

RESEAUX SOCIAUX ET RESEAUX SPATIAUX

Analyse des interactions élémentaires, directes et indirecte

Claude GRASLAND

5 novembre 2016

1.2. Relations sociales : l'exemple des moines de Sampson

```
library(igraph)
setwd("/Users/claudegrasland1/Documents/cg/cours/ASPS2016/NETWORK/sampson")
```

Parmi de très nombreux exemples de réseaux individuels, nous avons choisi de présenter un des plus intéressant et dramatique à la fois: l'enquête faite par un apprenti sociologue dans un monastère catholique avant et pendant une crise qui devait aboutir à la dissolution de la communauté observée.

Un “classique” de la sociologie des réseaux

Sampson recorded the social interactions among a group of monks while resident as an experimenter on vision, and collected numerous sociometric rankings. During his stay, a political “crisis in the cloister” resulted in the expulsion of four monks (Nos. 2, 3, 17, and 18) and the voluntary departure of several others - most immediately, Nos. 1, 7, 14, 15, and 16. (In the end, only 5, 6, 9, and 11 remained).

Most of the present data are retrospective, collected after the breakup occurred. They concern a period during which a new cohort entered the monastery near the end of the study but before the major conflict began. The exceptions are “liking” data gathered at three times: SAMPLK1 to SAMPLK3 - that reflect changes in group sentiment over time (SAMPLK3 was collected in the same wave as the data described below). Information about the senior monks was not included.

Four relations are coded, with separate matrices for positive and negative ties on the relation. Each member ranked only his top three choices on that tie. The relations are esteem (SAMPES) and disesteem (SAMPDES), liking (SAMPLK) and disliking (SAMPDLK), positive influence (SAMPIN) and negative influence (SAMPNIN), praise (SAMPPR) and blame (SAMPNPR). In all rankings 3 indicates the highest or first choice and 1 the last choice. (Some subjects offered tied ranks for their top four choices).

References : + Breiger R., Boorman S. and Arabie P. (1975). An algorithm for clustering relational data with applications to social network analysis and comparison with multidimensional scaling. *Journal of Mathematical Psychology*, 12, 328-383. + Sampson, S. (1969). *Crisis in a cloister*. Unpublished doctoral dissertation, Cornell University.

Chargement et visualisation des données

En tant que “classique” de la sociologie des réseaux, les données recueillies par Sampson ont souvent utilisées pour tester de nouveaux concepts, de nouvelles procédures d'analyse. On s'en est beaucoup servi notamment pour tester les algorithmes de recherche de communautés à l'intérieur d'un graphe compte tenu de l'histoire dramatique du monastère (expulsion puis scission de la communauté). Mais aussi compte-tenu du fait qu'on dispose d'une classification empirique réalisée par Sampson à partir de ces observations qualitatives, et qui peut être confrontée aux mesures quantitatives déduites de l'analyse des réseaux. Quelle que soit le graphe de relation analysé, on pourra la confronter aux quatre groupes définis empiriquement par Sampson et pour lesquels on va fixer des couleurs

```
setwd("/Users/claudegrasland1/Documents/cg/cours/ASPS2016/NETWORK/sampson")
code<-read.table("sampcode.csv",sep=";", header=TRUE, stringsAsFactors = F)
code$color="red"
code$color[code$statut2=="OL"]<-"blue"
code$color[code$statut2=="IN"]<-"green"
code$color[code$statut2=="EX"]<-"yellow"
code
```

| ## | code | nom | statut | statut2 | color |
|-------|------|-------------|----------------------|---------|--------|
| ## 1 | 1 | John | JEUNES TURCS | JT | red |
| ## 2 | 2 | Greg | JEUNES TURCS | JT | red |
| ## 3 | 3 | Basil | EXCLUS & IMMATURES | EX | yellow |
| ## 4 | 4 | Peter | OPPOSITION LOYALISTE | OL | blue |
| ## 5 | 5 | Bonaventure | OPPOSITION LOYALISTE | OL | blue |
| ## 6 | 6 | Berth | OPPOSITION LOYALISTE | OL | blue |
| ## 7 | 7 | Mark | JEUNES TURCS | JT | red |
| ## 8 | 8 | Victor | INTERMEDIAIRE | IN | green |
| ## 9 | 9 | Ambrose | OPPOSITION LOYALISTE | OL | blue |
| ## 10 | 10 | Romuald | INTERMEDIAIRE | IN | green |
| ## 11 | 11 | Louis | OPPOSITION LOYALISTE | OL | blue |
| ## 12 | 12 | Winifred | JEUNES TURCS | JT | red |
| ## 13 | 13 | Amand | INTERMEDIAIRE | IN | green |
| ## 14 | 14 | Hugh | JEUNES TURCS | JT | red |
| ## 15 | 15 | Boniface | JEUNES TURCS | JT | red |
| ## 16 | 16 | Albert | JEUNES TURCS | JT | red |
| ## 17 | 17 | Elias | EXCLUS & IMMATURES | EX | yellow |
| ## 18 | 18 | Simplicius | EXCLUS & IMMATURES | EX | yellow |

