

Optimisation et défis sociétaux

Claire Mathieu



avec: 1. T. Ehrhard, S. Attias, E. Bampis, V. Cohen-Addad, B. Escoffier, F. Pascual, A. Pass-Lanneau, D. Saulpic
2. H. Gimbert, S. Mauras
3. C. Avin, B. Keller, Z. Lotker, D. Peleg, Y-A Pignolet

3 types de contributions

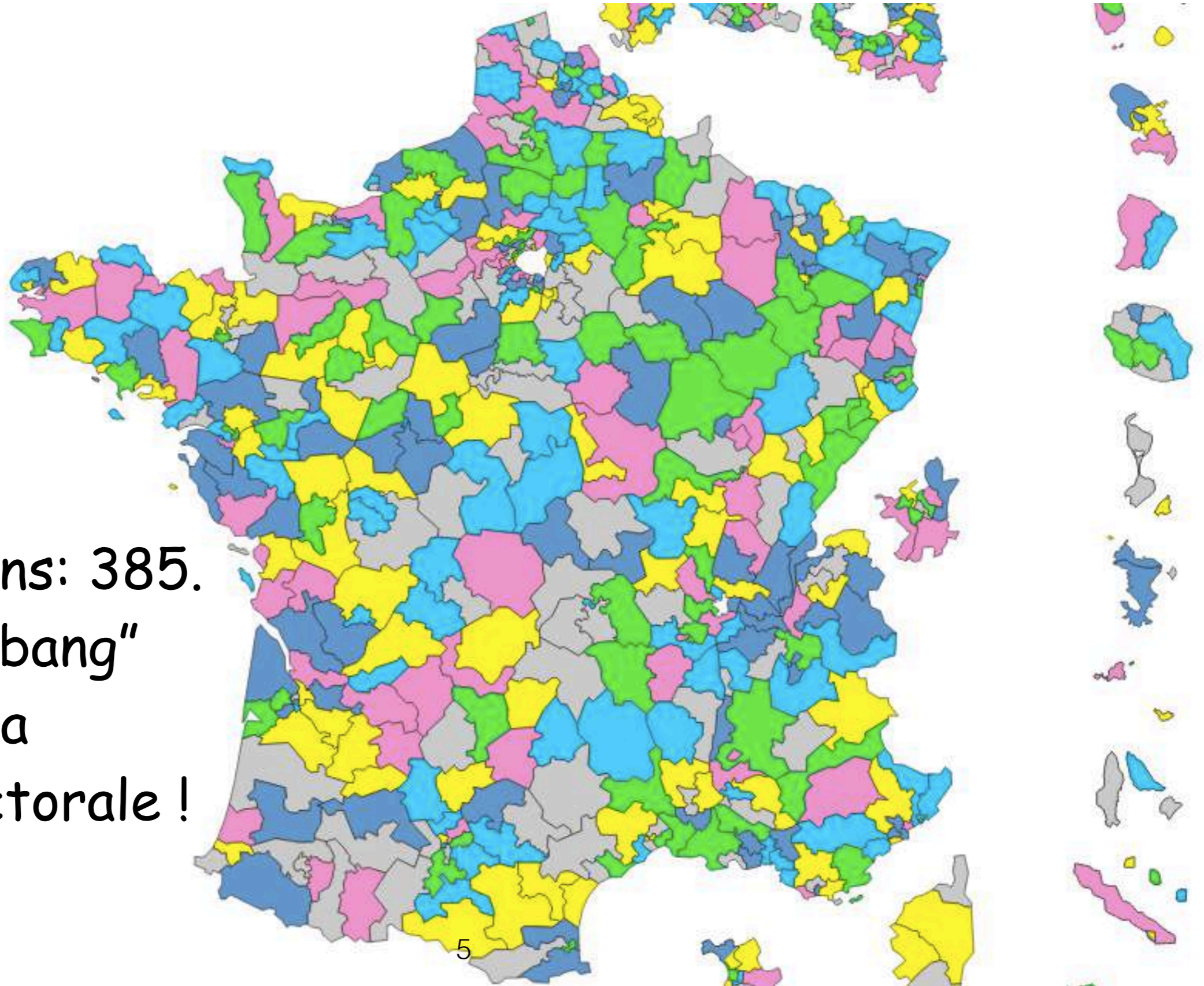
1. **conception** de solutions: découpage en circonscriptions
2. **analyse**: mariage stable pour Parcoursup
3. **modélisation**: plafond de verre dans le monde académique

1. Découpage en circonscriptions

2017



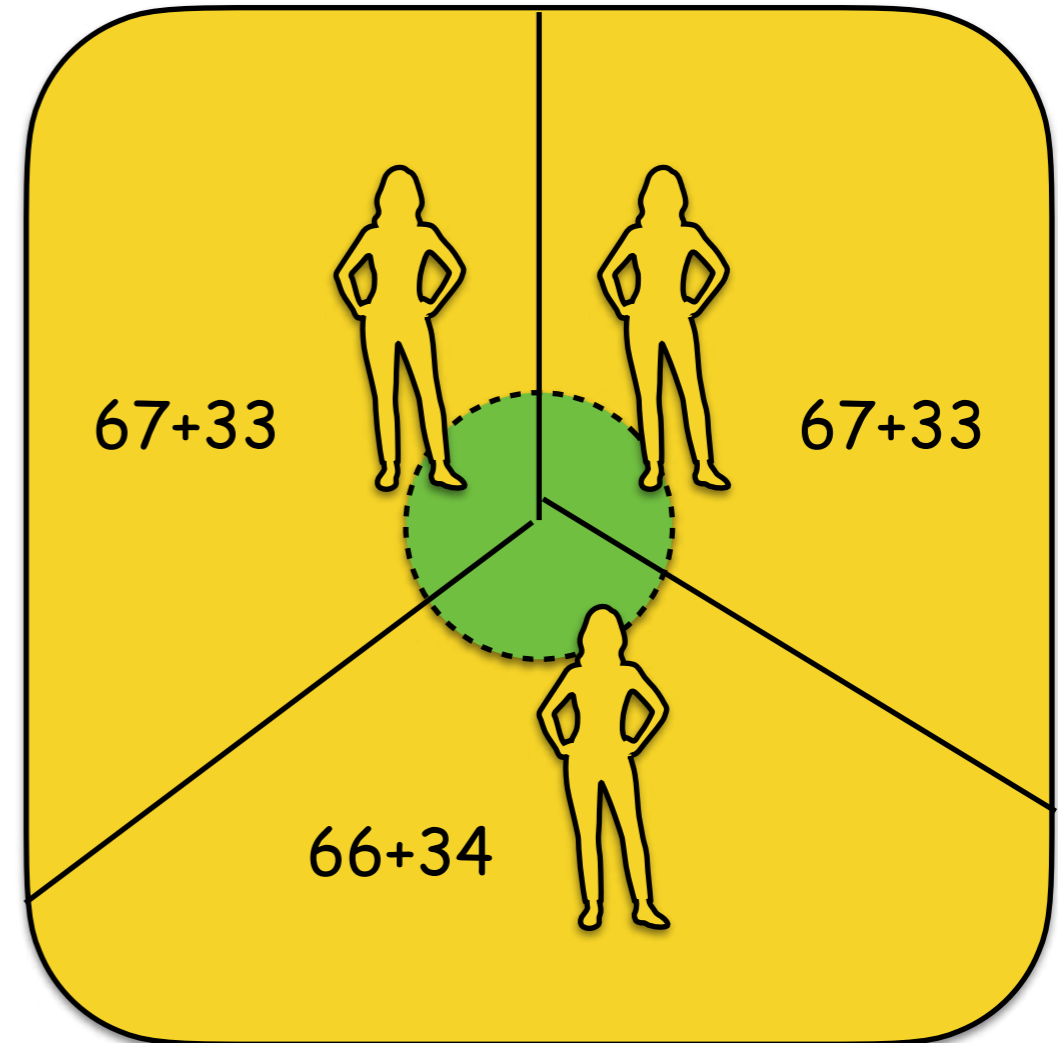
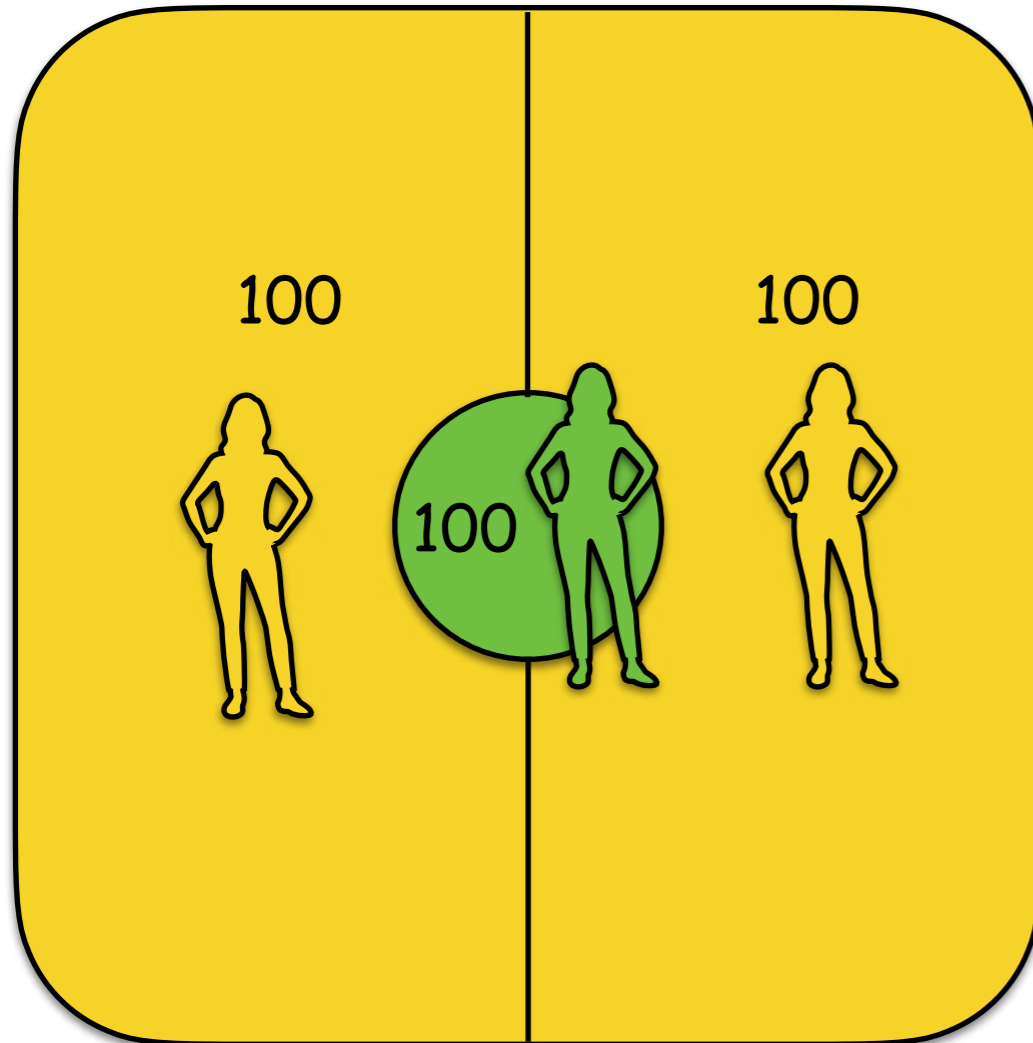
577 circonscriptions législatives



1/3 en moins: 385.
Un "big bang"
de la
carte électorale !

