

margarida$avoid_people[!is.na(margarida$avoid_people)] <-
  residuals(lm(avoid_people~conc, data = margarida))

margarida$procrastinating[!is.na(margarida$procrastinating)] <-
  residuals(lm(procrastinating~conc, data = margarida))

margarida$social_media[!is.na(margarida$social_media)] <-
```

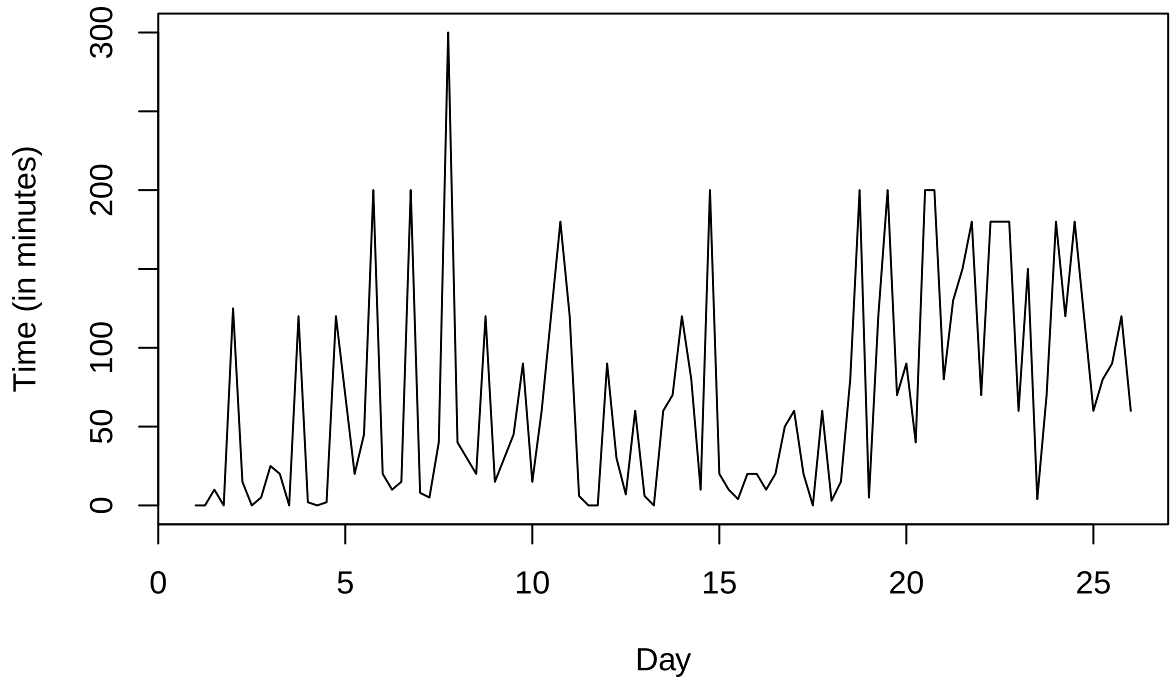

Figure 3: Margarida - Socializing

```
  residuals(lm(social_media~conc, data = margarida))
```

```
moh_stationarity <- data.frame(
  "social_media" = summary(lm(social_media~conc, data=mohammad))$coef["conc",4],
  "socializing" = summary(lm(socializing~conc, data = mohammad))$coef["conc",4],
  "avoid_people" = summary(lm(avoid_people~conc, data = mohammad))
  $coef["conc",4],
  "avoid_activities" = summary(lm(avoid_activities ~ conc, data = mohammad))
  $coef["conc",4],
  "procrastinating" = summary(lm(procrastinating~conc, data = mohammad))
  $coef["conc",4],
  "positive_mood" = summary(lm(positive~conc, data = mohammad))$coef["conc",4],
  "negative_mood" = summary(lm(negative ~ conc, data = mohammad))$coef["conc",4]
)

moh_stationarity %>%
  kable(caption = "Mohammad - p-values of Stationarity Check") %>%
  kable_styling(latex_options = c("striped", "HOLD_position", "scale_down"))
```

Table 2: Mohammad - p-values of Stationarity Check

social_media	socializing	avoid_people	avoid_activities	procrastinating	positive_mood	negative_mood
0.0630517	0.9227612	0.1884441	0.3750369	0.1301331	0.6164476	0.0015054

```r
##De-Trend
mohammad$negative[!is.na(mohammad$negative)] <-
  residuals(lm(negative~conc, data = mohammad))
```

We detrended the variables that violated the stationarity assumption.

5 Network Estimation

In order to explore the different estimation techniques for time-series networks, we estimated the temporal as well as contemporaneous networks of Margarida and Mohammad with both the regularized (`graphicalVAR`) and the maximum likelihood estimation (`gvar`). For visual comparison we used the average layout and the same maximum value in the `qgraph` function. To make our networks more inclusive, we choose a colorblind theme.

First, we pre-defined the variables corresponding to the nodes in the time-series networks (i.e., social media use (in minutes), socializing (in minutes), behavioral avoidance (avoiding people, avoiding activities, procrastination), positive and negative mood). We grouped them within a list for subsequent graphical choices.