Choix de la relation : “X a de l’estime pour Y” (sampes)

Cet exemple permet de montrer une deuxième méthode de chargement des données relatives à un graphe, non plus sous forme de paires d’individus mais de matrice origine-destination.

```
link<-read.table("sampes.csv",sep=";", header=TRUE, stringsAsFactors = F)
mat<-as.matrix(link[,-1])
rownames(mat)<-link[,1]
colnames(mat)<-link[,1]
mat[1:5,]
```

| ## | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 |
|------|---|---|---|---|---|---|---|---|---|----|----|----|----|----|----|----|----|----|
| ## 1 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 3 | 0 | 0 | 0 | 2 | 0 | 0 | 0 | 0 | 0 | 0 |
| ## 2 | 3 | 0 | 0 | 0 | 1 | 0 | 0 | 2 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| ## 3 | 3 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 2 | 0 | 0 | 0 | 1 | 1 |
| ## 4 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 2 | 3 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| ## 5 | 0 | 0 | 0 | 3 | 0 | 0 | 0 | 0 | 1 | 0 | 2 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |

Les lignes correspondent aux origines (individus déclarant) et les colonnes aux destinations (individus déclarés). En principe chaque moine devrait citer 3 individus par ordre d’importance décroissante en notant 3 le premier choix, 2, le deuxième choix, 1 le dernier choix. Mais en pratique la consigne n’a pas toujours été respecté, puisque certains moines ont cités plus ou moins de trois individus. La prise en compte des rangs complique par ailleurs beaucoup l’analyse et la plupart des travaux menés sur ces données préfèrent se ramener à un graphe binaire, soit en ne retenant que le premier choix, soit en donnant un poids identique à tous les individus cités. C’est cette seconde solution que nous avons adopté ici.

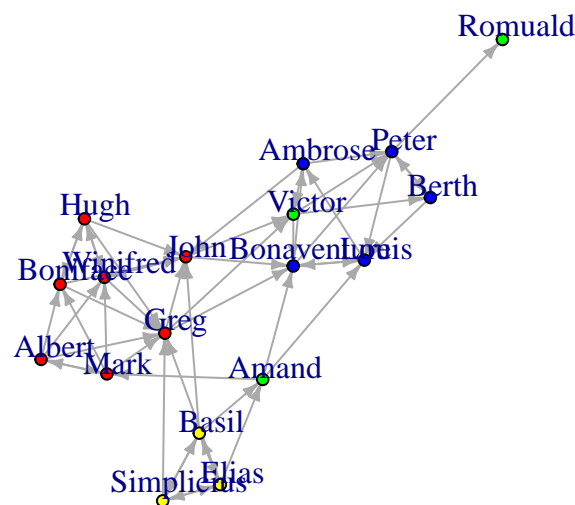
```
mat[mat>0]<-1
mat[1:5,]
```

```
##   1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18
## 1 0 0 0 0 1 0 0 1 0 0 0 1 0 0 0 0 0
## 2 1 0 0 0 1 0 0 1 0 0 0 0 0 0 0 0 0
## 3 1 0 0 0 0 0 0 0 0 0 0 1 0 0 0 1 1
## 4 0 0 0 0 0 1 0 0 0 1 1 0 0 0 0 0 0
## 5 0 0 0 1 0 0 0 0 1 0 1 0 0 0 0 0 0
```

Analyse du graphe non symétrique des relations d'estime

La matrice d'estime décrit une relation binaire (0/1) mais non symétrique que l'on peut visualiser à l'aide d'un graphe comportant des flèches de direction.

```
net<-graph_from_adjacency_matrix(mat,mode=c("directed"),diag=FALSE)
par(mfrow=c(1,1))
plot(net, vertex.color=code$color,
      vertex.label.dist=0.4,
      vertex.label.cex=1,
      edge.arrow.size=0.4,
      vertex.size=5,
      vertex.label=code$nom)
```



On ne discutera pas ici les paramètres de visualisation (pour plus de détail, voir le site <http://kateto.net/network-visualization>) et on se concentrera sur l'interprétation des relations entre individus. À première vue, on retrouve bien les regroupements décrits par Sampson avec une forte densité de lien entre les *jeunes turcs* (en rouge), *l'opposition loyaliste* (en bleu), les *exclus et immatures* (en jaune). Les *intermédiaires* (en vert) apparaissent bien en situation de pont entre les deux groupes principaux, tout au moins dans le cas de Victor et Amand. Mais Romuald apparaît surtout en position ultra-périphérique par rapport à l'ensemble de la communauté. Il est vrai qu'il n'a cité personne (!) et a été cité une seule fois.

