

# Election Topic Network - Czech Presidential Election 2018

*Matous Pilnacek*

*24 .7. 2018*

## Contents

<b>Introduction</b>	<b>1</b>
<b>Theory</b>	<b>1</b>
<b>Data</b>	<b>1</b>
<b>Results</b>	<b>2</b>
Zeman . . . . .	4
Drahoš . . . . .	6
<b>Validity</b>	<b>8</b>
<b>References</b>	<b>8</b>

## Introduction

### Theory

Postoj je v sociálně-psychologické literatuře někdy konceptualizován jako síť emocí, úvah a konceptů (Lodge and Taber 2013).

Dalege et al. (2017) ukazuje, že struktura sítě postojů vůči prezidentskému kandidátovi má vliv na chování voličů. Čím je síť hustší, tím má větší efekt a větší efekt mají také uzly, které mají centrálnější roli. K vytvoření sítě používá uzavřené otázky hodnotící charakter kandidáta (upřímný, inteligentní...). U každého kandidáta je tak stejný počet uzlů a jejich propojení je vytvořeno pomocí polychorických korelací.

V této studii navazujeme na síťové pojetí postoje a posouváme ho jiným směrem. Zaprvé se od postojů ke kandidátovi posouváme k otázce konceptů/témat spojených s volbami (a tedy blíže k teorii agenda-setting). A zadruhé nepoužíváme uzavřené ale otevřené otázky, což může přinést větší možnost zahrnout více podstatných témat, ale přináší další metodologické obtíže.

### Data

Vycházíme z povolební studie českých prezidentských voleb 2018 CVVM. Data ke stažení v datovém archivu ČSDA.

Byla položena otázka: *Která témata byla pro Vás nejdůležitější v prezidentských volbách před druhým kolem voleb? Vyjmenujte je v pořadí, jak byla důležitá pro Vás osobně* S možností uvést dvě témata.

Témata byla zakódována podle kódového klíče (je potřeba rozmyslet, že daný kódový klíč je vhodný).

## Results

```
library(foreign)
library(readxl)
library(plyr)
library(igraph)
library(knitr)
library(ggplot2)

NS1802 = as.data.frame(read.spss("../data/NS_1802_DataFile-FINAL.sav",
                                use.value.labels = F, reencode = "utf8"))

topics = read_excel("../data/temata_2kolo.xlsx", sheet = "temata")

NS1802$topic1 = NS1802$PS.23aa
NS1802[NS1802$topic1 %in% c(98,99,0,96,97), "topic1"] = NA
NS1802$topic2 = NS1802$PS.23ab
NS1802[NS1802$topic2 %in% c(98,99,0,96,97), "topic2"] = NA

# NS1802$vote = mapvalues(NS1802$PV.149B, c(96,99,0,9,1), c(NA,NA,NA, "Zeman", "Drahoš"))
NS1802$topic1 = mapvalues(NS1802$topic1, topics$val, topics$label)

## The following `from` values were not present in `x`: 98, 99
NS1802$topic2 = mapvalues(NS1802$topic2, topics$val, topics$label)

## The following `from` values were not present in `x`: 98, 99, 21
ties = data.frame ( table ( NS1802$topic1, NS1802$topic2 ) )
ties = ties[ties$Freq > 0,]
names(ties)[3] = "weight"
kable(head(ties))
```

	Var1	Var2	weight
4	Důchody, důchodová reforma	Bezpečnost, terorismus	3
5	Ekonomika	Bezpečnost, terorismus	3
6	EU, euro	Bezpečnost, terorismus	1
8	Kritika současné vlády	Bezpečnost, terorismus	1
9	Migrace	Bezpečnost, terorismus	4
11	Osobnost kandidáta	Bezpečnost, terorismus	2

```
nodes = data.frame ( table ( c(NS1802$topic1, NS1802$topic2 ) ) )
kable(head(nodes))
```

Var1	Freq
Bezpečnost, terorismus	33
Církevní restituce	7
Daně	7
Důchody, důchodová reforma	24
Ekonomika	38
EU, euro	159

```
g = graph_from_data_frame(ties, directed = FALSE, vertices = nodes)
g = simplify(g, edge.attr.comb = list("sum"))
plot(g, vertex.size = log(V(g)$Freq)*4, edge.width = E(g)$weight)
```