```r
vars <- c("social_media",
          "socializing",
          "avoid_people",
          "avoid_activities",
          "procrastinating",
          "positive",
          "negative")



groups <- list(
  "Social media use" = 1,
  "Socializing" = 2,
  "Behavioral avoidance" = c(3:5),
  "Mood" = c(6, 7)
)
```

```r
colGrp <- c("lightgoldenrod", "tan2", "mediumaquamarine", "plum3")
```

5.1 Regularized Estimation using `graphicalVAR`

First we used the regularized estimator `graphicalVAR`. By indicating the beepvariable *time* we ensured that the algorithm corrects for the night between measures, thereby meeting the assumption of equidistant measures. Setting the gamma-parameter to zero, we obtain a network optimizing the BIC criterion.
We first estimated the networks and then visually compared them between subjects.

```r
network_mar <- graphicalVAR(margarida, vars = vars,
          beepvar = "time", dayvar = "day", gamma = 0)
```

```
##    |                                                                      |
```

```r
network_moh <- graphicalVAR(mohammad, vars = vars,
         beepvar = "time", dayvar = "day", gamma = 0)
```

```
##    |                                                                              |
```

```r
#gamma = 0 -> optimizes BIC criterion (not default EBIC)

L1 <- averageLayout(network_mar$PCC, network_moh$PCC)
```

5.1.1 Contemporaneous Networks

First, we compare the contemporaneous networks of Mohammad and Margarida. We make sure to use an average layout to ensure (visual) comparability. We manually set the same `maximum` argument in the `qgraph()`function.
To make our networks more inclusive, we choose a colorblind theme.

```r
qgraph(network_mar$PCC,
       layout = L1,
       theme = "colorblind",
       labels = vars,
       maximum = max(network_mar$PCC, network_moh$PCC),
       groups = groups,
       color = colGrp,
       legend.cex = 0.5,
       bg="transparent",
       threshold = 0.05)
```

```r
qgraph(network_moh$PCC,
       layout = L1,
       theme = "colorblind",
       labels = vars,
       maximum = max(network_mar$PCC, network_moh$PCC),
       groups = groups,
       color = colGrp,
       bg = NULL,
       transparency = TRUE,
       legend.cex = .5,
       bg="transparent",
       threshold = 0.05)
```

5.1.2 Correlating weights matrices

```r
corMat <- cor(network_mar$PCC, network_moh$PCC)
#partial contemporaneous correlations (PCC)


