Social Networks Analysis

Suzana Iacob

13/12/2019

Social Network of the US Senate

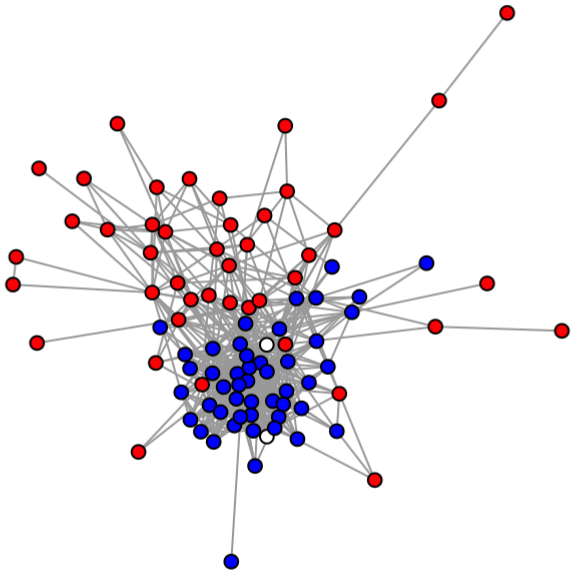
We will explore the social network of the US Senate induced by bill co-sponsorship. You will focus on the 115th Congress, which began on January 3, 2017, and concluded on January 3, 2019. The data were obtained from https://www.gpo.gov/fdsys/bulkdata/BILLSTATUS/115 (https://www.gpo.gov/fdsys/bulkdata/BILLSTATUS/115).

The data is for each senator and their party (Democrat, Republican or Independent). The edges of the network is for senators which co-sponsored at least 8 bills together.

```
senators <- read.csv("senators.csv")
senatorLinks <- read.csv("senateCosponsorship.csv", stringsAsFactors=FALSE)
G <- graph.data.frame(senatorLinks, directed=FALSE, senators)
comp = components(G)
in.max.comp = comp$membership == which.max(comp$csize)
sg = induced_subgraph(G, in.max.comp)
sg.Senators = senators[in.max.comp,]
table(sg.Senators$party)
```

```
##
##  D  I  R
## 48  2 43
```

```
set.seed(144)
color.plot <- rep('white',nrow(sg.Senators))
color.plot[sg.Senators$party=="R"] <- 'red'
color.plot[sg.Senators$party=="D"] <- 'blue'
color.plot[sg.Senators$party=="I"] <- 'white'
plot(sg, vertex.label=NA, vertex.size=5, vertex.color = color.plot)
```



The 2 parties (Republican and Democrat) are clearly differentiated and most senators co-sponsor bills alongside their party members. There is some overlap, perhaps bills that have been sponsored by both parties, and some senators who collaborate both within their party as well as with the other party. There are a few Republicans which have more links to Democrats than other Repoublicans. Democrats seem more clustered together meaning they potentially sponsored more bills with each other (this is counterintuitive since US had a Republican president in this period of time), but this could mean that the Republicans sponsored bills that are not in this network as they are fewer than 8 with the same person.

There are 2 Independent senators who seem to be more close to the Democrats but not necessarily close to each other. One of the Independents co-sponsored bills with both Republicans and Democrats, but it is difficult to intrpret a party of just 2 members.

There are also a few outliers (senators who only sponsored few bills with another senator - one edge represents 8 bills). We also know there are senators who do not appear in the graph as they do not have at least 8 bills in common with any other senator, meaning some members of the senate collaborate a lot less or simply do not sponsor bills.