Un sujet sensible

charcutage électoral ("gerrymandering") : découpage ayant pour objectif de donner l'avantage à un parti donné



2021 : réforme enterrée "sans fleurs ni couronnes"

2024: on en reparle

Les règles du découpage électoral en France

MODE D'ÉLECTION

Élection au **suffrage universel direct**, pour un **mandat de 5 ans**, dans chacune des 577 circonscriptions électorales



RÉPARTITIONS DES ÉLUS



539

en France métropolitaine



27

en Outre-mer



11

pour les citoyens Français
établis à l'étranger

1. Répartition des sièges entre les départements
2. Délimitation des circonscriptions

Répartition des sièges entre les départements

1. Chaque département a au moins 1 député
2. Chaque département a un nombre entier de sièges
3. Répartition au prorata de la population

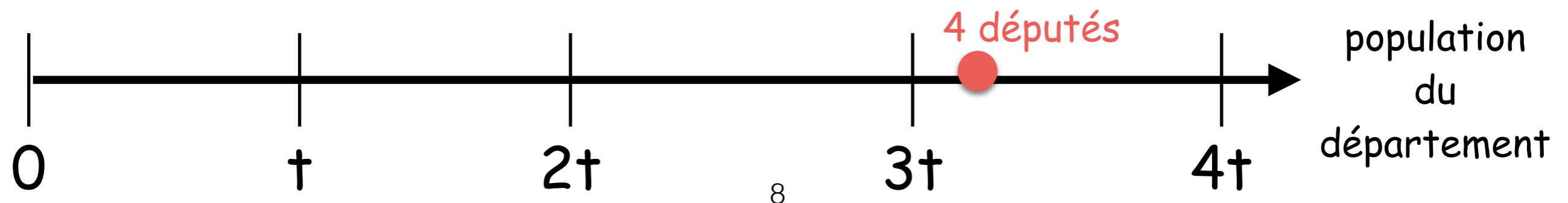
En France : répartition par tranches.

Pour chaque département

1 député jusqu'à t habitants, puis

1 député supplémentaire par tranche ou fraction de tranche de t habitants

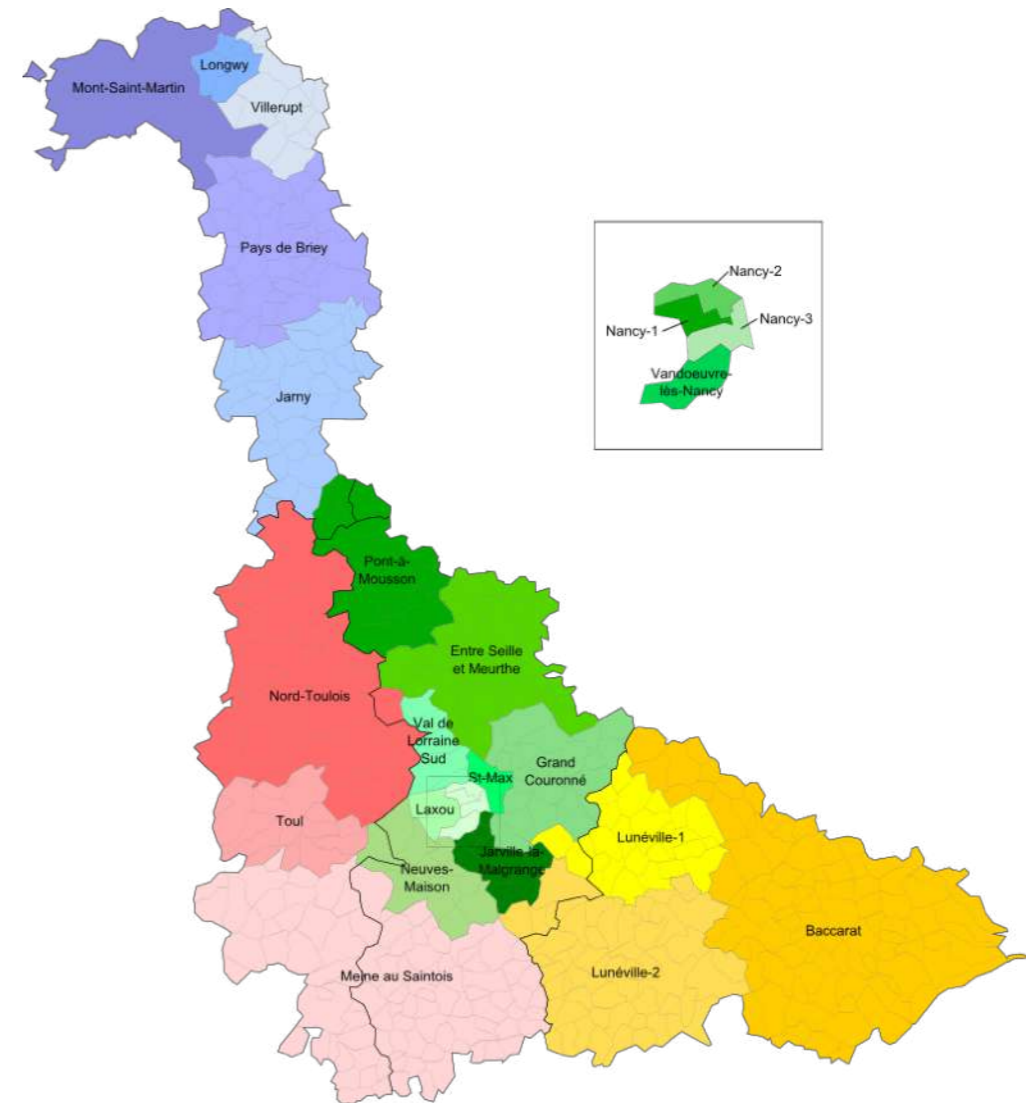
Le paramètre t est choisi de façon à obtenir 577 députés



NOMBRE DE CIRCONSCRIPTIONS PAR DÉPARTEMENT



Découpage d'un département en circonscriptions: le rôle des cantons

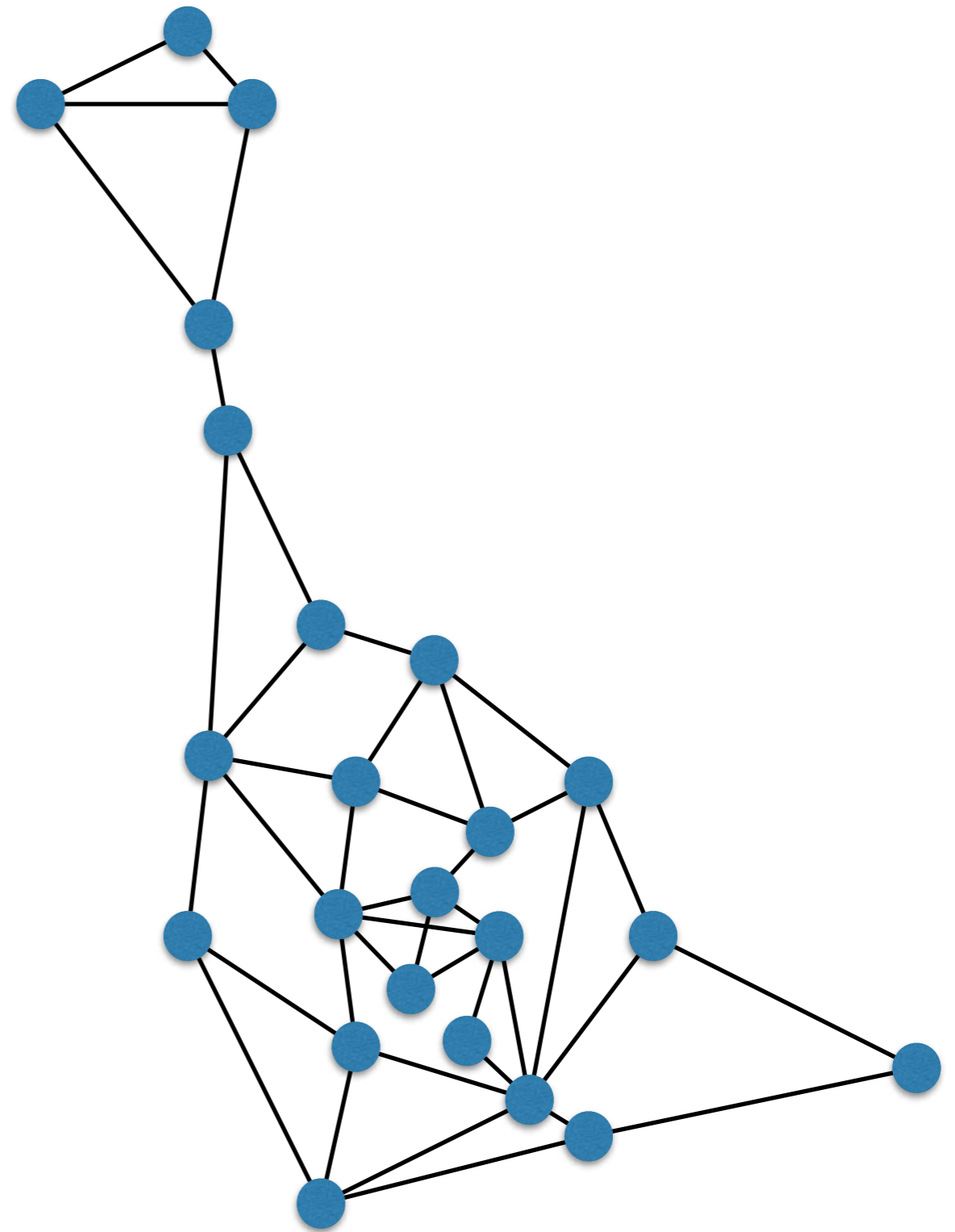
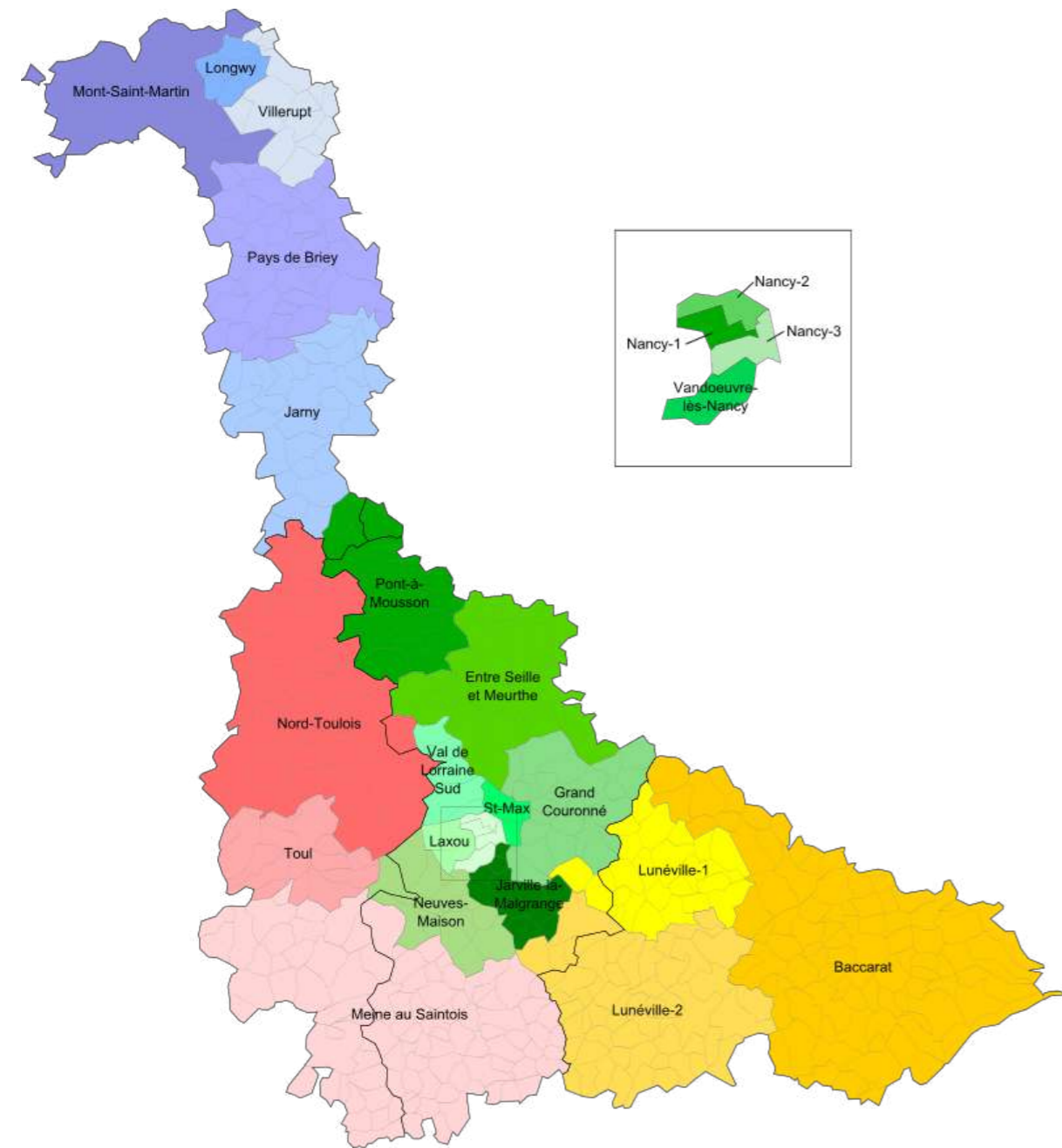