corMat %>%
  kable(caption = "Correlation of Weights Matrices, graphicalVAR") %>%
```

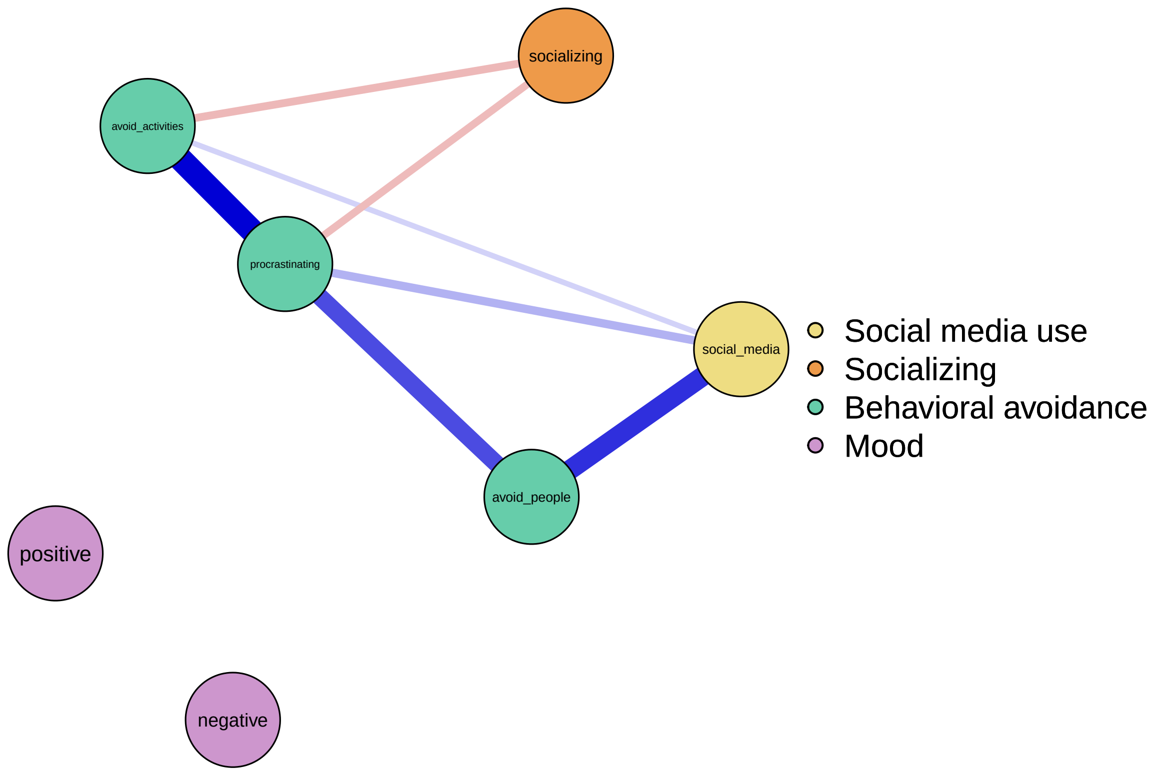

Figure 4: Margarida - Contemporaneous Network, using graphicalVAR

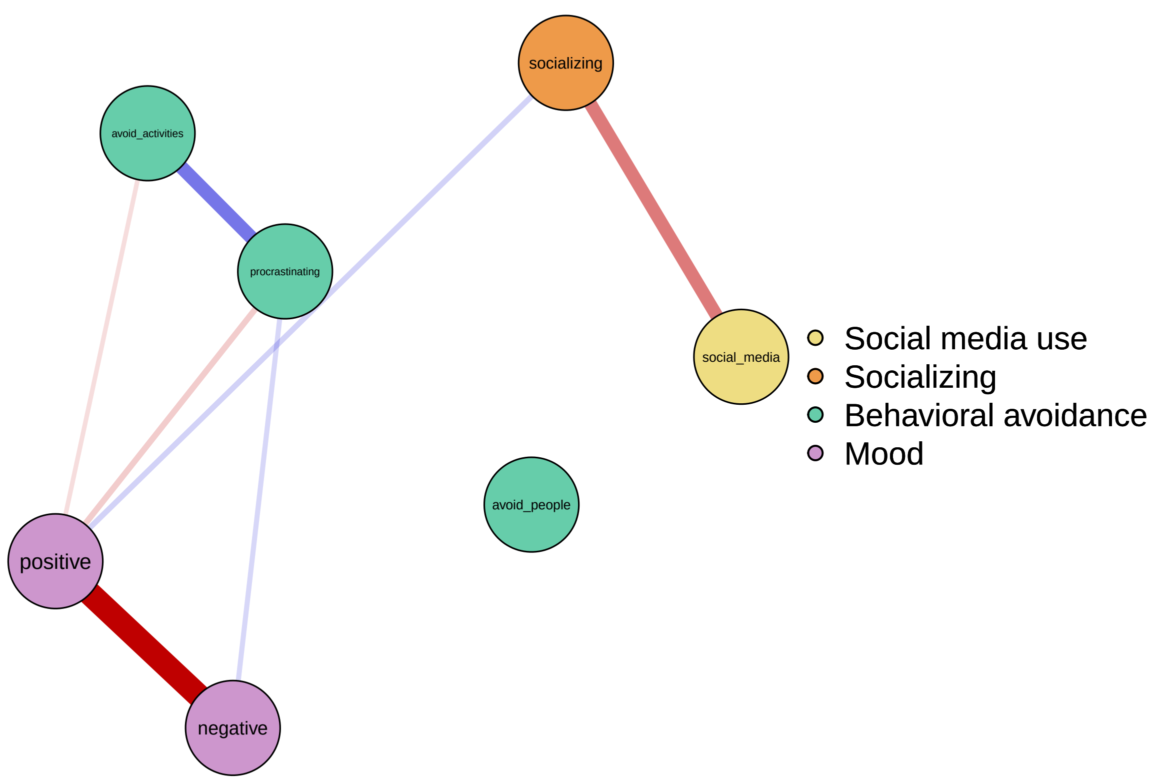

Figure 5: Mohammad - Contemporaneous Network, using graphicalVAR

```r
kable_styling(latex_options = c("striped", "HOLD_position", "scale_down"))
```

Table 3: Correlation of Weights Matrices, graphicalVAR

	social_media	socializing	avoid_people	avoid_activities	procrastinating	positive	negative
social_media	0.2697066	0.1572276	0.2697066	0.2241709	0.0064665	0.1164939	0.2860224
socializing	-0.2580874	-0.1504541	-0.2580874	-0.6420781	-0.6075588	-0.0156466	-0.3432273
avoid_people	0.2928900	-0.7260653	0.4199833	0.5969724	0.1525099	0.1394238	0.4857010
avoid_activities	0.4907730	-0.0592871	0.1901625	0.8729122	-0.1552687	-0.0624058	0.3179101
procrastinating	0.5018847	-0.1238404	0.2429874	-0.1700472	0.6633601	0.0654345	0.1971937
positive	-0.2581861	-0.1505117	-0.2581861	0.1806836	-0.0193008	0.6390628	-0.2095281
negative	-0.1666667	-0.4647259	1.0000000	0.3898026	0.5143401	-0.2019356	0.9887241

The correlations between the contemporaneous networks estimated via `graphicalVAR` show that our two subjects' weights matrices were not highly correlated in most cases. The highest correlations could be seen between the same nodes of both subjects (e.g. r(positive-positive) = .81).

5.1.3 Temporal Networks