Degree centrality

```
sg.degree <- degree(sg)
sg.Senators[order(sg.degree,decreasing=TRUE)[1:10],c("name", "state", "party")]
```

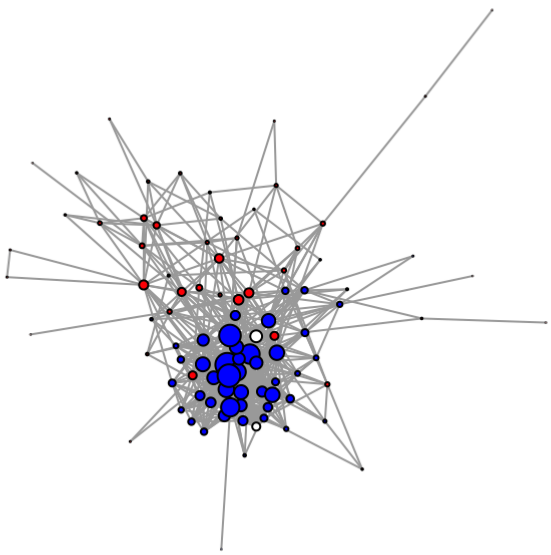
	name <fctr>	**state** <fctr>	**party** <fctr>
5	Richard Blumenthal	CT	D
101	Elizabeth Warren	MA	D

	name <fctr>	state <fctr>	party <fctr>
56	Amy Klobuchar	MN	D
2	Tammy Baldwin	WI	D
37	Kirsten Gillibrand	NY	D
9	Sherrod Brown	OH	D
62	Edward Markey	MA	D
67	Jeff Merkley	OR	D
104	Ron Wyden	OR	D
19	Christopher Coons	DE	D

1-10 of 10 rows

```
sg.Senators$degree = degree(sg)
```

```
set.seed(144)
plot(sg, vertex.label=NA, vertex.size=0.2*sg.degree, vertex.color = color.plot)
```



Degree Centrality refers to the number of connections. We see that Democrats have a much higher degree centrality, all top 10 being democrats. In fact the senator ranking 22nd is the first Republican in the degree centrality ranking and the senator ranking 18 is Independent (the larger white dot). This means democrats co-sponsor many more bills among each other.

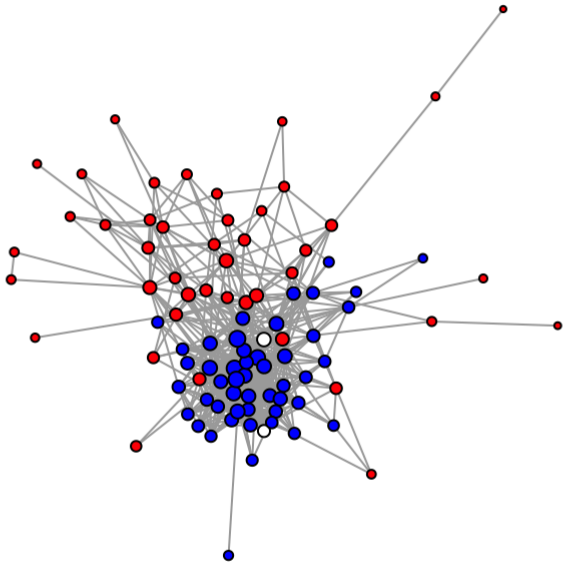
Closeness centrality

```
nodes.in.subgraph <- components(sg)$membership == which.max(components(sg)$csize)
subgraph <- induced_subgraph(sg, nodes.in.subgraph)
cl <- rep(0, nrow(sg.Senators))
cl[nodes.in.subgraph] <- closeness(subgraph)
sg.Senators[order(cl,decreasing = T)[1:10],c("name", "state", "party")]
```

	name <fctr>	**state** <fctr>	**party** <fctr>
56	Amy Klobuchar	MN	D
5	Richard Blumenthal	CT	D
101	Elizabeth Warren	MA	D
2	Tammy Baldwin	WI	D
19	Christopher Coons	DE	D
9	Sherrod Brown	OH	D
62	Edward Markey	MA	D
37	Kirsten Gillibrand	NY	D
94	Jon Tester	MT	D
104	Ron Wyden	OR	D