Meurthe-et-Moselle (54) :
pop. 733000, 6 circonscriptions

23 cantons

-Règle: La division d'un canton ne doit intervenir que « dans une mesure limitée »

déc. n°86-208 DC du 2 juillet 1986.



- Règle: Les circonscriptions « doivent être constituées par un territoire continu »

déc. n°86-208 DC du 2 juillet 1986 et déc. n°2008-573 DC du 8 janvier 2009

Entrée : Graphe
sommet : canton
Chaque sommet a un poids :
population du canton

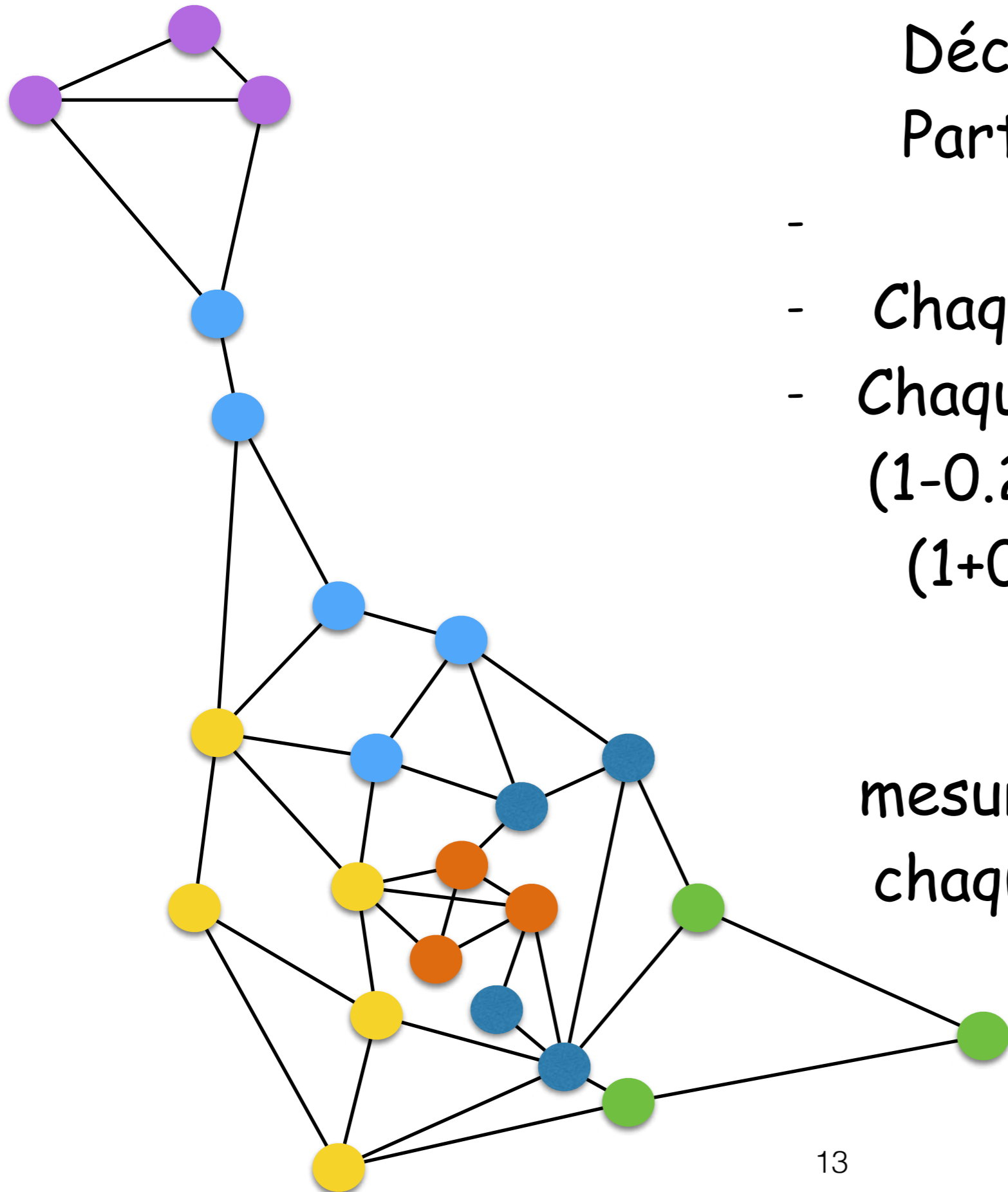
- Règle: La population d'une circonscription ne peut s'écarter de plus de 20 % de
- la population moyenne des circonscriptions au sein du département

déc. n°86-208 DC du 2 juillet 1986

Bon découpage: poids équilibrés

- Objectif: Le découpage doit se faire selon des bases « essentiellement démographiques » ... « respectant au mieux l'égalité devant le suffrage »

déc. n°86-208 DC du 2 juillet 1986 ; déc. n°86-218 DC du 18 novembre 1986 ; déc. n°2008- 573 DC du 8 janvier 2009 et déc. n°2010-602 DC du 18 février 2010



Découpage **admissible**:
 Partition des sommets

- En 6 parties
- Chaque partie est connexe
- Chaque partie a poids entre $(1-0.2) * (\text{population}/6)$ et $(1+0.2) * (\text{population}/6)$

mesure de **déséquilibre**: c tq
 chaque partie a poids entre $(1-c) * \text{moyenne}$ et $(1+c) * \text{moyenne}$

Algorithme de calcul des découpages

Pour chaque département :

- Définir le graphe correspondant et calculer la population moyenne m par circonscription, ainsi que les minimum ($0.8m$) et maximum ($1.2m$) admissibles
- Pour i allant de 1 au nombre de circonscriptions du département :
 - Pour chaque ensemble admissible de $i-1$ circonscriptions :
 - faire un **parcours en profondeur avec backtracking**, à partir du plus petit sommet non encore couvert, pour générer l'ensemble des circonscriptions i admissibles (connexité et contrainte démographique) et compatibles avec les circonscriptions $1,2,\dots, i-1$.
- Calculer le déséquilibre démographique de chaque découpage généré

Quelques cas particuliers

circonscriptions hors métropole ; départements non connexes (Charente-Maritime, Vaucluse, Vendée) ; villes de Paris, Lyon

Déséquilibre démographique minimum

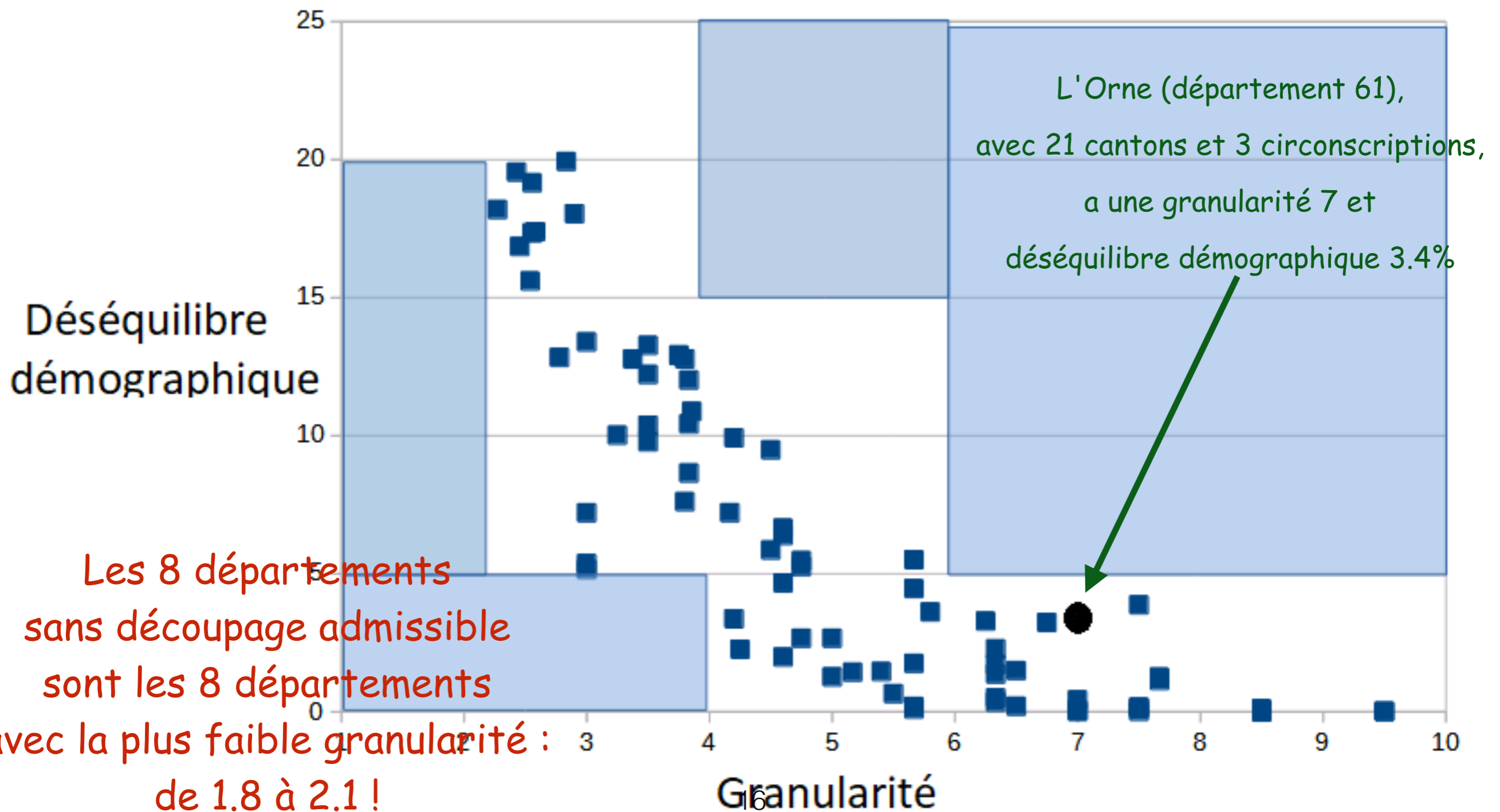
Déséquilibre démographique (%)	≤ 1%	≤ 5%	≤ 10%	≤ 15%	≤ 20%	>20%	>25%
Nombre de départements	22	44	59	72	81	8	7
% de départements	25%	49%	66%	81%	91%	9%	

- Huit départements n'ont **aucun** découpage admissible.
- Si le seuil de 20% était diminué à 15%, le nb de départements sans découpage admissible doublerait
- Si le seuil de 20% était augmenté à 25%: -1
- **20% est un bon choix !**

Comment expliquer la variation du déséquilibre minimum ?

