Affinity groups according to individual preferences

Alain Guénoche

CNRS, Institut de Mathématiques de Marseille
alain.guenoche@univ-amu.fr

Avignon, 2016

Problem

Let $X = \{$ n individuals $\}$

- x_+ (ranked) list of those x would like to be with
- x_- (ranked) list of those x rejects out of his group

$\{2n$ lists$\}$ $\rightarrow$ Affinity values (≥ 0 and < 0), non symmetrical

$$a : X \times X \rightarrow \mathbb{Z}$$

Pb: Compute a partition of X in p classes

- satisfying (at the best) affinity preferences
- in (more or less) balanced classes

Many applications

- built teams in business companies
- spread workers in workshops or offices
- make working groups .. at school

Partition $\Rightarrow$ Equivalence Relation $\Rightarrow$ Make affinities symmetrical

- If there is a rejection iff $a(x, y) < 0$ or $a(y, x) < 0$

$$A(x, y) = \min\{a(x, y), a(y, x)\}$$

- Else, there in an attraction : $a(x, y) \geq 0$ and $a(y, x) \geq 0$

$$A(x, y) = \frac{1}{2}[a(x, y) + a(y, x)]$$

A primary classroom ($n = 25$)

- At most 3 (ordered) attractions (> 0) and rejections (< 0)
- weighted $\pm(3, 2, 1)$, indifference $= 0$

1	0	0	0	0	0	1	0	3	0	-1	0	0	0	2	0	0	0
2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
3	3	0	0	0	-1	0	-3	0	0	0	0	0	0	0	0	0	0
4	1	0	-3	0	-1	0	0	3	0	0	0	0	0	0	0	0	0
5	1	0	0	-1	0	3	0	0	0	0	0	0	0	2	0	-2	0
6	1	0	0	2	3	0	0	0	-1	0	0	0	0	0	0	0	0
7	0	0	0	0	-2	0	0	0	0	2	0	-3	0	1	0	0	0
8	0	3	0	0	0	0	0	0	-3	0	0	2	0	0	0	0	0
9	0	3	1	0	-1	0	0	2	0	0	0	-3	0	0	0	0	0
10	0	0	0	0	-1	3	0	-3	0	0	0	0	0	1	0	0	0
11	-2	0	0	0	-1	-3	0	0	0	0	0	0	0	0	2	3	0
12	0	3	0	0	0	0	-1	1	-2	0	0	0	0	0	0	-3	0
13	0	0	3	0	-1	0	-2	0	0	0	0	0	0	0	0	0	0
14	0	0	0	0	0	0	2	0	-1	3	0	0	0	0	0	0	0
15	0	0	0	0	-1	0	-2	0	0	0	3	0	0	0	0	2	0
etc ..																	

Symmetrical affinity values ($\times 2$)

1	0	0	3	1	1	2	0	3	0	-1	-4	0	0	2	0	0
2	0	0	0	0	0	0	0	3	3	0	0	3	0	0	0	0
3	3	0	0	-6	-1	0	-6	0	1	0	0	0	3	0	0	3
4	1	0	-6	0	-2	2	0	3	0	0	0	0	0	0	0	0
5	1	0	-1	-2	0	6	-4	0	-1	-1	-1	0	-1	2	-1	-4
6	2	0	0	2	6	0	0	0	-1	3	-6	0	0	0	0	0
7	0	0	-6	0	-4	0	0	0	0	2	0	-6	-4	3	-4	0
8	3	3	0	3	0	0	0	0	-6	-6	0	3	0	0	0	0
9	0	3	1	0	-1	-1	0	-6	0	0	0	-6	0	-1	0	0
10	-1	0	0	0	-1	3	2	-6	0	0	0	0	0	4	0	0
11	-4	0	0	0	-1	-6	0	0	0	0	0	0	0	0	5	3
12	0	3	0	0	0	0	-6	3	-6	0	0	0	0	0	0	-6
13	0	0	3	0	-1	0	-4	0	0	0	0	0	0	0	0	0
14	2	0	0	0	2	0	3	0	-1	4	0	0	0	0	0	0
15	0	0	0	0	-1	0	-4	0	0	0	5	0	0	0	0	4
etc ..																