1-10 of 10 rows

```
set.seed(144)
plot(sg, vertex.label=NA, vertex.size=900*cl, vertex.color = color.plot)
```



```
sg.Senators$closeness = cl*1000
```

High Closeness is someone who needs few introductions to reach everyone else. Visually we cannot differentiate as well, since the numbers are not that different in absolute value. This also means that overall all senators are fairly well connected, meaning the network has many edges. Again only Democrats make the top 10, the first republican is at number 13 and the first independent at 14. This means Democrats only need a few introductions to meet not just all other Democrats but also the Republicans, meaning they are connected to the key republicans, who in turn are connected to the rest of the republicans.

Betweenness centrality

```
sg.betweenness <- betweenness(sg)
sg.Senators[order(sg.betweenness,decreasing = T)[1:10],c("name", "state", "party")]
```

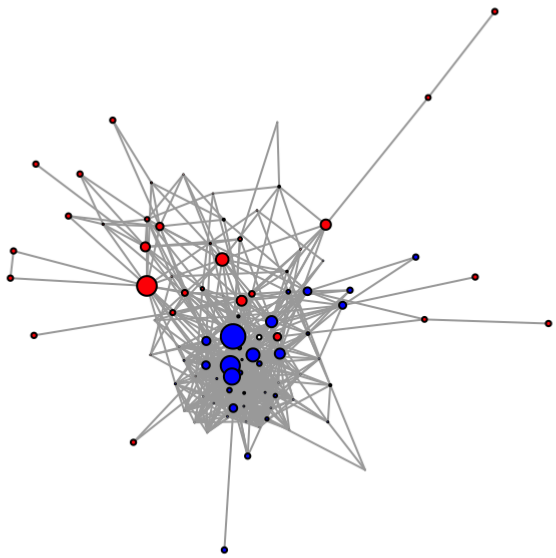
	name <fctr>	**state** <fctr>	**party** <fctr>
56	Amy Klobuchar	MN	D
81	Marco Rubio	FL	R
5	Richard Blumenthal	CT	D
101	Elizabeth Warren	MA	D
2	Tammy Baldwin	WI	D
42	Orrin Hatch	UT	R

	name <fctr>	state <fctr>	party <fctr>
94	Jon Tester	MT	D
24	Michael Crapo	ID	R
104	Ron Wyden	OR	D
36	Cory Gardner	CO	R

1-10 of 10 rows

```
sg.Senators$betweenness = betweenness(sg)*1000
```

```
set.seed(144)
plot(sg, vertex.label=NA, vertex.size=0.02*sg.betweenness,vertex.color = color.plot)
```



Betweenness measures how a person bridges otherwise disconnected groups in a network. Here we see a much better mix of Democrats and Republicans, which makes sense. The two groups are disconnected in general and connected by a few key individuals. The first independent is at position 25 which is somewhat surprising (as a member of neither party they could have a bigger impact by being the link between otherwise disconnected groups)

Page Rank

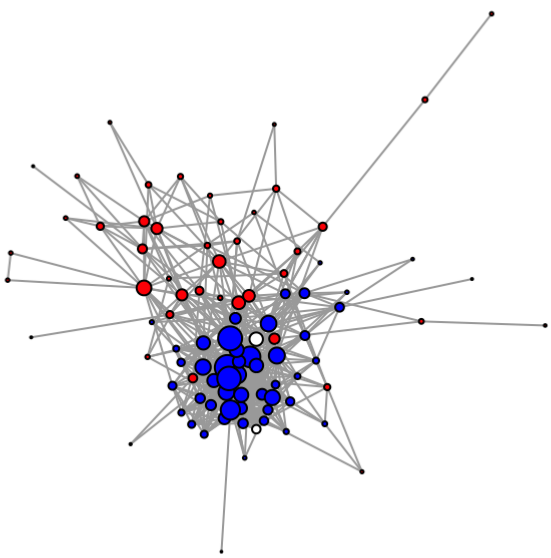
```
page.rank.score = page.rank(sg)$vector
sg.Senators[order(page.rank.score,decreasing = T)[1:10],c("name", "state", "party")]
```