Granularité : nombre moyen de cantons par circonscription

Observation : plus la granularité est élevée, plus le découpage est équilibré



Nombre de découpages admissibles pour chaque département

Très hétérogène !
Trop grand pour analyse manuelle : intérêt de découpage assisté par ordinateur

Maintenant que l'ensemble des découpages législatif est disponible, tout est possible !

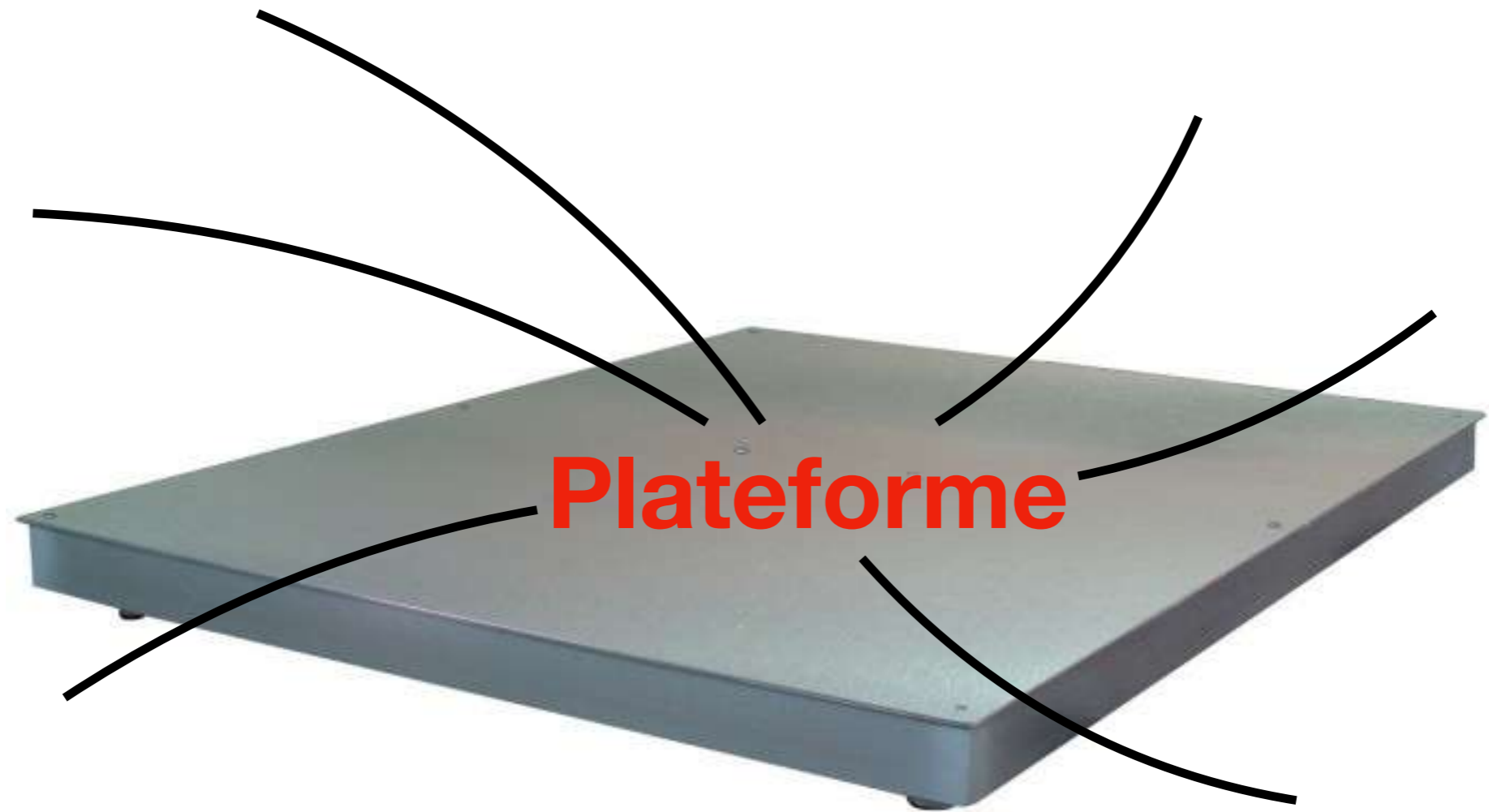
n° INSEE	NC	NDA	n° INSEE	NC	NDA	n° INSEE	NC	NDA
13	16	0	4	2	612	18	3	5 474
59	21	0	70	2	704	29	7	5 475
77	11	0	31	11	768	34	9	8 222
78	12	0	82	3	799	45	6	13 603
91	11	0	79	3	1 015	72	5	15 322
92	13	0	25	5	1 033	50	4	15 836
93	13	0	60	7	1 170	86	4	17 789
95	10	0	8	3	1 187	35	9	18 300
2A	2	2	26	4	1 394	30	6	18 540
2B	2	5	1	6	1 442	38	10	22 080
90	2	8	53	3	1 457	44	12	23 701
74	7	17	88	3	1 499	37	5	23 970
39	2	101	67	9	1 524	12	3	25 850
36	2	114	49	7	1 583	47	3	26 661
9	2	119	10	3	1 877	22	5	30 114
5	2	126	56	6	2 002	51	5	30 901
68	6	138	19	2	2 188	33	13	34 376
94	11	142	3	3	2 303	14	6	36 045
41	3	154	11	3	2 341	27	5	94 206
55	2	157	43	2	2 566	80	5	94 648
15	2	176	66	4	2 876	81	3	96 638
83	9	181	65	2	2 888	21	5	110 653
40	4	193	2	5	3 000	87	3	165 313
61	3	202	16	3	3 241	24	4	165 658
28	4	261	42	6	3 831	64	6	282 457
7	3	288	89	3	3 886	71	5	1 511 042
52	2	333	57	8	3 975	62	12	3 494 226
32	2	481	54	6	4 279	63	6	> 500 000
58	2	488	73	4	4 318	76	10	> 500 000
46	2	531	6	9	4 888			

2. Parcoursup

**Formations
en recherche d'étudiants**

**Candidats
en recherche de formations**

“Marier” formations et étudiants



Algorithme de Gale et Shapley

Entrée :

Classement des candidats par les formations
Hiérarchisation des formations par les candidats
Chaque formation a une capacité

Répéter:

1. Chaque formation envoie des propositions jusqu'à capacité (résiduelle)
2. Chaque candidat examine les propositions reçues plus celle (temporairement) acceptée précédemment et les rejette toutes sauf sa préférée, qu'il accepte **temporairement**.

Terminaison: quand il ne se passe plus rien

Les acceptations temporaires deviennent définitives

Propriétés

Temps polynomial

**Pas de *paire bloquante*:
(candidat, formation)
se préférant mutuellement
à leur assignation**

L'autre algorithme de Gale et Shapley

Entrée :
idem

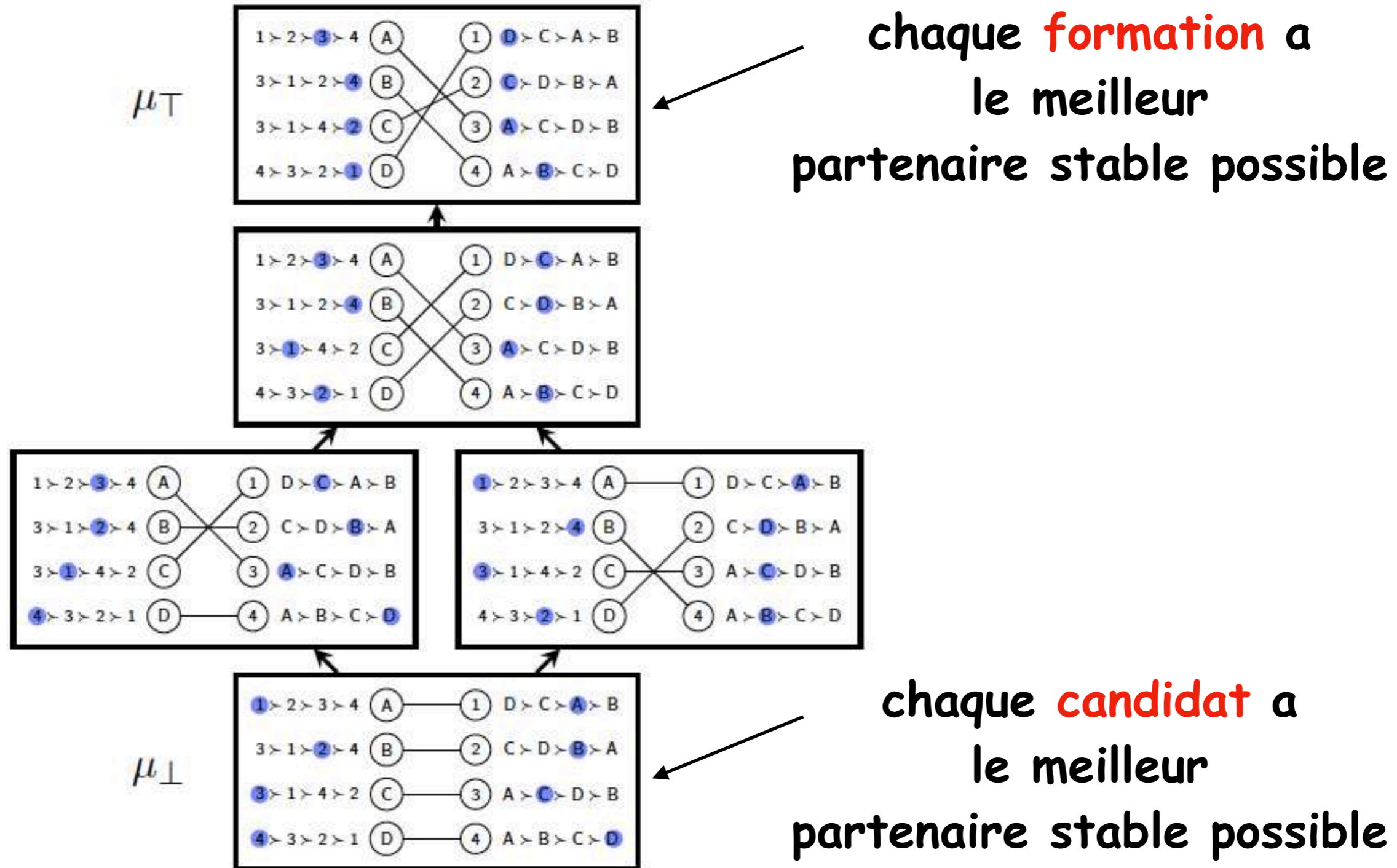
Répéter:

1. Chaque candidat envoie une candidature à la formation suivante sur sa liste
2. Chaque formation regarde les candidatures plus celles (temporairement) acceptées précédemment, rejette les moins bien classées qui dépassent sa capacité et accepte **temporairement** les autres.