Multicriteria Optimization Problem

Let $P = \{X_1, X_2, \ldots, X_p\}$ be a partition on X

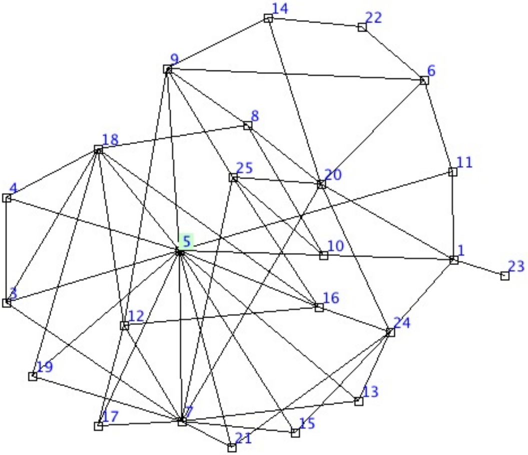
1. Rejections first
   - An undirected rejection graph
   - A coloring problem

$$R(P) = \sum_{k=1}^{p} \sum_{x,y \in X_k, A(x,y)<0} -A(x,y)$$

2. Combining rejections and attractions

$$W(P) = \sum_{k=1}^{p} \sum_{x,y \in X_k} A(x,y)$$

The rejection graph of the CE2 classroom



A clique of order 4 $\{3,4,5,18\} \Rightarrow \chi \geq 4$

Rejections first

Coloring the rejection graph

Optimal coloring NP-Hard $\Rightarrow$ Heuristic *Dsatur* (Brelaz 1979)

$$Sat(x) = \text{Nb. of colors in } \Gamma(x)$$

While vertices are not colored
- Select x with maximum Sat value then maximum degree
- Color x with the first possible color
- For any y adjacent to x
  - y is saturated for this color
  - $Dg(y) := Dg(y) - 1$
  - Update $Sat(y)$

Equitable classes

Dsatur has been extended to an optimal coloring (Brelaz, 1979)

but

Classes are not balanced

Equitable coloring problem :

$$\forall \{i, j\},\ |card(X_i) - card(X_j)| \leq 1$$

I.L.P. (C. Ribero et al., 2014), B & B (Mendez-Diaz et al., 2015)

$$\forall p, \exists G \text{ such that } \chi(G) = 2 \text{ and } \chi_{eq}(G) > p$$

Balancing classes

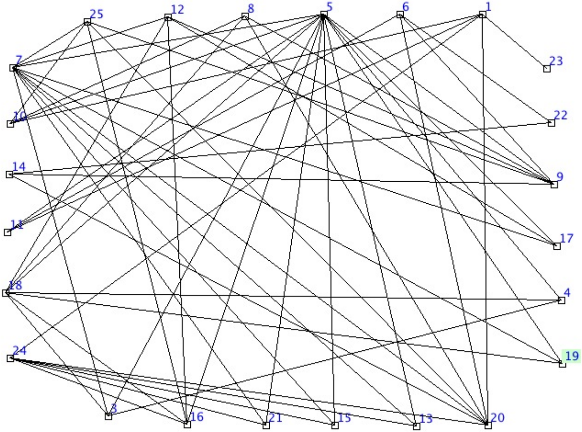
1. Balancing procedure (*optimizing W)*

   While (gainmax $>$ 0)
   For all x
   For any class without rejection of x
   Let P' be the partition after x transfer
   $Dif = W(P') - W(P)$
   $gain(x) := Max(Dif)$
   $gainmax := Max(gain(x))$
   If ($gainmax > 0$) transfer x

2. Modify *Dsatur*

   Colour x with the less used possible color

Pupil groups by Dsatur + Balancing procedure



Four groups ($R = 0$ and $W = 37.5$)

Simulations I

Two strategies

1. Applying *Dsatur* + *Balancing procedure*
2. Applying *Modified Dsatur* + *Balancing procedure*

For random affinity graphs with 100 vertices

N_r	W_1	$\%Equi_1$	4	5	6	W_2	$\%Equi_2$	4	5	6
3	93.8	75	94	6	0	74.6	83	90	10	0
4	139.7	73	0	98	2	103.9	89	1	98	1
5	177.0	69	0	60	40	141.9	62	0	49	51

N_r : Nb. of attractions and rejections per vertex
W_i : Affinity weights
$\%Equi_i$: Equitable partition rate ($cardmax - cardmin \leq 1$)
$4, 5, 6$: Percentage of Pb. giving 4, 5, 6 classes

How far is Dsatur from χ ?