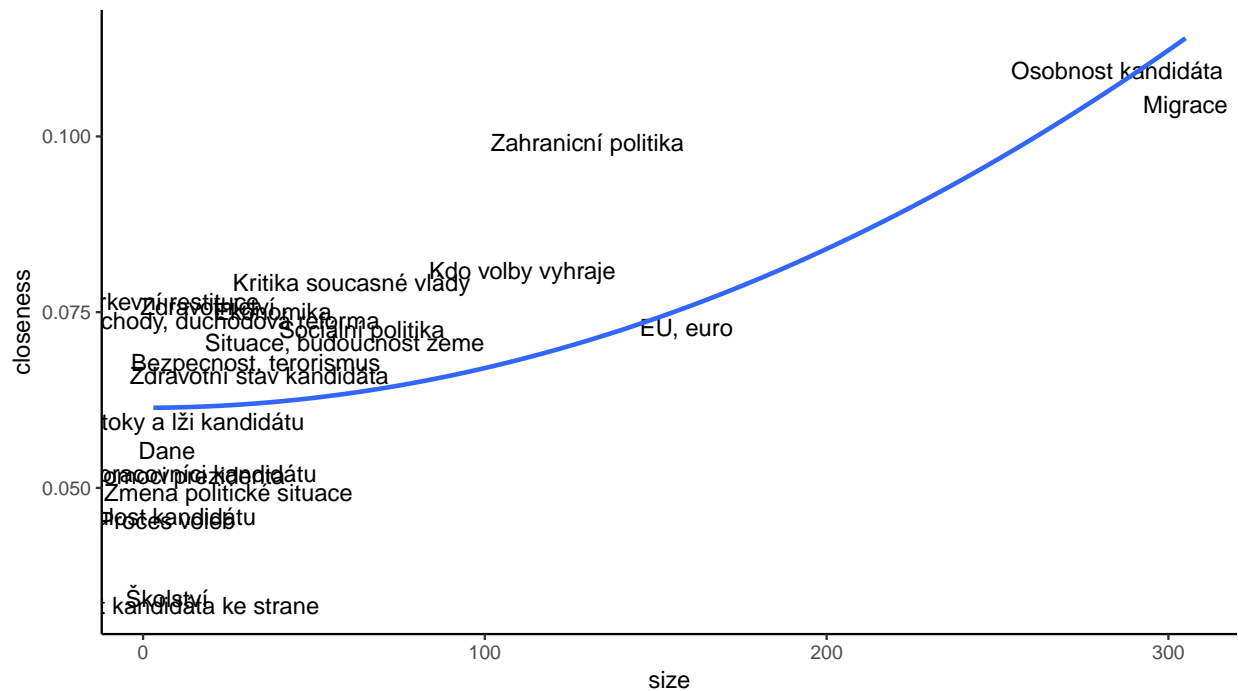


```
closeness = data.frame(closeness(g, weights = 1/E(g)$weight))
closeness = cbind(rownames(closeness), closeness)

closeness = merge(nodes, closeness, by.x = "Var1", by.y = "rownames(closeness)")
names(closeness) = c("label", "size", "closeness")
```

Vztah mezi dvěma tématy vzniká díky tomu, že jsou 1. pro danou osobu obě významná 2. je mezi nimi v paměti asociační vazba

```
ggplot(closeness, aes(x=size, y=closeness)) + geom_text(label=closeness$label) +
  stat_smooth(method = "lm", formula = y ~ I((x^2)), size = 1, se = FALSE) + theme_bw() +
  theme(panel.border = element_blank(), panel.grid.major = element_blank(),
```



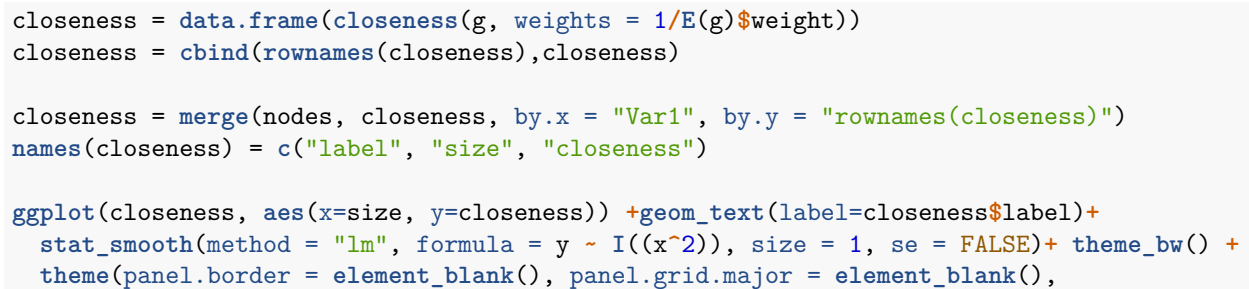
Takto agregovaná síť je problematická, protože agreguje dohromady všechny respondenty, kteří mohou mít ale síť postojů výrazně odlišnou. Je proto vhodné síť rozdělit mezi více skupin. V našem případě na voliče Zemana a Drahoše