Terminaison: idem

Propriétés: idem

Le treillis des mariages stables



Comparaison entre le pire et le meilleur mariages stables

Presque identique en pratique:
presque tous les étudiants (> 99.9%)
obtiennent la même formation

Empirical studies [Roth and Peranson, 1999]

National Resident Matching Program: physicians in the USA.

“The set of stable matchings, and the opportunities for strategic manipulation, are surprisingly small.”

Lemme

si tous les candidats ont la même liste de préférences,
alors le mariage stable est unique.

Lemme: si tous les candidats ont la même liste de préférences, alors le mariage stable est unique.

Pourquoi le meilleur et le pire mariages stables sont ils similaires en pratique ?

Intuition: les candidats ont des listes de préférences similaires

Roth Peranson 1999: "One factor ... is the **correlation of preferences**... When preferences are highly correlated (i.e., when similar programs tend to agree which are the most desirable applicants, and applicants tend to agree which are the most desirable programs), the set of stable matchings is small."

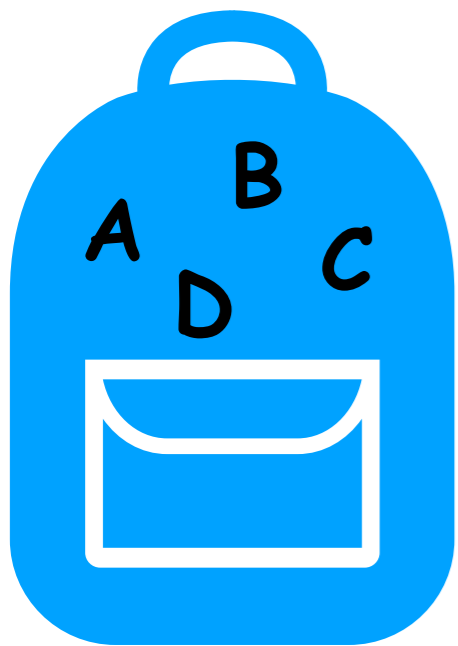
Modèle de popularité [Immorlica Mahdian 2005]

Definition. Define the popularities p_A, p_B, p_C, p_D of A, B, C, D.

$$\mathbb{P}(D \succ A, B, C) = \frac{p_D}{p_A + p_B + p_C + p_D}$$

Sample using the popularities, without replacement.

$$\mathbb{P}(D \succ C \succ A \succ B) = \frac{p_D}{p_A + p_B + p_C + p_D} \cdot \frac{p_C}{p_A + p_B + p_C} \cdot \frac{p_A}{p_A + p_B}$$



Exemple

Si les popularités des formations sont
 $A:1, B:1/2, C:1/4, D:1/8, \dots$
alors toutes les hiérarchisations seront
similaires à A B C D ...

Théorème: Si les hiérarchisations des formations par les candidats suivent une loi géométrique, alors en moyenne tous les partenaires stables d'un candidat ont le même rang, à $O(1)$ près, dans sa liste de préférence.

[Gimbert M. Murras]

Remarque: les classements des candidats par les universités sont arbitraires !

3. Plafond de verre

What is the glass ceiling effect?

Federal Glass Ceiling Commission, US Gov, 1994:

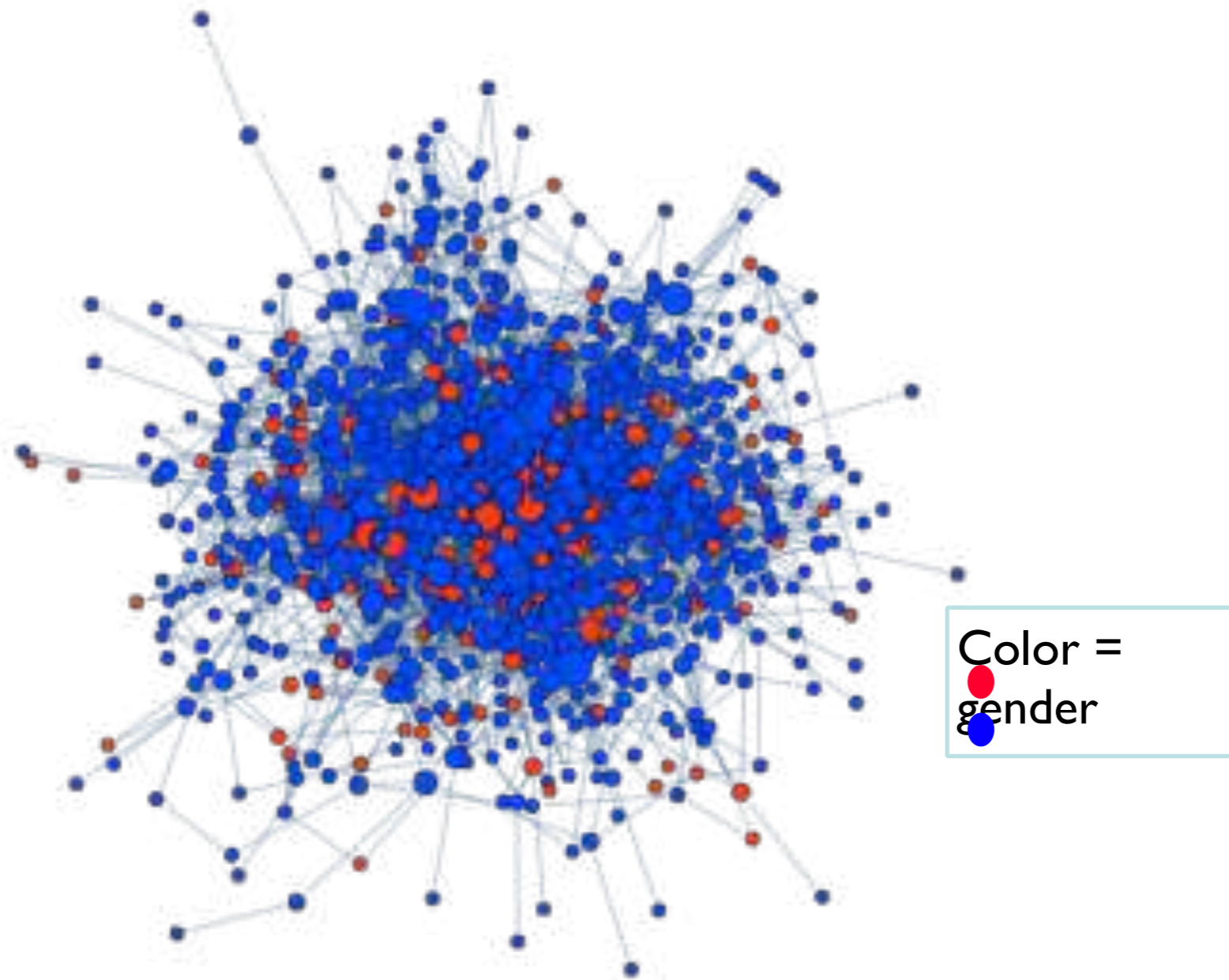
“the unseen, yet unbreakable barrier that keeps minorities and women from rising to the upper rungs of the corporate ladder, regardless of their qualifications or achievements”

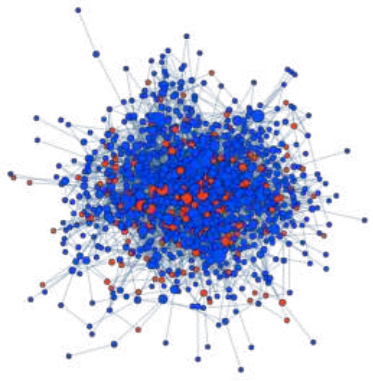
Merriam-Webster: *“an unfair system or attitudes that prevents some people (such as women or a certain race) from getting the most powerful jobs”*

What is our main goal?

Model glass ceiling in a social network, analyze it, compare with empirical data

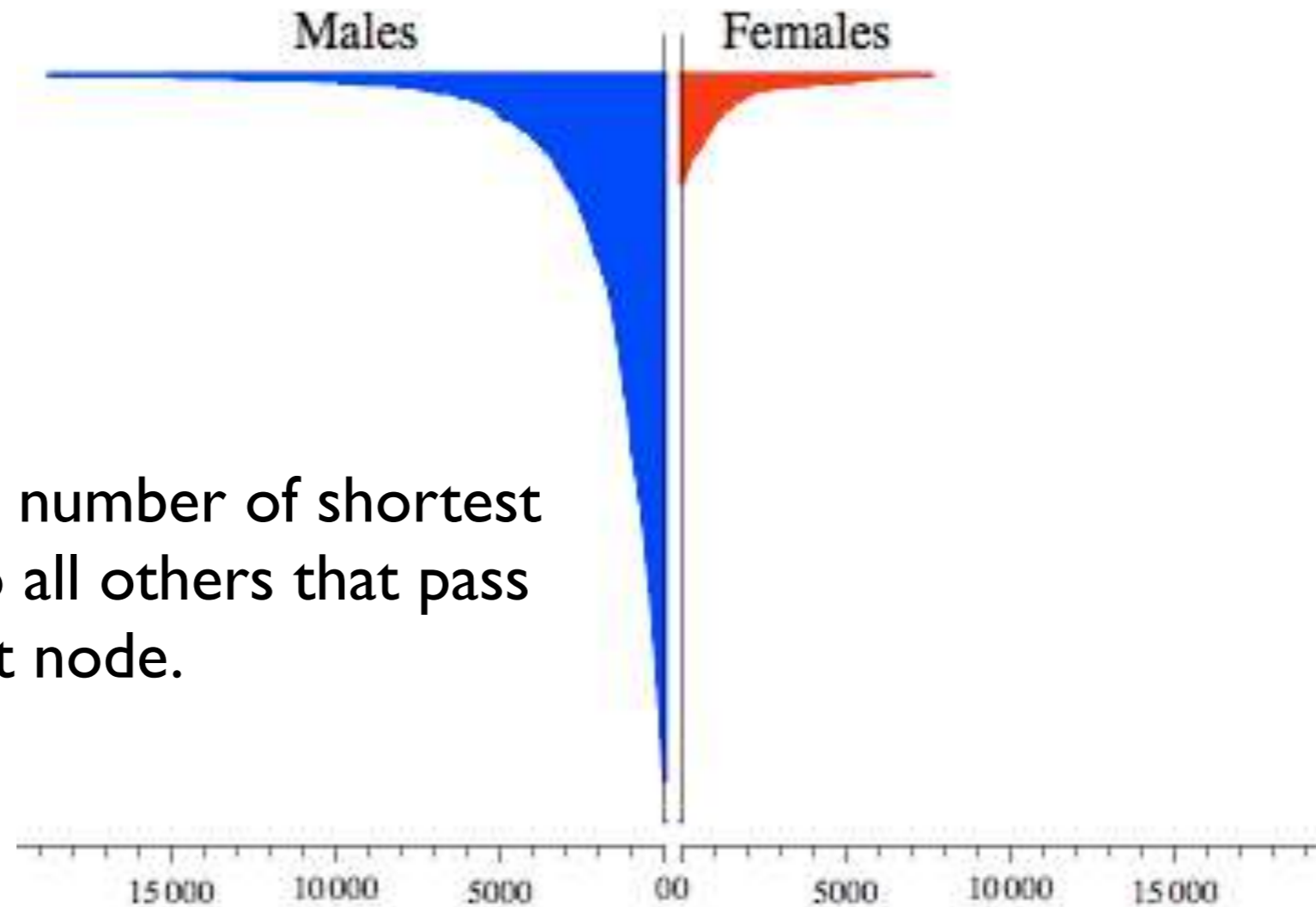
A social network: DBLP





Are men and women different?

