Derviving Voting Blocs

Contents

Overview

This section generates historical voting blocs from Eurovision voting data. The social network methods used to identify the voting blocs are cited from multiple academic papers. The voting blocs can be used as independent variables to explain contest scores. The countries will be clustered together into voting blocs using the average vote score. The connectedness of the clusters / communities will be evaluated using modularity. The modularity of a graph with respect to some division (or vertex types) measures how good the grouping is. The higher the modularity the greater the division between communities and the better grouping

```
# Load in relevant libraries
library(rmarkdown)
library(knitr)
library(igraph)
library(dplyr)

# load in the historic voting data for deriving the voting blocs
past_voting_data <- read.csv(file = "./data/ESC_hist_voting_data.csv", header = T)
```

Data Profiling

```
# the head of the data
head(past_voting_data)
```

From.country	To.country	Points	Year	X.semi...final	Edition	Jury.or.Televoting
Germany	Belgium	7	1975	f	1975f	J
Sweden	Belgium	2	1975	f	1975f	J
The Netherlands	Belgium	5	1975	f	1975f	J
Germany	Finland	12	1975	f	1975f	J

From.country	To.country	Points	Year	X.semi...final	Edition	Jury.or.Televoting
Ireland	Finland	5	1975	f	1975f	J
Israel	Finland	8	1975	f	1975f	J

```
# structure of the data
str(past_voting_data)
```

```
## 'data.frame':    12292 obs. of  7 variables:
##  $ From.country      : chr  "Germany" "Sweden" "The Netherlands" "Germany" ...
##  $ To.country        : chr  "Belgium" "Belgium" "Belgium" "Finland" ...
##  $ Points            : int  7 2 5 12 5 8 10 1 3 12 ...
##  $ Year              : int  1975 1975 1975 1975 1975 1975 1975 1975 1975 1975 ...
##  $ X.semi...final    : chr  "f" "f" "f" "f" ...
##  $ Edition           : chr  "1975f" "1975f" "1975f" "1975f" ...
##  $ Jury.or.Televoting: chr  "J" "J" "J" "J" ...
```

```
# summary statistics of the data
summary(past_voting_data)
```

```
##  From.country        To.country           Points           Year
##  Length:12292       Length:12292       Min.   : 1.000   Min.   :1975
##  Class :character   Class :character   1st Qu.: 3.000   1st Qu.:1996
##  Mode  :character   Mode  :character   Median : 6.000   Median :2006
##                                        Mean   : 5.822   Mean   :2002
##                                        3rd Qu.: 8.000   3rd Qu.:2011
##                                        Max.   :12.000   Max.   :2015
##  X.semi...final       Edition          Jury.or.Televoting
##  Length:12292       Length:12292       Length:12292
##  Class :character   Class :character   Class :character
##  Mode  :character   Mode  :character   Mode  :character
##
##
##
```

```
# there is no missing data
anyNA(past_voting_data)
```

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

Construct the average vote matrix

```
# Average over the voting data
pvd <- past_voting_data %>%
       select(From.country, To.country, Points) %>%
       group_by(From.country, To.country) %>%
       summarise(Average.Points = mean(Points)) %>%
       as.data.frame()

# write the average voting data to a csv file
write.csv(x = pvd, file = "./data/average_points.csv", row.names = F)
```

Construct the Average Points Graph

```
# graph the data based on the weight of the average point score
# assumption: an average point score of 8 or more is a sign of bloc voting
G <- graph_from_data_frame(d = pvd[pvd$Average.Points >= 8, 1:2], directed = T)
E(G)$weight <- as.numeric(pvd[pvd$Average.Points >= 8, 3])
# check the graph is weighted
if (is_weighted(G) == FALSE){
  stop("Voting Network Graph is Unweighted.")
}
```

Derive the Voting Blocs

Edge Betweeness Clustering

```
# Extract the voting blocs
# 21 groups
# global modularity = 0.057
# Perform Edge between clustering
com1 <- cluster_edge_betweenness(graph = G, weights = E(G)$weight)
# calculate modularity of the graph
modularity(com1)
```

```
## [1] 0.06721339
```

```
# get the number of groups
max(com1$membership)
```