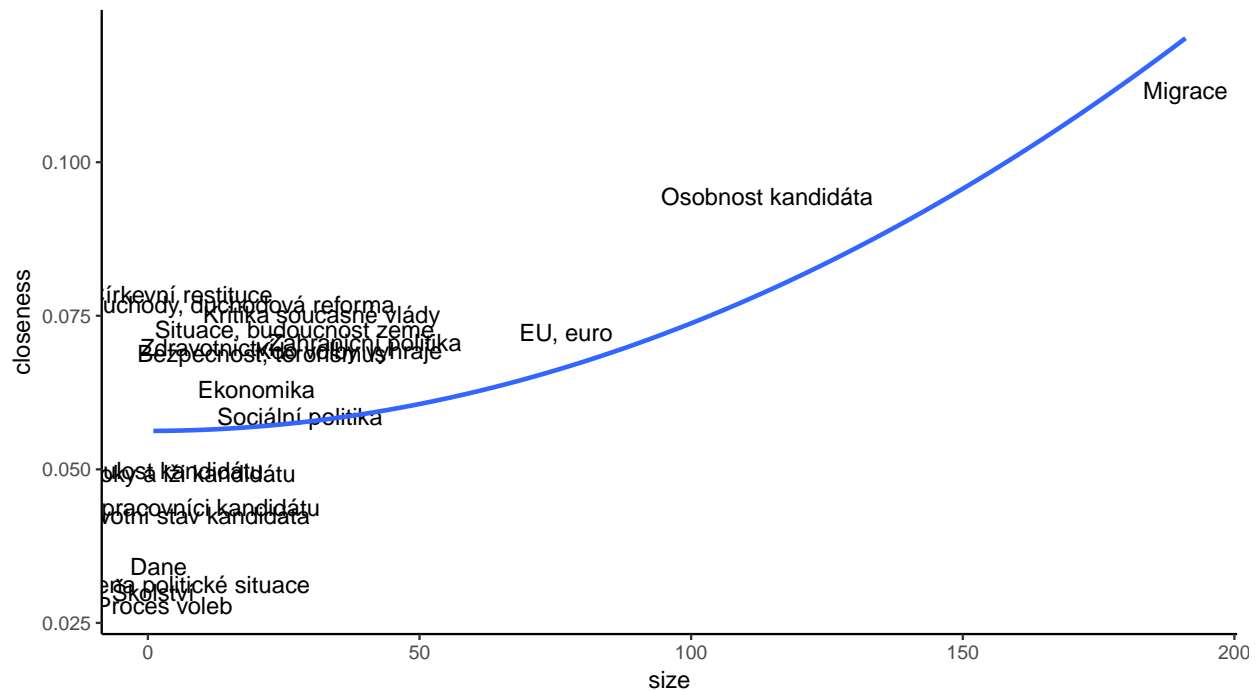
## Zeman

```
zeman = NS1802[NS1802$PV.149B %in% 9,]
ties = data.frame ( table ( zeman$topic1, zeman$topic2 ) )
ties = ties[ties$Freq > 0,]
names(ties)[3] = "weight"

nodes = data.frame ( table ( c(zeman$topic1, zeman$topic2) ) )

g = graph_from_data_frame(ties, directed = FALSE, vertices = nodes)
g = simplify(g, edge.attr.comb = list("sum"))
plot(g, vertex.size = log(V(g)$Freq)*4, edge.width = E(g)$weight)
```