(from the viewpoint of DBLP)

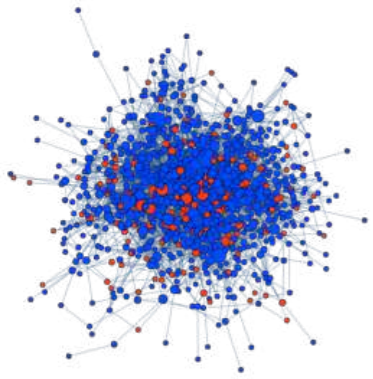


Betweenness centrality = number of shortest paths from all vertices to all others that pass through that node.

DBLP - top coauthorship network

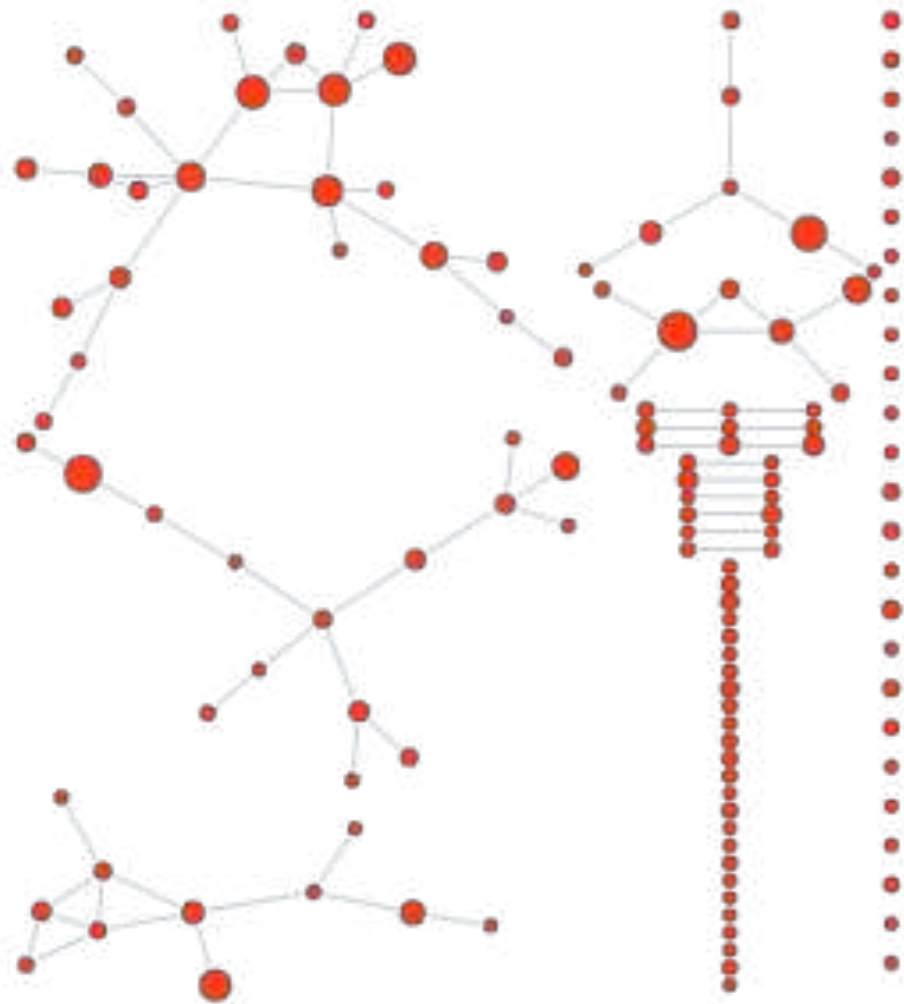
Are men and women different?

(from the viewpoint of DBLP)