```r
L2 <- averageLayout(network_mar$PDC, network_moh$PDC)
# partial directed correlations (PDC)
qgraph(network_mar$PDC,
       layout = L2,
       theme = "colorblind",
       labels = vars,
       groups = groups,
       color = colGrp,
       bg = NULL,
       transparency = TRUE,
       legend.cex = .5,
       bg="transparent",
       threshold = 0.05)
```

```r
qgraph(network_moh$PDC,
       layout = L2,
       theme = "colorblind",
       labels = vars,
       groups = groups,
       color = colGrp,
       bg = NULL,
       transparency = TRUE,
       legend.cex = .5,
       bg="transparent",
       threshold = 0.05)
```

Using the regularized estimation we get very sparse networks.
When optimizing the **EBIC** which we did in a previous analysis, we even found both networks to be empty. This is due to the edges being pulled to zero by the algorithm.

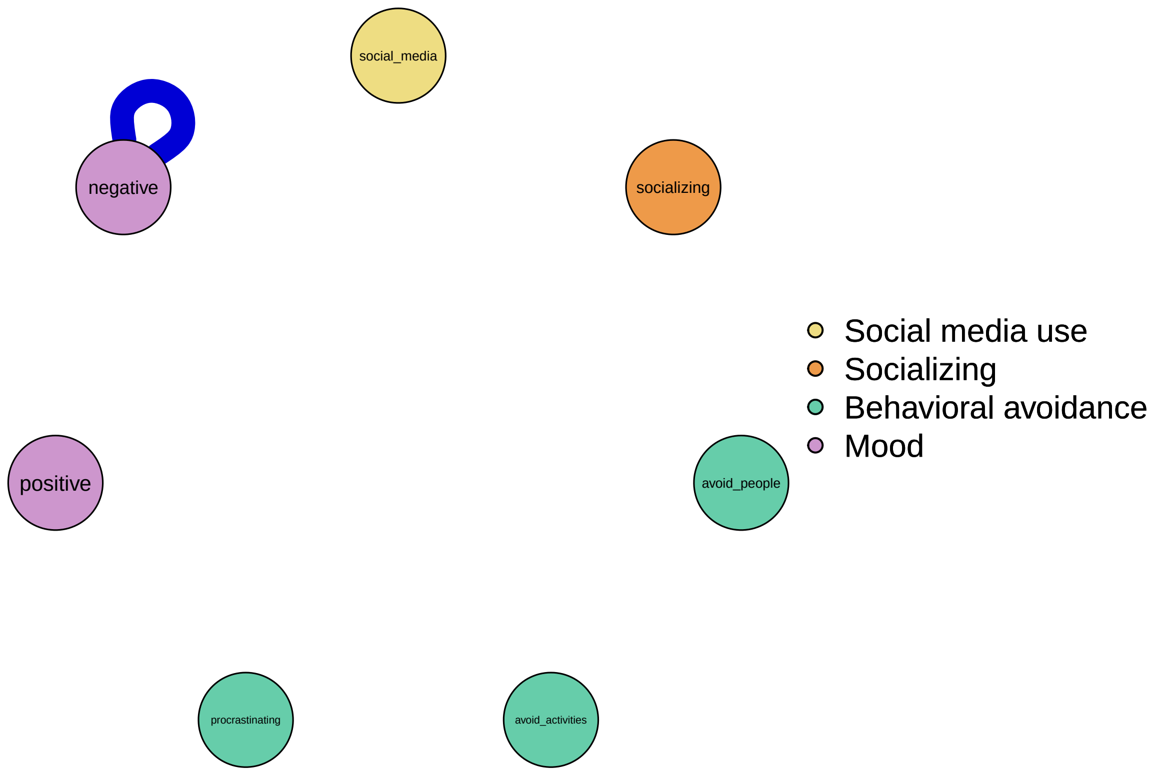

Figure 6: Margarida - Temporal Network, using graphicalVAR

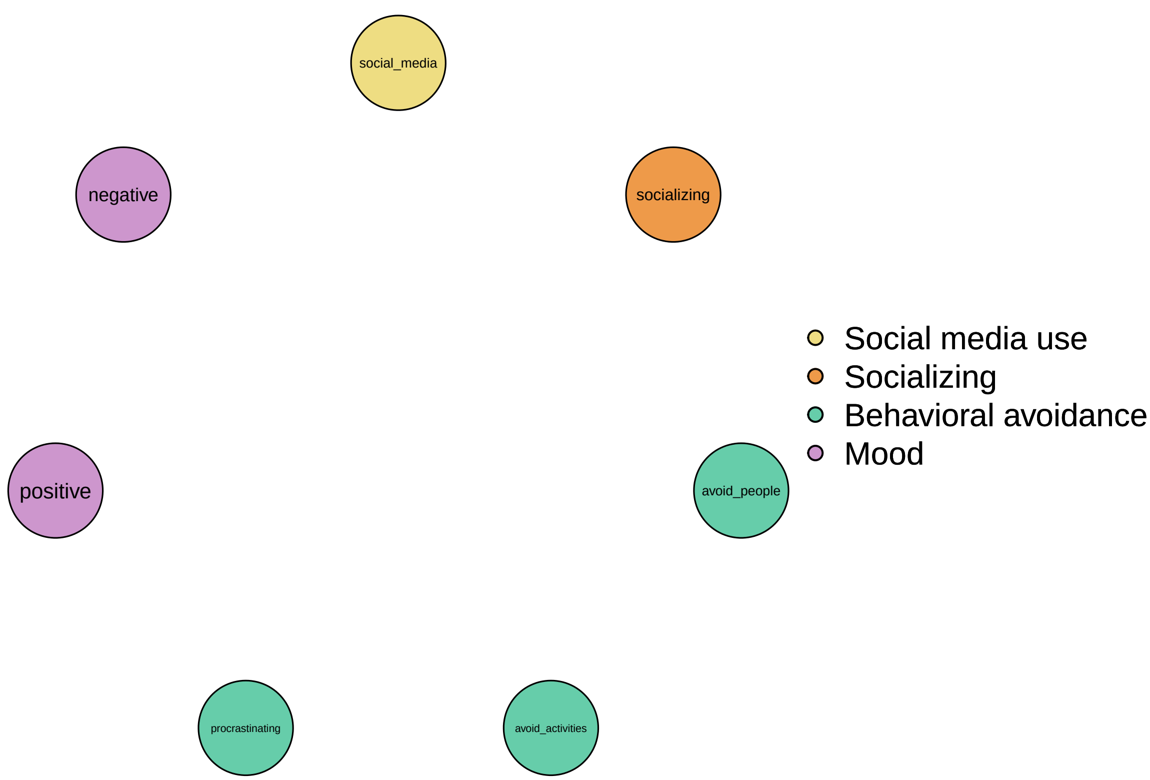

Figure 7: Mohammad - Temporal Network, using graphicalVAR

5.2 Maximum Likelihood Estimation using `gvar`

In a second step, we re-run the analysis done before, using `gvar`. This function estimates the network via maximum likelihood estimation and might yield different results than an analysis with `graphicalVAR`.

5.2.1 Model Estimation

As before, we first estimate the networks for both, Mohammad and Margarida.