Avant d'aller plus loin dans l'analyse, il est important de faire un point sur les **risques de surinterprétation des positions des individus dans un graphe**. Il existe en effet de nombreux algorithmes de positionnement des sommets, certains chercheurs pouvant être tentés d'en tester plusieurs jusqu'à obtenir celui "qui les arrange", c'est-à-dire qui correspond le mieux aux idées préconçues qu'ils ont du phénomène étudié. Or,

la recherche ne consiste pas à valider ses préjugés ! Pas plus qu'à produire de jolies images en accord avec nos préjugés ! L'algorithme utilisé par défaut dans la plupart des logiciels - dont igraph- est celui de Fruchterman-Rheingold (1991). Il essaye de rapprocher les sommets fortement reliés et d'éloigner ceux qui ne le sont pas, tout en maintenant autant que possible une distance identique entre chacun des sommets reliés. Il n'y a toutefois pas de solution exacte et beaucoup de variantes existent. On recommandera donc de toujours indiquer l'algorithme utilisé pour visualiser un graphe et de citer la publication où cet algorithme est décrit. Soit dans notre cas :

Fruchterman, Thomas M. J.; Reingold, Edward M. (1991), "Graph Drawing by Force-Directed Placement", *Software – Practice & Experience*, Wiley, 21 (11): 1129–1164, [doi:10.1002/spe.4380211102](https://doi.org/10.1002/spe.4380211102).

Plutôt que de s'extasier sur le fait de retrouver les classes prévues par Sampson, il est plus intéressant de calculer des indices et de vérifier par exemple quels sont les individus ayant le plus fort indice de centralité de degré ou d'intermédiarité parmi les moines. Le graphe étant orienté, on utilisera ici le nombre de degrés entrants pour évaluer la popularité de chaque individu.

```
tabres<-code[,c(2,3)]
tabres$POPULARITE<-degree(net,mode="in")
tabres$INTERMEDIARITE<-betweenness(net)
tabres[order(tabres$INTERMEDIARITE,decreasing=T),]
```

| ## | nom | statut | POPULARITE | INTERMEDIARITE |
|-------|-------------|----------------------|------------|----------------|
| ## 1 | John | JEUNES TURCS | 6 | 47.8333333 |
| ## 5 | Bonaventure | OPPOSITION LOYALISTE | 5 | 39.0333333 |
| ## 2 | Greg | JEUNES TURCS | 7 | 38.7833333 |
| ## 9 | Ambrose | OPPOSITION LOYALISTE | 2 | 31.7333333 |
| ## 12 | Winifred | JEUNES TURCS | 4 | 28.1166667 |
| ## 8 | Victor | INTERMEDIAIRE | 3 | 20.6666667 |
| ## 4 | Peter | OPPOSITION LOYALISTE | 4 | 20.5000000 |
| ## 11 | Louis | OPPOSITION LOYALISTE | 4 | 17.7000000 |
| ## 13 | Amand | INTERMEDIAIRE | 2 | 16.1666667 |
| ## 7 | Mark | JEUNES TURCS | 2 | 12.7666667 |
| ## 14 | Hugh | JEUNES TURCS | 2 | 9.5000000 |
| ## 3 | Basil | EXCLUS & IMMATURES | 2 | 4.9166667 |
| ## 17 | Elias | EXCLUS & IMMATURES | 2 | 2.8333333 |
| ## 15 | Boniface | JEUNES TURCS | 3 | 2.6166667 |
| ## 18 | Simplicius | EXCLUS & IMMATURES | 2 | 0.5000000 |
| ## 6 | Berth | OPPOSITION LOYALISTE | 2 | 0.3333333 |
| ## 10 | Romuald | INTERMEDIAIRE | 1 | 0.0000000 |
| ## 16 | Albert | JEUNES TURCS | 1 | 0.0000000 |

Il apparaît alors que l'individu le plus central en matière d'intermédiarité est John qui sert véritablement de plaque tournante entre les "Jeunes Turcs" et l'"Opposition Loyaliste". On notera également le cas intéressant d'Ambrose qui a une forte intermédiarité malgré sa faible centralité de degré.