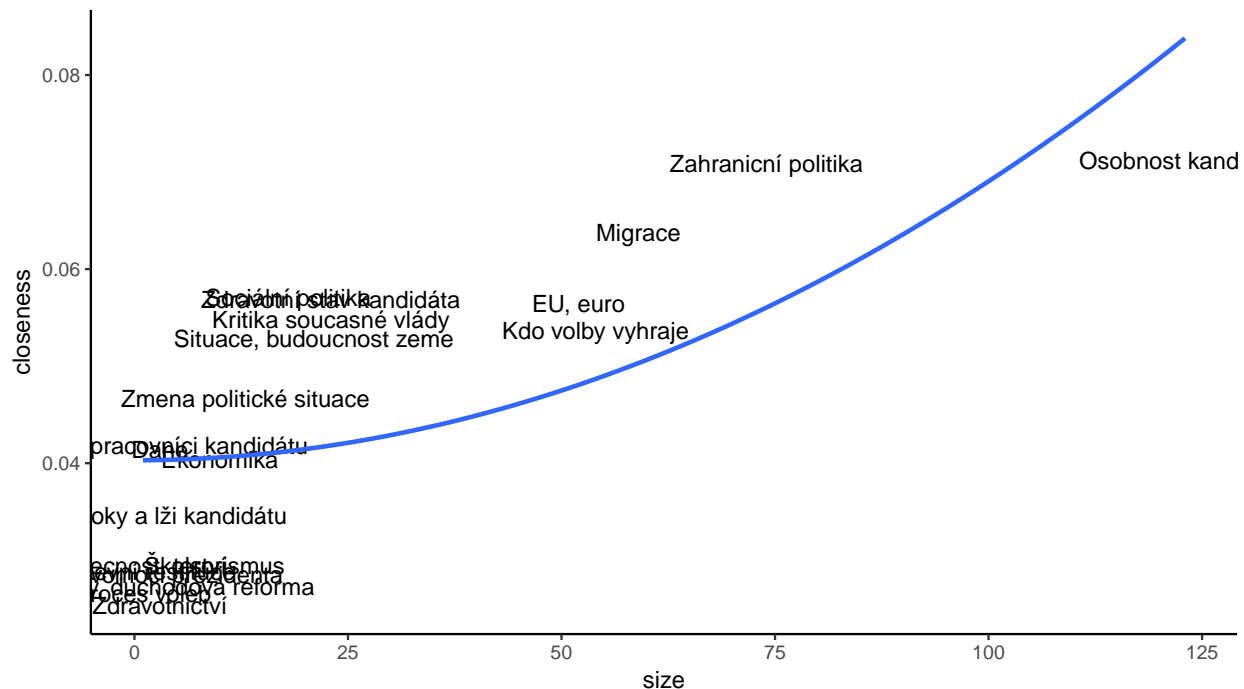
## Drahoš

```
drahos = NS1802[NS1802$PV.149B %in% 1,]
ties = data.frame ( table ( drahos$topic1, drahos$topic2) )
ties = ties[ties$Freq > 0,]
names(ties)[3] = "weight"

nodes = data.frame ( table ( c(drahos$topic1, drahos$topic2) ) )

g = graph_from_data_frame(ties, directed = FALSE, vertices = nodes)
g = simplify(g, edge.attr.comb = list("sum"))
plot(g, vertex.size = log(V(g)$Freq)*4, edge.width = E(g)$weight)
```





## Validity

Jak ověřit důležitost zapojení v síti? **Návrh: pomocí regrese**

Hledá se vhodná závislá proměnná.

Vhodné by mohlo být například zapojení do diskuze pod internetovým zpravodajstvím na iDNES.cz během druhého kola prezidentské volby pro podporce Zemana a Drahoše (počet lajků/dislajků u komentářů pro Zemana/Drahoše). Články by byly tagovány k jednotlivým tématům ze sítě.

Nezávislé proměnné by byla velikost uzlu, closeness a míra zájmu politiků lidí, kteří dané téma zmiňovali. Otázka je, zda closeness hraje roli a jak velkou.

## References

Dalege, Jonas, Denny Borsboom, Frenk van Harreveld, Lourens J. Waldorp, and Han L. J. van der Maas. 2017. "Network Structure Explains the Impact of Attitudes on Voting Decisions." *Scientific Reports* 7 (1). doi:10.1038/s41598-017-05048-y.

Lodge, Milton, and Charles S. Taber. 2013. *The Rationalizing Voter*. 1 edition. Cambridge University Press.