Applying

1. *Strict Dsatur + Balancing procedure*
2. *Modified Dsatur + Balancing procedure*

to random graphs with 100 vertices and $\chi = 5$

N_a	$\%Equi_1$	5	6	7	8	9	$\%Equi_2$	5	6	7	8	9
200	100	100	0	0	0	0	100	100	0	0	0	0
400	83	97	3	0	0	0	98	99	1	0	0	0
600	93	0	88	12	0	0	96	0	78	22	0	0
800	73	0	5	71	24	0	84	0	2	59	38	1
1000	47	6	0	9	55	30	44	6	2	7	57	28

N_a : Nb. of edges per graph
$\%Equi$: percentage of Equitable partitions

percentage of Pb. with 5, 6, 7, 8 or 9 groups

Combining attractions and rejections

Optimizing

$$\max_{P \in \mathcal{P}} W(P) = \sum_{k=1}^{p} \sum_{(x,y) \in X_k} A(x,y)$$

A classical NP-Hard problem for graph (network) partitioning
- when vertex paires are valued by *Modularity*
- with positive and negative values

- Hierarchical ascending method (Girwan & Newmann, 2002)
- Random walk (Pons & Latapy, 2006)
- The Louvain method (Blondel et al. 2008)
- Bootstrap clustering (Gambette & Guénoche 2011)
- etc ..

Two simple algorithms

1. Hierarchical ascending method :
   - Starting from the atomic partition
   - merge two classes making maximum affinity gain
   - until the required nb. of classes
2. Transfert method (as Louvain)
   - Start from a balanced (random) partition
   - Transfer vertex making maximum affinity gain (W)
   - while there is a > 0 gain

$$+$$

Balancing procedure and *Rejection elimination procedure*

Eliminating rejections

Rejection elimination procedure

While (*gainmax* > 0)

For all x rejected in its class c

$R(x) :=$ Sum of the x rejection weights in c

For any other class q

Let $R'(x)$ the x rejection weights in q

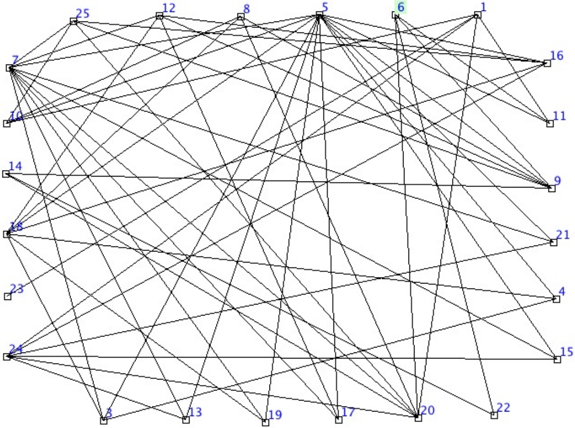
$Dif = R'(x) - R(x)$

$gain(x) := \text{Max}_q\{Dif\}$

$gainmax := \text{Max}_x\{gain(x)\}$

If (*gainmax* > 0) transfer x into class q

Other pupil groups



Four better groups ($R = 0$ et $W = 45.0$)

Simulation II

Random affinity graphs with 100 vertices
- N_r attractions and rejections
- p classes are expected

N_r	p	W_1	Gap_1	R_1	W_2	Gap_2	R_2
3	4	113.4	4.1	6.5	109.4	3.4	5.9
4	5	168.0	4.0	4.3	164.7	3.4	4.1
5	5	206.8	4.0	14.4	209.7	3.6	12.9
5	6	225.6	4.3	3.2	221.4	3.5	3.3

W : Affinity weights
Gap : Difference between the largest and the smallest classes
R : Average rejection weights (on the whole graph)

Conclusions

If rejections are mandatory (and can be satisfied)

Dsatur strict + *Balancing procedure* give

- few classes
- a good equilibrium
- but smallest affinity weights

If rejections are not mandatory

Optimizing affinity weights gives better results

The transfer method produces

- classes which are better balanced (in the average)
- less rejections