```
#Margarida
network_mar2 <- gvar(margarida, vars = vars, dayvar = "day", beepvar = "time",
estimator = "FIML") %>% runmodel
#Mohammad
network_moh2 <- gvar(mohammad, vars = vars, dayvar = "day", beepvar = "time",
estimator = "FIML") %>% runmodel
```

Then, we store the weights matrices for the contemporaneous and temporal network for both subjects.

```
#Margarida's weights-matrices
mar_matrix <- getmatrix(network_mar2, "omega_zeta") #contemp.
mar_matrix2 <- getmatrix(network_mar2, "PDC") #temporal

#Mohammad's weights-matrices
moh_matrix <- getmatrix(network_moh2, "omega_zeta") #contemp.
moh_matrix2 <- getmatrix(network_moh2, "PDC") #temporal

max1 <- max(moh_matrix, mar_matrix)
max2 <- max(moh_matrix2[,-(1:2)], mar_matrix2[,-(1:2)])
  #omitting cols 1 and 2 because they give NaN estimations
L3 <- averageLayout(mar_matrix, mar_matrix2, moh_matrix, moh_matrix2)
```

5.2.2 Contemporaneous Networks

```
qgraph(mar_matrix,
       theme = "colorblind",
       labels = vars,
       layout = L3,
       groups = groups,
       color = colGrp,
       maximum =max1,
       legend.cex = .5,
       bg="transparent",
       legend = FALSE, #set to FALSE in order to obtain a graph without a legend for the poste
       vsize = 15,
       threshold = 0.05
       )
```