Transformation en graphe symétrique par la méthode MIN ou MAX

Il est possible de transformer le graphe précédent en graphe symétrique, mais à condition de changer le sens de la relation observée et, du même coup son interprétation. Deux possibilités au moins permettent d'opérer cette transformation qui revient à remplacer la relation asymétrique $A(i,j)$ par une nouvelle relation symétrique $S(i,j)$

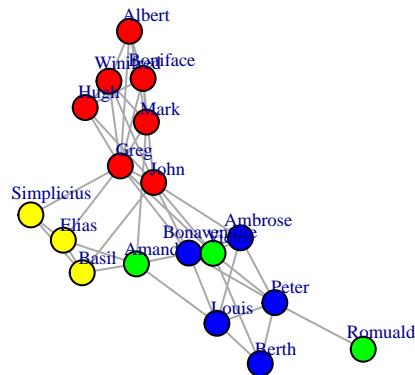
- méthode MIN : il n'y a une relation entre i et j que si cette relation est réciproque. $S(i,j) = \min(A(i,j), A(j,i))$

- méthode MAX : il y a une relation entre i et j si cette relation existe dans au moins une des deux directions $S(i,j) = \text{MAX}(A(i,j), A(j,i))$

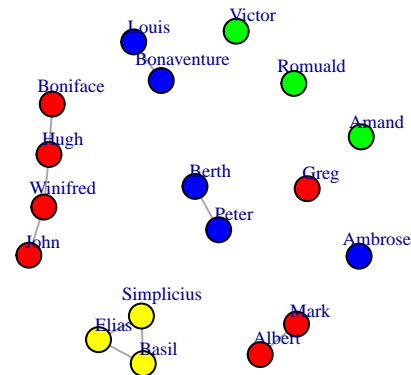
```
par(mfrow=c(1,2))
# create maximum symmetric network
netmax<-graph_from_adjacency_matrix(mat,mode=c("max"),diag=FALSE)
plot(netmax, vertex.color=code$color,
     vertex.label.dist=0.7,
     vertex.label.cex=0.6,
     vertex.label=code$nom,
     main="Réseau maximal")

# create minimum symmetric network
netmin<-graph_from_adjacency_matrix(mat,mode=c("min"),diag=FALSE)
plot(netmin, vertex.color=code$color,
     vertex.label.dist=0.7,
     vertex.label.cex=0.6,
     vertex.label=code$nom,
     main="Réseau minimal")
```

Réseau maximal



Réseau minimal



Si le réseau maximal n'apporte pas d'informations très nouvelles (il y a tout de même quelques changements dans les indices de centralité), le réseau minimal met en valeur la relative rareté des relations d'estimes réciproques. Et le fait que ces relations d'estime réciproque se retrouvent exclusivement à l'intérieur des groupes observés empiriquement par Sampson, avec notamment la triade formée par le groupe des exclus et immatures.

```
library(knitr)
```

```
## Warning: package 'knitr' was built under R version 3.2.5
```

```
list.files()
```

```
## [1] "network_sampson_files"      "network_sampson.pdf"
## [3] "network_sampson.R"         "network_sampson.Rmd"
## [5] "pgm_sampson_loadnetwork.R" "sampcode.csv"
```

```
## [7] "sampdes.csv"          "sampdlk.csv"
## [9] "sampes.csv"           "sampin.csv"
## [11] "samplk1.csv"          "samplk2.csv"
## [13] "samplk3.csv"          "sampnin.csv"
## [15] "sampnpr.csv"          "samppr.csv"
## [17] "sampson.xls"
```

```
purl("network_sampson.Rmd")
```

```
##
##
## processing file: network_sampson.Rmd
##
|
|                                     | 0%
|
|....                               | 6%
|
|.....                             | 12%
|
|.....                             | 18%
|
|.....                             | 24%
|
|.....                             | 29%
|
|.....                             | 35%
|
|.....                             | 41%
|
|.....                             | 47%
|
|.....                             | 53%
|
|.....                             | 59%
|
|.....                             | 65%
|
|.....                             | 71%
|
|.....                             | 76%
|
|.....                             | 82%
|
|.....                             | 88%
|
|.....                             | 94%
|
|.....                             | 100%

## output file: network_sampson.R
## [1] "network_sampson.R"
```