	name <fctr>	state <fctr>	party <fctr>
5	Richard Blumenthal	CT	D
56	Amy Klobuchar	MN	D
101	Elizabeth Warren	MA	D
2	Tammy Baldwin	WI	D
37	Kirsten Gillibrand	NY	D
9	Sherrod Brown	OH	D
104	Ron Wyden	OR	D
62	Edward Markey	MA	D
94	Jon Tester	MT	D
19	Christopher Coons	DE	D

1-10 of 10 rows

```
sg.Senators$pagerank = page.rank.score*100
```

```
set.seed(144)
plot(sg, vertex.label=NA, vertex.size=300*page.rank.score,vertex.color = color.plot)
```



Page rank marks the importance or popularity of senators. Again top 10 are Democrats and the first Republican is at 12, the first Independent at 16. The unconnected senators have low page rank as expected.

Page ra

	id <int>	name <fctr>	state <fctr>	party <fctr>	degree <int>	closeness <int>	betweenness <int>	pagerank <int>
1	5	Richard Blumenthal	CT	D	1	2	3	1
2	101	Elizabeth Warren	MA	D	2	3	4	3
3	56	Amy Klobuchar	MN	D	3	1	1	2
4	2	Tammy Baldwin	WI	D	4	4	5	4
5	37	Kirsten Gillibrand	NY	D	5	8	14	5
6	9	Sherrod Brown	OH	D	6	6	28	6
7	62	Edward Markey	MA	D	7	7	26	8
8	104	Ron Wyden	OR	D	8	9	9	7
9	67	Jeff Merkley	OR	D	9	18	31	11
10	19	Christopher Coons	DE	D	10	5	15	10

1-10 of 10 rows

We here display the top 10 senators by degree, with their ranks for closeness, betweeness and page rank. We see that degree and page rank are very similar. And closeness is quite similar with a few exceptions. Betweenness is a quite different measure and points out different senators.

Overall some key people in the senate are Amy Klobuchar and Richard Blumenthal who rank 1,2 or 3 for all measures, meaning they are very active both inside the democratic party as well as outside.

Discussion

```
attach(sg_rank)
newdata <- sg_rank[order(betweenness),]
detach(sg_rank)
head(newdata[newdata$degree>30,c("name", "state", "party", "degree", "betweenness")],
10)
```

	name <fctr>	state <fctr>	party <fctr>	degree <int>	betweenness <int>
56	Michael Crapo	ID	R	56	8
51	Roger Wicker	MS	R	51	11
42	Claire Mccaskill	MO	D	42	13
32	Charles Grassley	IA	R	32	16
44	Ted Cruz	TX	R	44	17
49	Heidi Heitkamp	ND	D	49	18
31	John Cornyn	TX	R	31	19
54	Thomas Tillis	NC	R	54	21

	name <fctr>	state <fctr>	party <fctr>	degree <int>	betweenness <int>
77	Joni Ernst	IA	R	77	22
87	Jeff Flake	AZ	R	87	23

1-10 of 10 rows

We obtained the results by first making new columns for the rank for reach measure. Then by ordering in order of betweenness rank and taking the senators that have degree rank 30 or above (quite low).

High betweenness and low degree means these senators connect with few people in general but are key links, enabling collaboration between the two parties. They are mostly republicans, meaning they are the group that connects the rest of the republicans to the important senate bills. The states are Texas, North Carolina, Arizona, mostly states in the south. These senators are critical links to the senate functions.

Community detection

```
set.seed(144)
spinglass = cluster_spinglass(subgraph, spins = 100)
community = spinglass$membership
table(community)
```

```
## community
##  1  2  3  4  5  6
## 32 18 24  2 16  1
```

We obtain 6 communities.