```
qgraph(moh_matrix,
       labels = vars,
       theme = "colorblind",
```

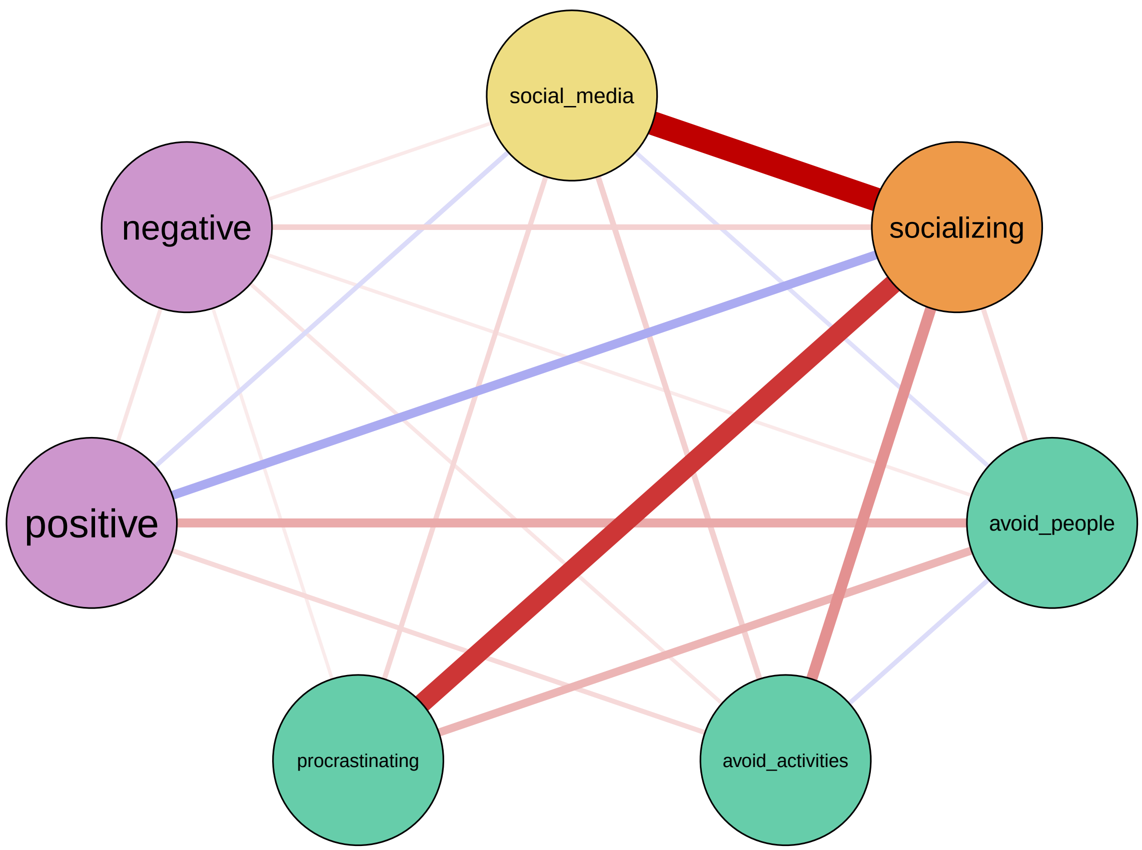

Figure 8: Margarida - Contemporaneous Network, using gvar

```
        layout = L3,
        groups = groups,
        color = colGrp,
        maximum = max1,
        legend.cex = .5,
        bg="transparent",
        legend = FALSE,
        vsize = 15,
        threshold = 0.05
)
```

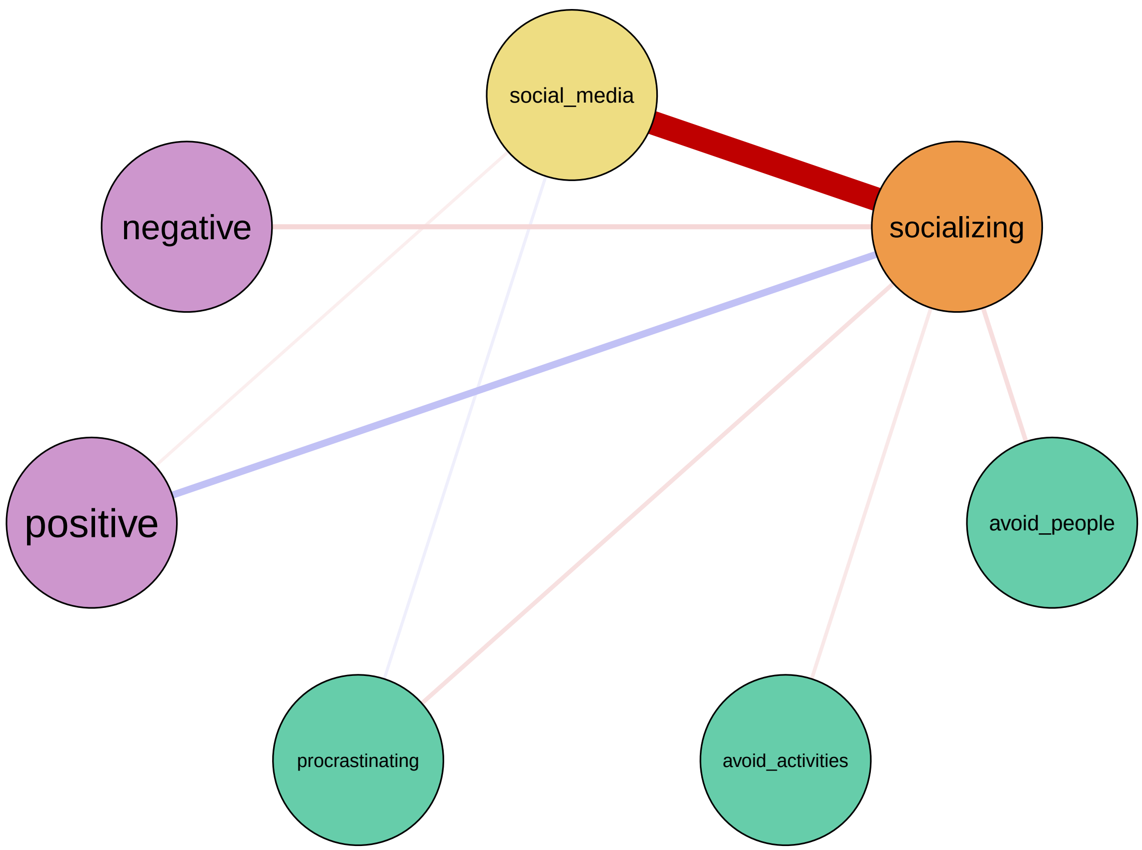

Figure 9: Mohammad - Contemporaneous Network, using gvar

The two networks show opposite effects in the relationship between social media use and real-life socializing. Mohammad's network showed a strong negative correlation, while Margarida's showed a weaker positive one. A possible explanation relates to how much social interactions are centered around social media use. For example, when checking something funny on social media with your friends.
Socializing also had opposite relations with behavioral avoidance components. Margarida had a positive correlation between socializing and behavioral avoidance, and these were strongly positively correlated between them. This means that, for Margarida, socializing with others is related to more behavioral avoidance. Mohammad, on the other hand, showed a negative correlation between socializing and behavioral avoidance, and the individual components were weakly correlated.
Finally, we also noticed a key difference in the relation between socializing and mood in the two networks.

Margarida shows an overall negative correlation between mood and socializing, but positive between social media use and positive mood. Mohammad's network showed the opposite effect. This could be because of differences in the trait of introversion, or also differences in how social media is used (e.g. to talk to others which increases positive mood despite being on social media).

5.2.3 Temporal Networks