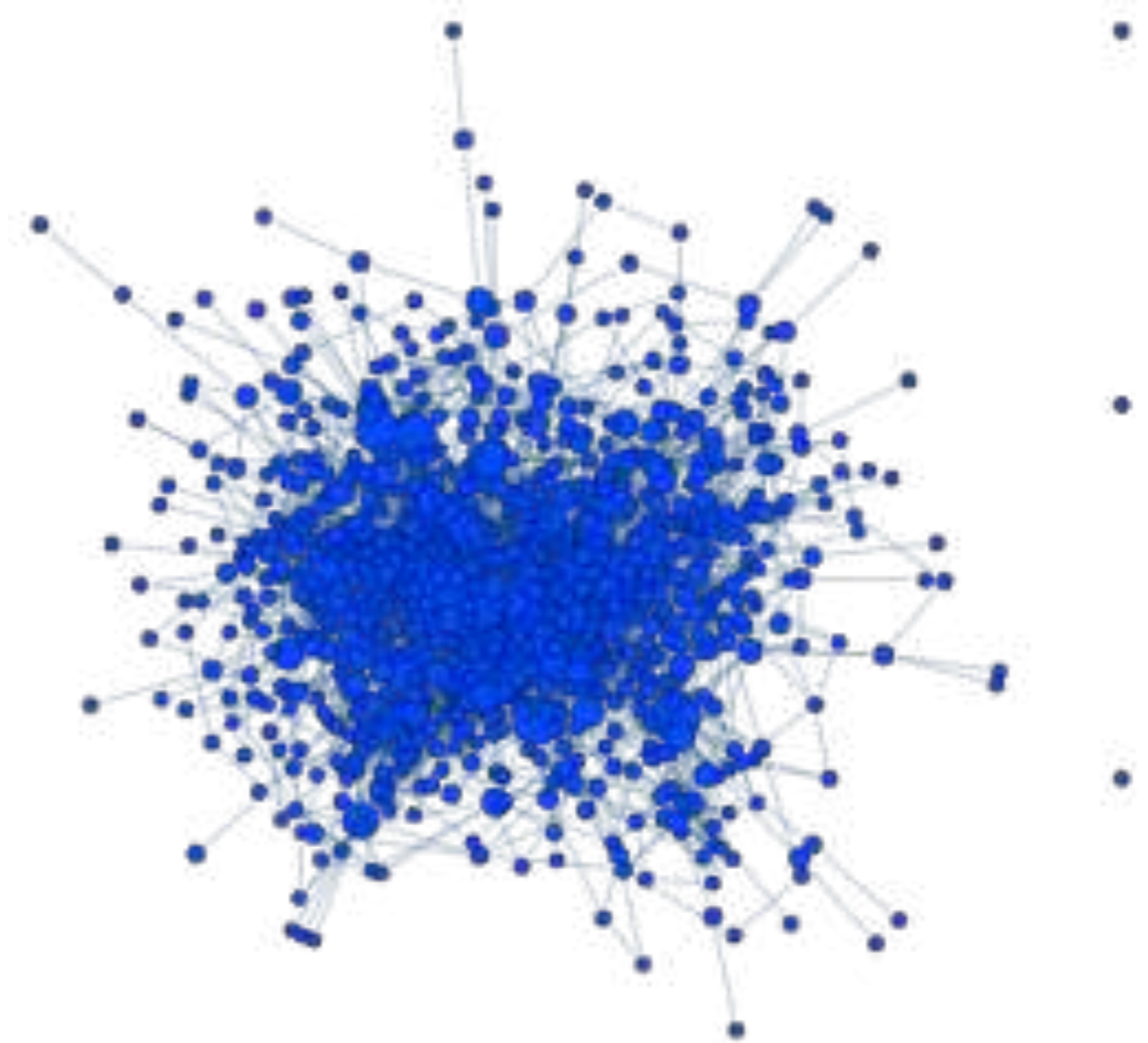


Induced subgraphs

Female network

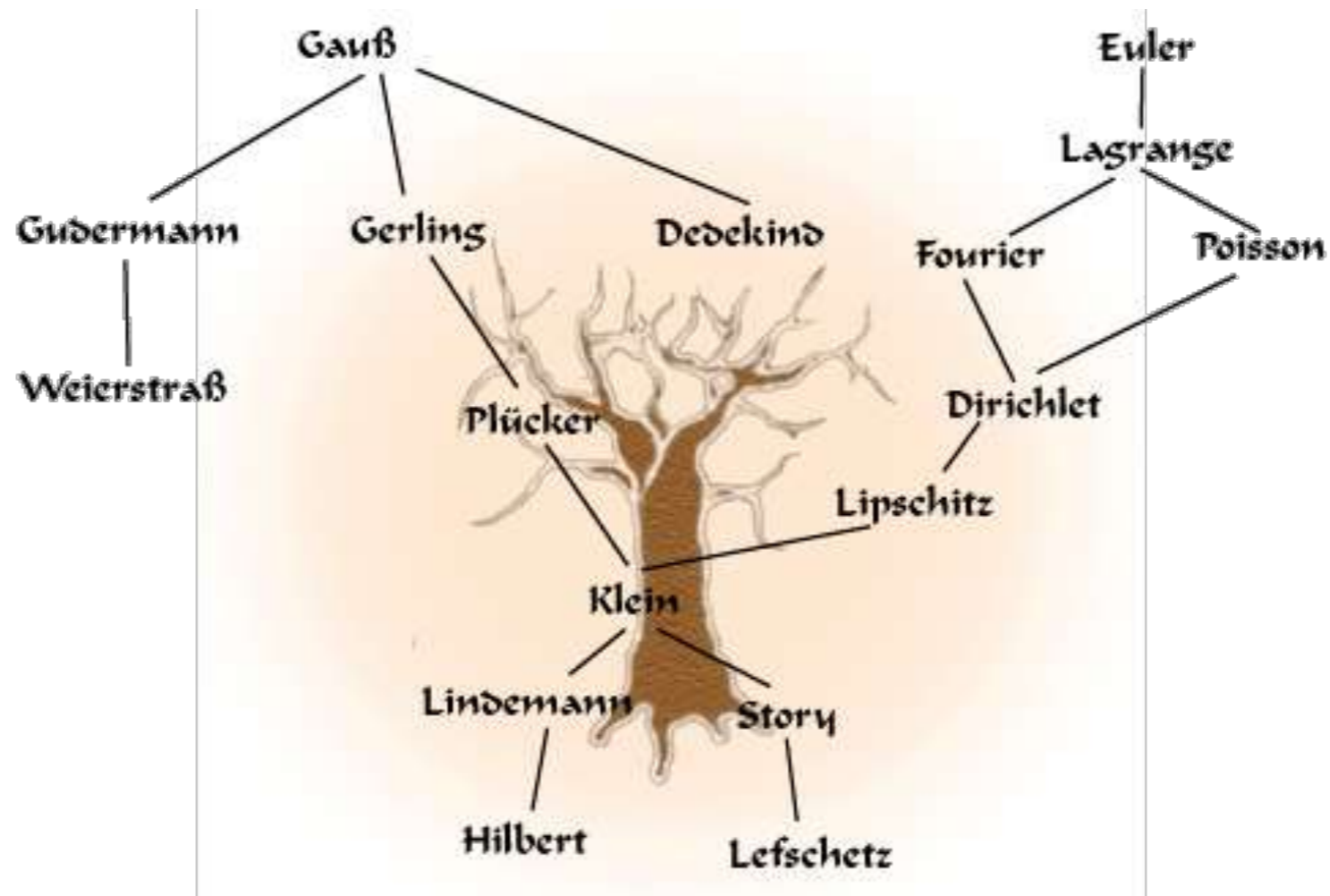


Male network



DBLP - top coauthorship network

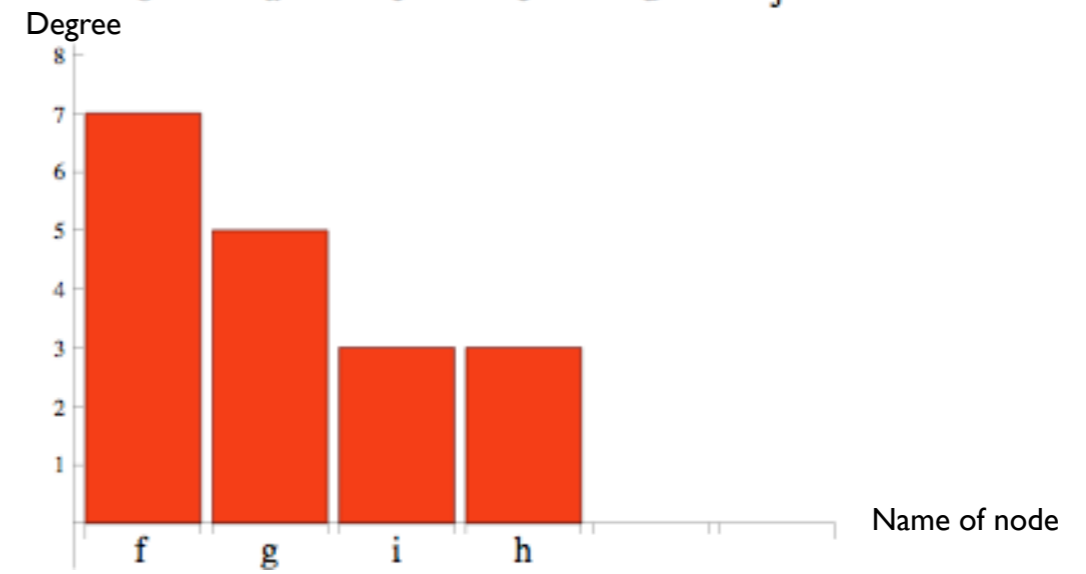
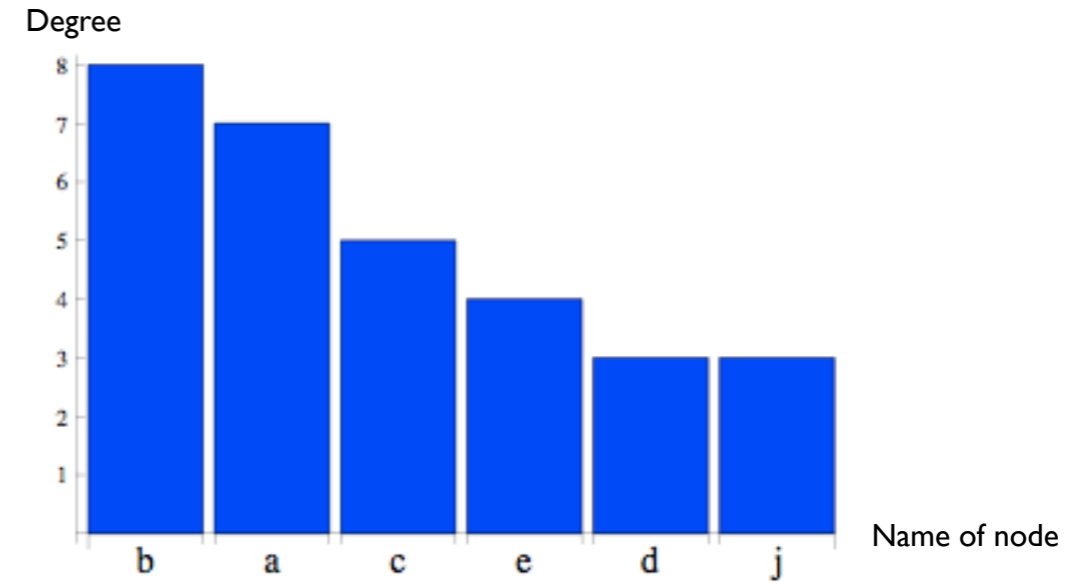
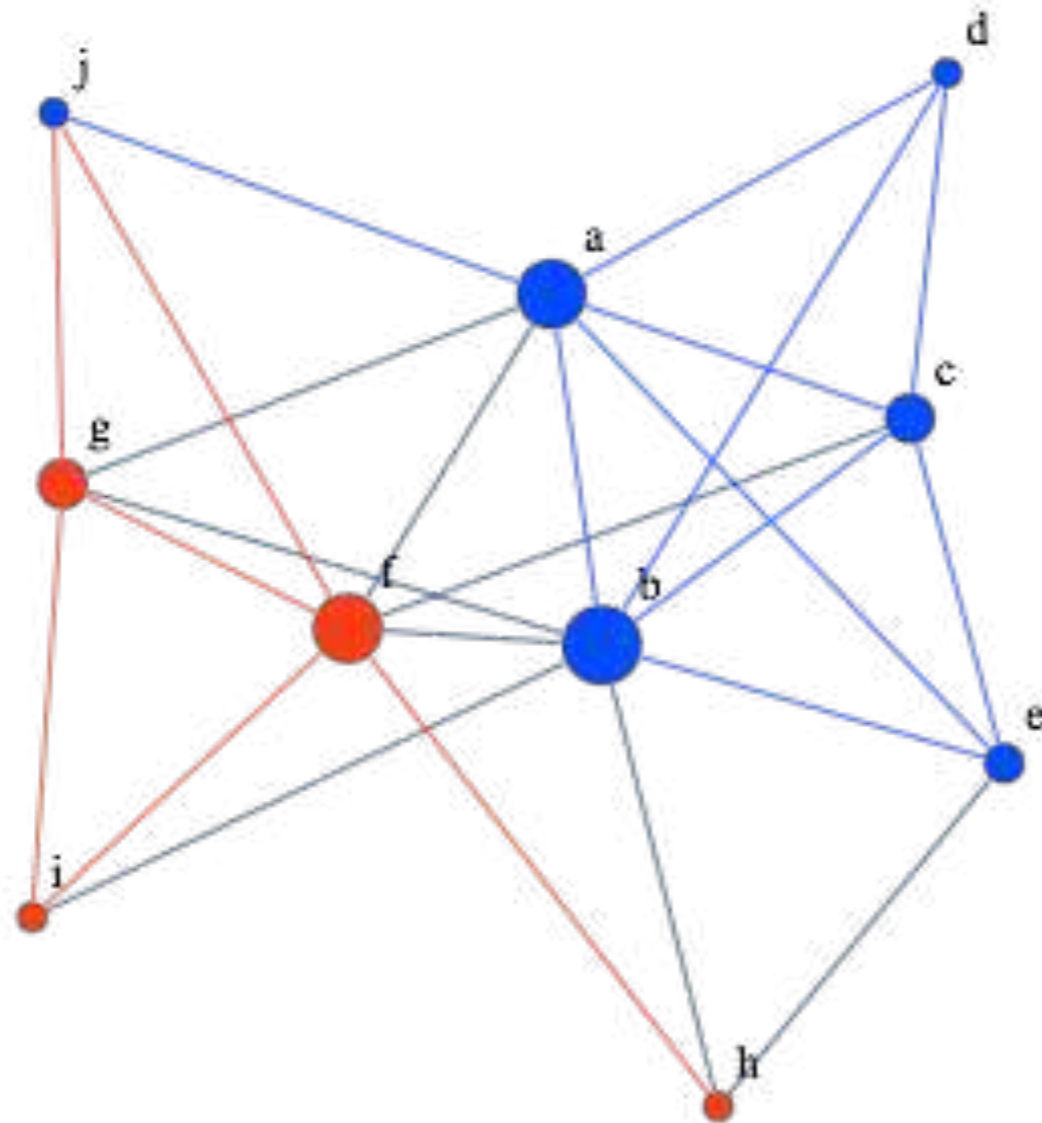
Another social network: Student-Mentor (advisor) graph in academia



- people: nodes
- edge: advisor-advisee, mentor-mentee

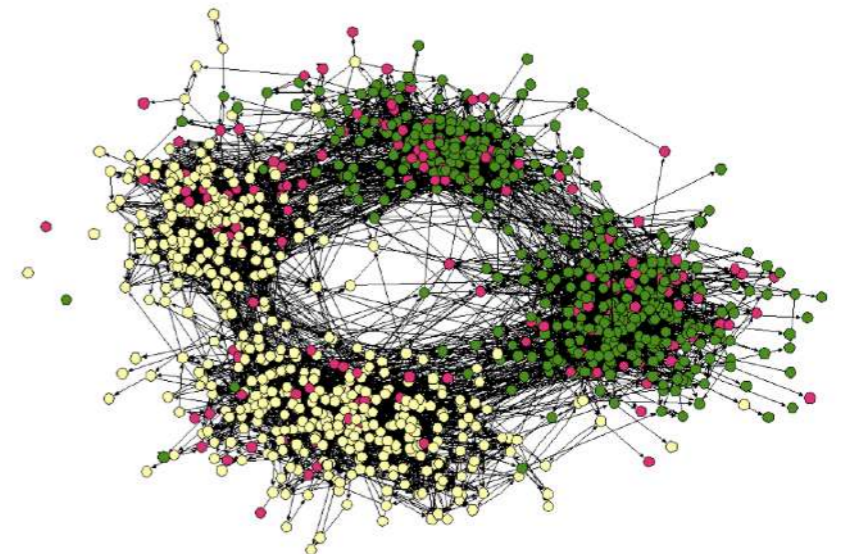
From DBLP, extract gender and Student-Mentor graph

“Most powerful”, “upper rungs”: high degree



Social concepts used for the model

- “rich get richer” - preferential attachment
- homophily - tendency of individuals to associate and bond with similar individuals.
- minority



Model glass ceiling in a social network: Biased Preferential Attachment (BPA) - $G(n, r, \rho)$

1. New node (student) arrives. **Minority** parameter:
 - **Red** with probability $r < 1/2$
 - **Blue** with probability $1 - r$
2. Selects a neighbor (advisor) by **preferential attachment** $\frac{\delta_t(v)}{\sum_{i \in V(t)} \delta_t(i)}$
3. Checks neighbor's color. **Homophily** effect:
Same color: accept
Different color: accept with probability ρ ,
otherwise go to step 2.

of (red, blue) nodes with degree at least k : $\text{top}_k(\mathbf{R}), \text{top}_k(\mathbf{B})$

- **Tail glass ceiling:** $G(n)$ exhibits **tail glass ceiling effect** for the red nodes if:

$$\lim_{G \rightarrow \infty} \frac{\text{top}_k(\mathbf{R})}{\text{top}_k(\mathbf{B})} \rightarrow 0$$

while $k=k(n)$ is s.t.

$$\text{top}_k(\mathbf{B}) \rightarrow \infty$$

- **Moment glass ceiling:** ratio of expected square degree of red vs. of blue vertices

Main Theorem

- If $0 < r < \frac{1}{2}$ and $0 < \rho < 1$ then: **glass ceiling**
- If no minority ($r = \frac{1}{2}$) or no homophily ($\rho = 1$) then:
no glass ceiling
- If no preferential attachment (new vertex selects advisor uniformly) then: **no moment glass ceiling.**