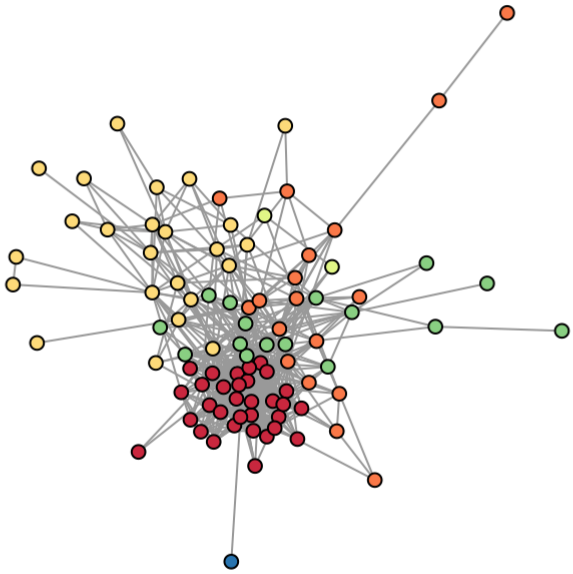
```
table(sg.Senators$party[nodes.in.subgraph], community)
```

```
##    community
##       1  2  3  4  5  6
##   D 29  7  1  1  9  1
##   I  1  0  0  0  1  0
##   R  2 11 23  1  6  0
```

```
modularity <- spinglass$modularity
modularity
```

```
## [1] 0.2823829
```

```
color.communities <- brewer.pal(max(community),"Spectral")
set.seed(144)
plot(sg, vertex.label=NA, vertex.size=5, vertex.color=color.communities[community])
```



Discussion

We see that the communities are clearly differentiated by party. The first one has mostly democrats, the next two ones have mostly republicans, and the smaller ones are a mix. Interestingly there is a community of 1 sentor.

This can explain why the republicans are less active - they are not all part of the same community. The democrats are more unites and can sponsor more bills.

```
table(sg.Senators$party[nodes.in.subgraph], community)
```

```
##    community
##       1  2  3  4  5  6
##   D 29  7  1  1  9  1
##   I  1  0  0  0  1  0
##   R  2 11 23  1  6  0
```

Sates in the communities are quite mixed but we see many groups of 2 senatores from the same state in the same community.