```
qgraph(mar_matrix2,
       labels = vars,
       theme = "colorblind",
       layout = L3,
       groups = groups,
       color = colGrp,
       maximum = max2,
       legend.cex = .5,
       layoutScale = c(1,0.8),
       layoutOffset = c(0, -0.2),
       bg="transparent",
       legend = FALSE,
       vsize = 15,
       threshold = 0.05)
```

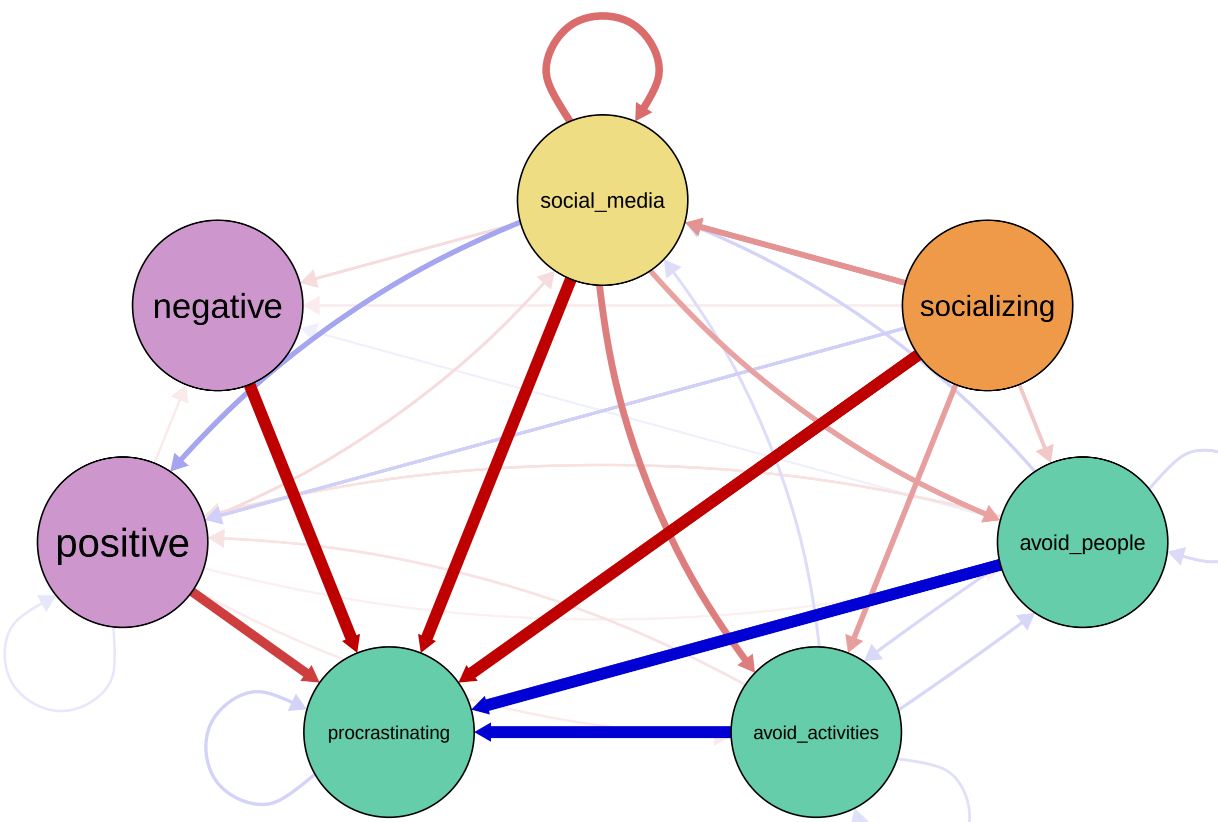

Figure 10: Margarida - Temporal Network, using gvar