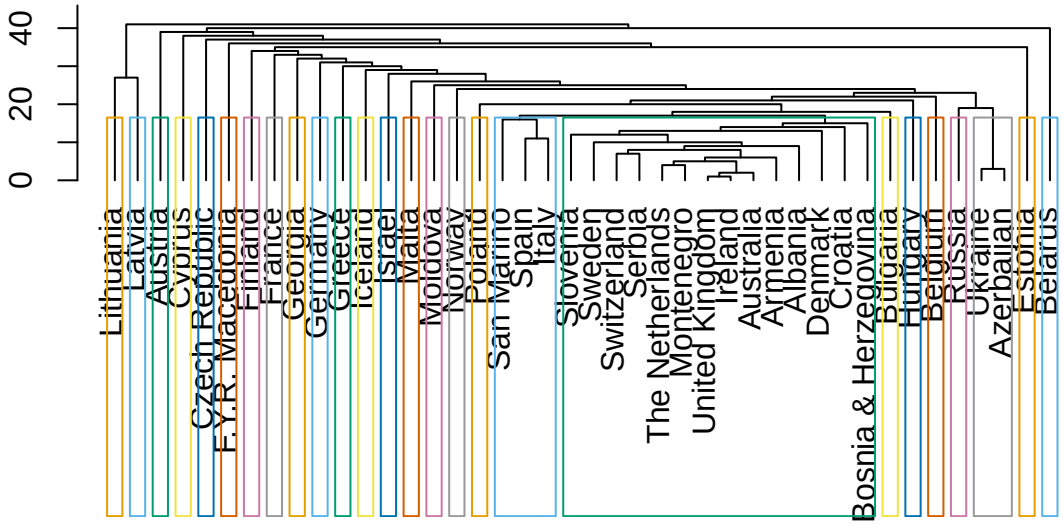
```
## [1] 26
```

```
# create a data frame out of the clustering results
com1df <- rbind(com1$names, com1$membership)
# assign the row names to the data frame
row.names(com1df) <- c("Country", "Group")
# show head of results
t(head(com1df))
```

```
##       Country                Group
##  [1,] "Albania"              "1"
##  [2,] "Armenia"              "1"
##  [3,] "Australia"            "1"
##  [4,] "Austria"              "2"
##  [5,] "Azerbaijan"           "3"
##  [6,] "Belarus"              "4"
##  [7,] "Belgium"              "5"
##  [8,] "Bosnia & Herzegovina" "1"
##  [9,] "Bulgaria"             "6"
## [10,] "Croatia"              "1"
## [11,] "Cyprus"               "7"
## [12,] "Czech Republic"       "8"
```

```
## [13,] "Denmark"              "1"
## [14,] "Estonia"              "9"
## [15,] "F.Y.R. Macedonia"     "10"
## [16,] "Finland"              "11"
## [17,] "France"               "12"
## [18,] "Georgia"              "13"
## [19,] "Germany"              "14"
## [20,] "Greece"               "15"
## [21,] "Hungary"              "16"
## [22,] "Iceland"              "17"
## [23,] "Ireland"              "1"
## [24,] "Israel"               "18"
## [25,] "Italy"                "19"
## [26,] "Latvia"               "20"
## [27,] "Lithuania"            "21"
## [28,] "Malta"                "22"
## [29,] "Moldova"              "23"
## [30,] "Montenegro"           "1"
## [31,] "Norway"               "24"
## [32,] "Poland"               "25"
## [33,] "Russia"               "26"
## [34,] "San Marino"           "19"
## [35,] "Serbia"               "1"
## [36,] "Slovenia"             "1"
## [37,] "Spain"                "19"
## [38,] "Sweden"               "1"
## [39,] "Switzerland"          "1"
## [40,] "The Netherlands"      "1"
## [41,] "Ukraine"              "3"
## [42,] "United Kingdom"       "1"
```

```
# construct a dendrogram (hierarchical clustering method)
plot_dendrogram(com1, main = "Dendrogram of EDge Betweeness Clustering")
```



Short Random Walks

```
# Extract voting blocs
# 6 groups
# global modularity = 0.3
# perform short random walks clustering
com2 <- cluster_walktrap(graph = G, weights = E(G)$weight)
# calculate modularity of the graph
modularity(com2)
```

```
## [1] 0.3910908
```

```
# get the number of groups
max(com2$membership)
```