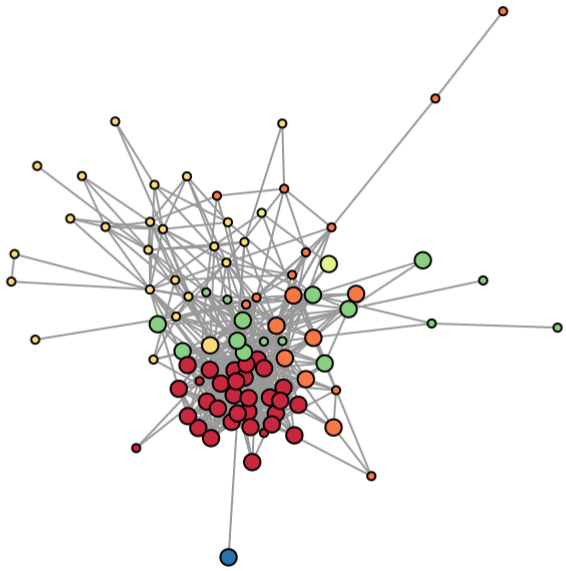
Community 1 has democrats from the noth-east, with some other states. Communities 2 and 3 have most republicans. Community 4 has 2 senators from Indiana, one democrat one republican. Community 6 is very diverse bith as parties as well as states.

```
table(sg.Senators$state[nodes.in.subgraph], community)
```

```
##     community
##       1 2 3 4 5 6
##    AK 0 2 0 0 0 0
##    AL 0 0 0 0 1 0
##    AR 0 0 2 0 0 0
##    AZ 0 2 0 0 0 0
##    CA 2 0 0 0 0 0
##    CO 0 2 0 0 0 0
##    CT 2 0 0 0 0 0
##    DE 2 0 0 0 0 0
##    FL 0 0 2 0 0 0
##    GA 0 0 1 0 0 0
##    HI 2 0 0 0 0 0
##    IA 0 0 0 0 2 0
##    ID 0 2 0 0 0 0
##    IL 2 0 0 0 0 0
##    IN 0 0 0 2 0 0
##    KS 0 0 2 0 0 0
##    KY 0 0 1 0 0 0
##    LA 0 0 2 0 0 0
##    MA 2 0 0 0 0 0
##    MD 2 0 0 0 0 0
##    ME 1 0 0 0 1 0
##    MI 2 0 0 0 0 0
##    MN 1 0 0 0 2 0
##    MO 0 0 1 0 1 0
##    MS 0 0 2 0 0 0
##    MT 0 2 0 0 0 0
##    NC 0 0 2 0 0 0
##    ND 0 0 0 0 2 0
##    NE 0 0 0 0 1 0
##    NH 1 0 0 0 1 0
##    NJ 2 0 0 0 0 0
##    NM 0 2 0 0 0 0
##    NV 0 2 0 0 0 0
##    NY 1 0 0 0 0 1
##    OH 1 0 0 0 1 0
##    OK 0 0 2 0 0 0
##    OR 1 1 0 0 0 0
##    PA 1 0 0 0 0 0
##    RI 2 0 0 0 0 0
##    SC 1 0 0 0 0 0
##    SD 0 0 2 0 0 0
##    TN 0 0 0 0 0 0
##    TX 0 0 2 0 0 0
##    UT 0 0 2 0 0 0
##    VA 0 0 0 0 2 0
##    VT 2 0 0 0 0 0
##    WA 1 1 0 0 0 0
##    WI 1 0 1 0 0 0
##    WV 0 0 0 0 2 0
##    WY 0 2 0 0 0 0
```

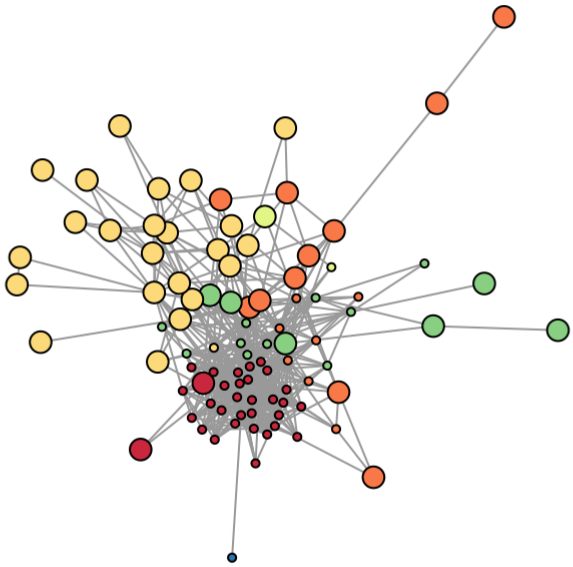
Most democrats are in communitty 1, but there are some in the other communities.

```
set.seed(144)
plot(sg, vertex.label=NA, vertex.size=ifelse(V(sg)$party == "D", 6, 3),# if the emplo
yee does trading, the node has size 10; and 3 otherwise
     vertex.color=color.communities[community])
```



We see the 2 communities and their overlap with Republicans:

```
set.seed(144)
plot(sg, vertex.label=NA, vertex.size=ifelse(V(sg)$party == "R", 8, 3),
     vertex.color=color.communities[community])
```



The idependents are part of different communities:

```
set.seed(144)
plot(sg, vertex.label=NA, vertex.size=ifelse(V(sg)$party == "I", 8, 3),
     vertex.color=color.communities[community])
```