```
qgraph(moh_matrix2,
       labels = vars,
       theme = "colorblind",
       layout = L3,
       groups = groups,
       color = colGrp,
       maximum = max2,
       legend.cex = .5,
       bg="transparent",
       legend = FALSE,
       vsize = 15,
       threshold = 0.05
       )
```

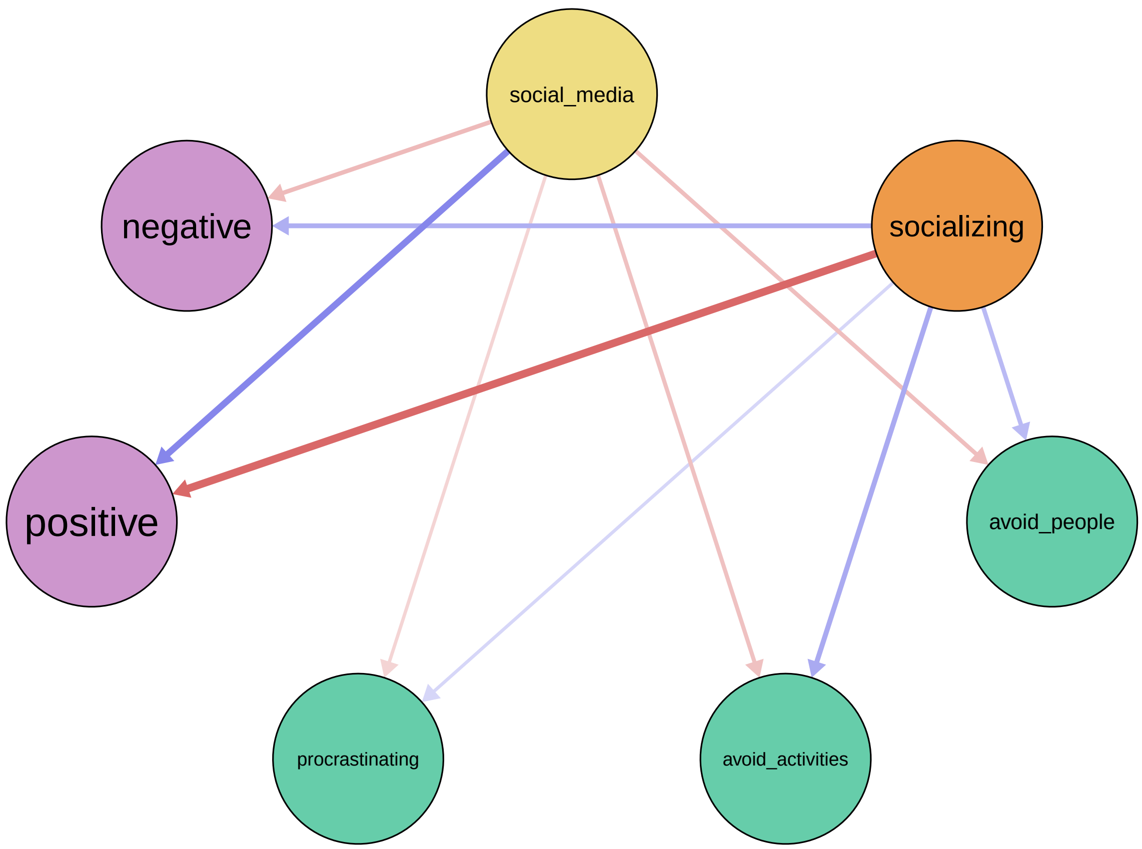

Figure 11: Mohammad - Temporal Network, using gvar

First, we noticed a key difference in the relationship between social media and socializing. Margarida's network showed a negative two-way effect between these variables, while Mohammad's showed no link. Similarly, in both networks there is an increase in behavioral avoidance after socializing. This could mean that, perhaps due to introversion, both participants found it necessary to avoid others and certain activities after having real-life interactions to "recharge".
In terms of mood, Mohammad's network showed that there was a decrease in both positive and negative mood after socializing, and an increase in positive mood after using social media. For Margarida, on the other hand, socializing tended to lead to an increase in negative mood and a decrease in positive mood later on. Additionally, using social media at one point did not seem to affect her mood later on, as there was no edge.
Finally, Margarida's network showed two weak self-loops. They show that procrastinating at one point led to more procrastinating later on, and socializing led to less socializing later on. Mohammad's network did not show any self-loops.

Interestingly, the relationships between our variables of interest indicated by the networks now look drastically different than above when we used `graphicalVAR`!

6 References

Fisher, A. J., Reeves, J. W., Lawyer, G., Medaglia, J. D., & Rubel, J. A. (2017). Exploring the idiographic dynamics of mood and anxiety via network analysis. *Journal of abnormal psychology, 126 (8)*, 1044.

Ottenbreit, N. D., & Dobson, K. S. (2004). Avoidance and depression: the construction of the Cognitive–Behavioral Avoidance Scale. *Behaviour research and therapy, 42(3)*, 293-313.