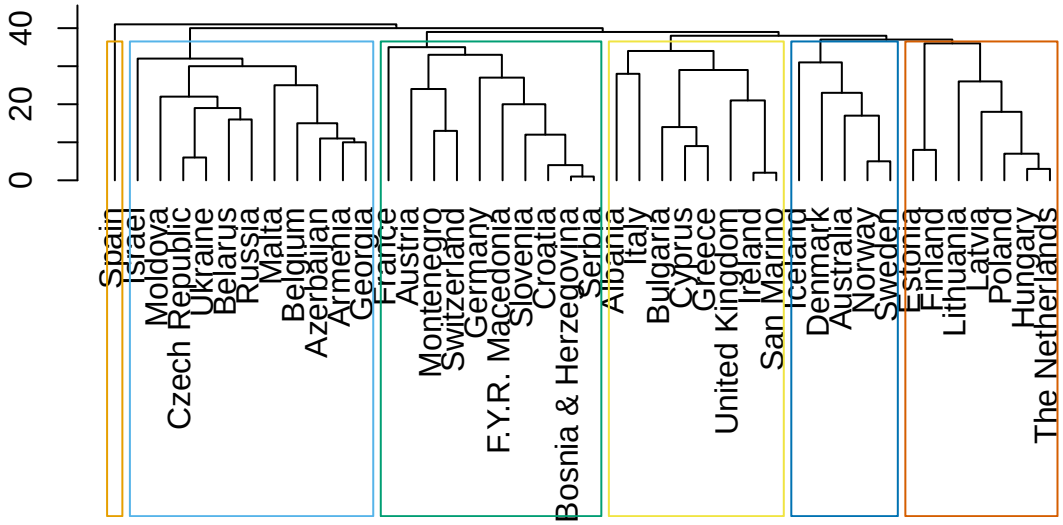
```
## [1] 6
```

```
# create a data frame out of the clustering results
com2df <- rbind(com2$names, com2$membership)
# assign the row names to the data frame
row.names(com2df) <- c("Country", "Group")
# show head of results
t(head(com2df))
```

```
##      Country                Group
```

```
##  [1,] "Albania"              "3"
##  [2,] "Armenia"              "4"
##  [3,] "Australia"            "5"
##  [4,] "Austria"              "2"
##  [5,] "Azerbaijan"           "4"
##  [6,] "Belarus"              "4"
##  [7,] "Belgium"              "4"
##  [8,] "Bosnia & Herzegovina" "2"
##  [9,] "Bulgaria"             "3"
## [10,] "Croatia"              "2"
## [11,] "Cyprus"               "3"
## [12,] "Czech Republic"       "4"
## [13,] "Denmark"              "5"
## [14,] "Estonia"              "1"
## [15,] "F.Y.R. Macedonia"     "2"
## [16,] "Finland"              "1"
## [17,] "France"               "2"
## [18,] "Georgia"              "4"
## [19,] "Germany"              "2"
## [20,] "Greece"               "3"
## [21,] "Hungary"              "1"
## [22,] "Iceland"              "5"
## [23,] "Ireland"              "3"
## [24,] "Israel"               "4"
## [25,] "Italy"                "3"
## [26,] "Latvia"               "1"
## [27,] "Lithuania"            "1"
## [28,] "Malta"                "4"
## [29,] "Moldova"              "4"
## [30,] "Montenegro"           "2"
## [31,] "Norway"               "5"
## [32,] "Poland"               "1"
## [33,] "Russia"               "4"
## [34,] "San Marino"           "3"
## [35,] "Serbia"               "2"
## [36,] "Slovenia"             "2"
## [37,] "Spain"                "6"
## [38,] "Sweden"               "5"
## [39,] "Switzerland"          "2"
## [40,] "The Netherlands"      "1"
## [41,] "Ukraine"              "4"
## [42,] "United Kingdom"       "3"
```

```
# construct a dendrogram (hierarchical clustering method)
plot_dendrogram(com2, main = "Dendrogram of Short Random Walk Clustering")
```



Construct voting blocs data frame

```
# UPDATE: for the propose of this research we shall only include hierarchical clustering methods
# (1) Edge-betweenness
# (2) Short Random Walks
# combine results over both clustering methods into a data frame
voting_bloc_data <- as.data.frame(cbind(com1$names, com1$membership, com2$membership))
# assign column names to the output data frame
colnames(voting_bloc_data) <- c("Country", "VBlocs1_EB", "VBlocs2_SRW")
# show head of results
head(voting_bloc_data)
```

Country	VBlocs1_EB	VBlocs2_SRW
Albania	1	3
Armenia	1	4
Australia	1	5
Austria	2	2
Azerbaijan	3	4
Belarus	4	4

```r
# generate summary of data
summary(voting_bloc_data)
```

```
##    Country           VBlocs1_EB        VBlocs2_SRW
##  Length:42          Length:42          Length:42
##  Class :character   Class :character   Class :character
##  Mode  :character   Mode  :character   Mode  :character
```

```r
# Writing Voting Bloc Data to csv file
write.csv(x = voting_bloc_data, file = "./data/voting_bloc_data.csv", row.names = F)
```