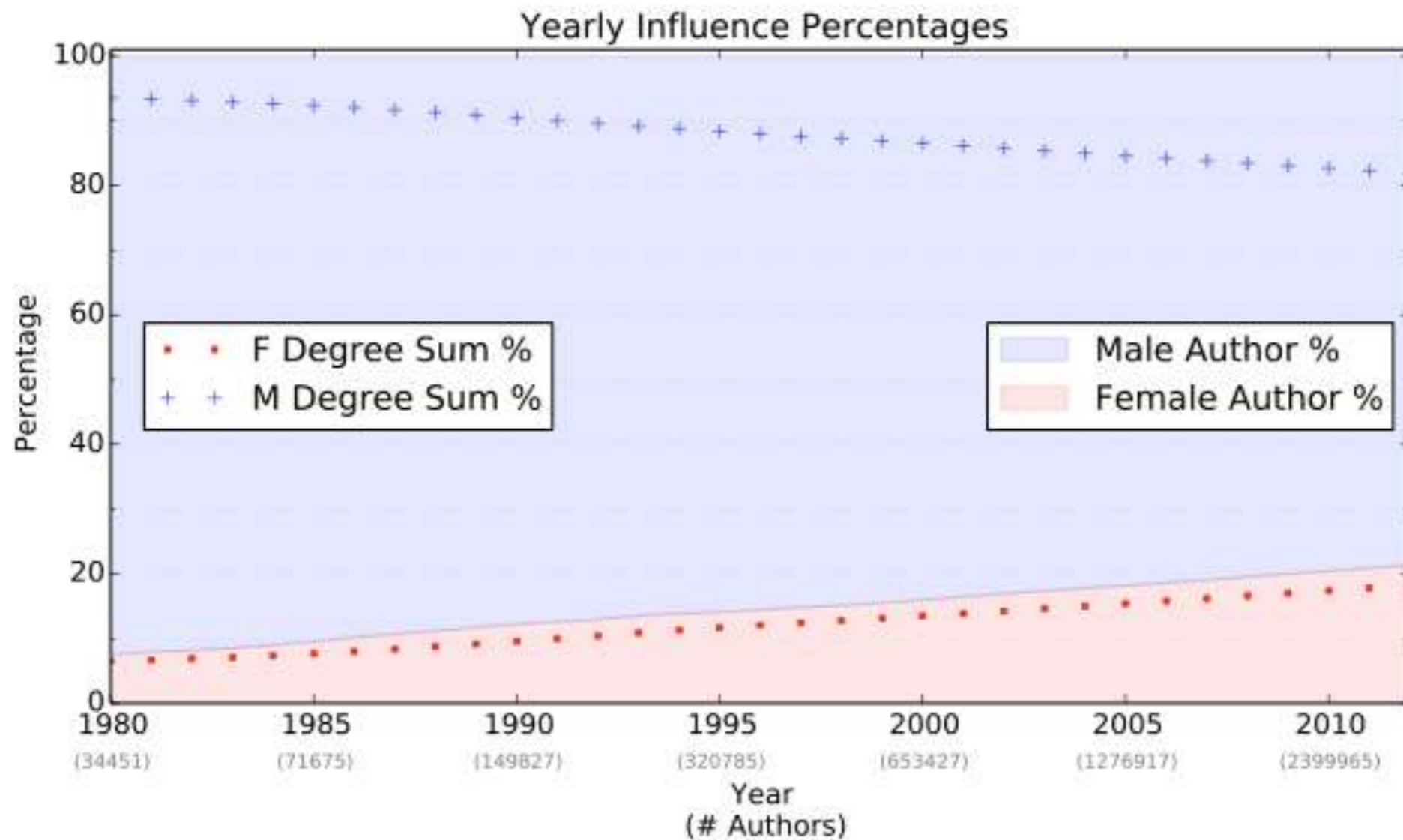
I.e.: three assumptions \rightarrow glass ceiling ,
any two assumptions \rightarrow no glass ceiling

Compare with **empirical** data :

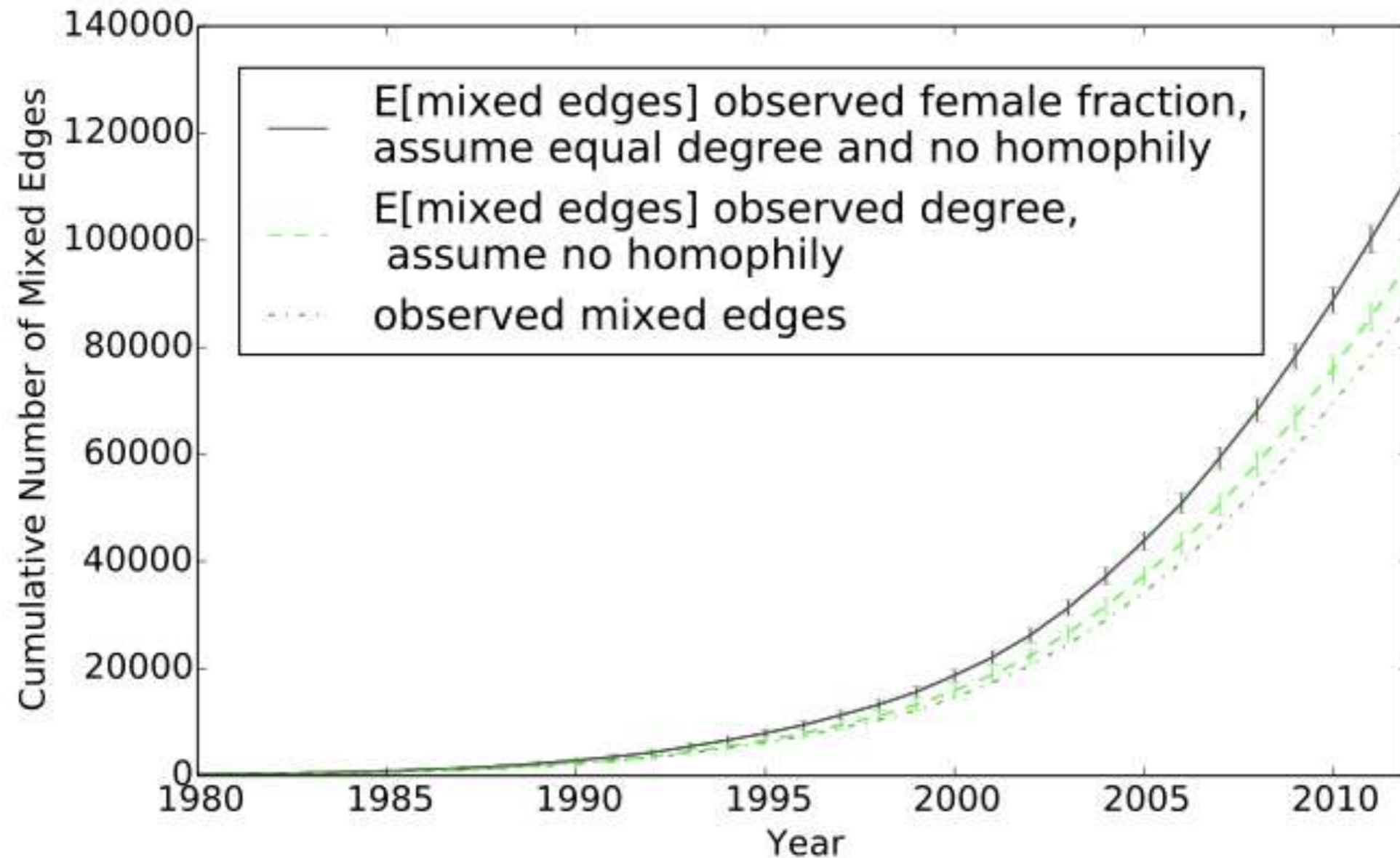
Any connection to real life? -

Student-advisor links in DBLP

Are women a minority in DBLP?



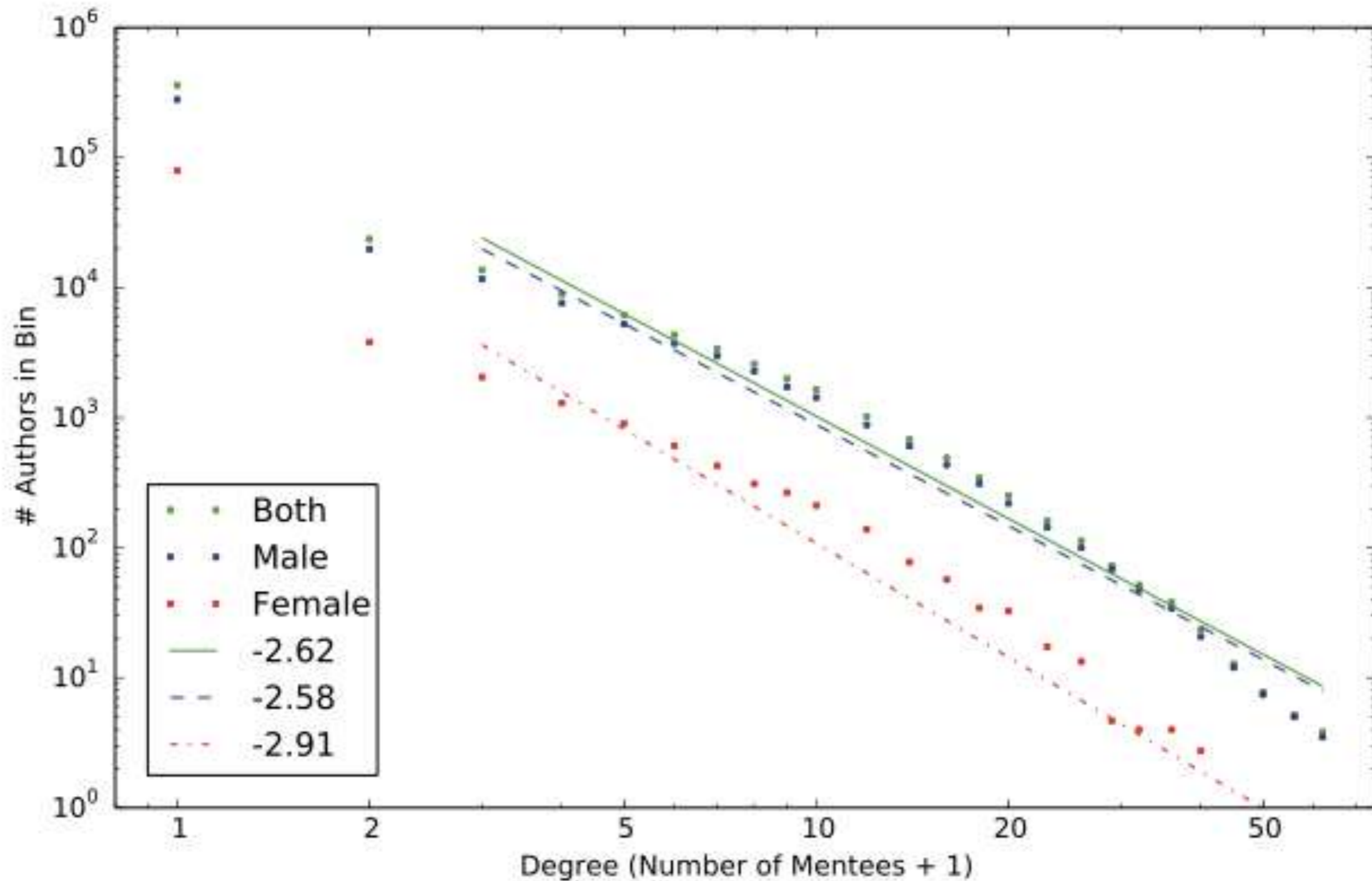
Is there homophily in DBLP?



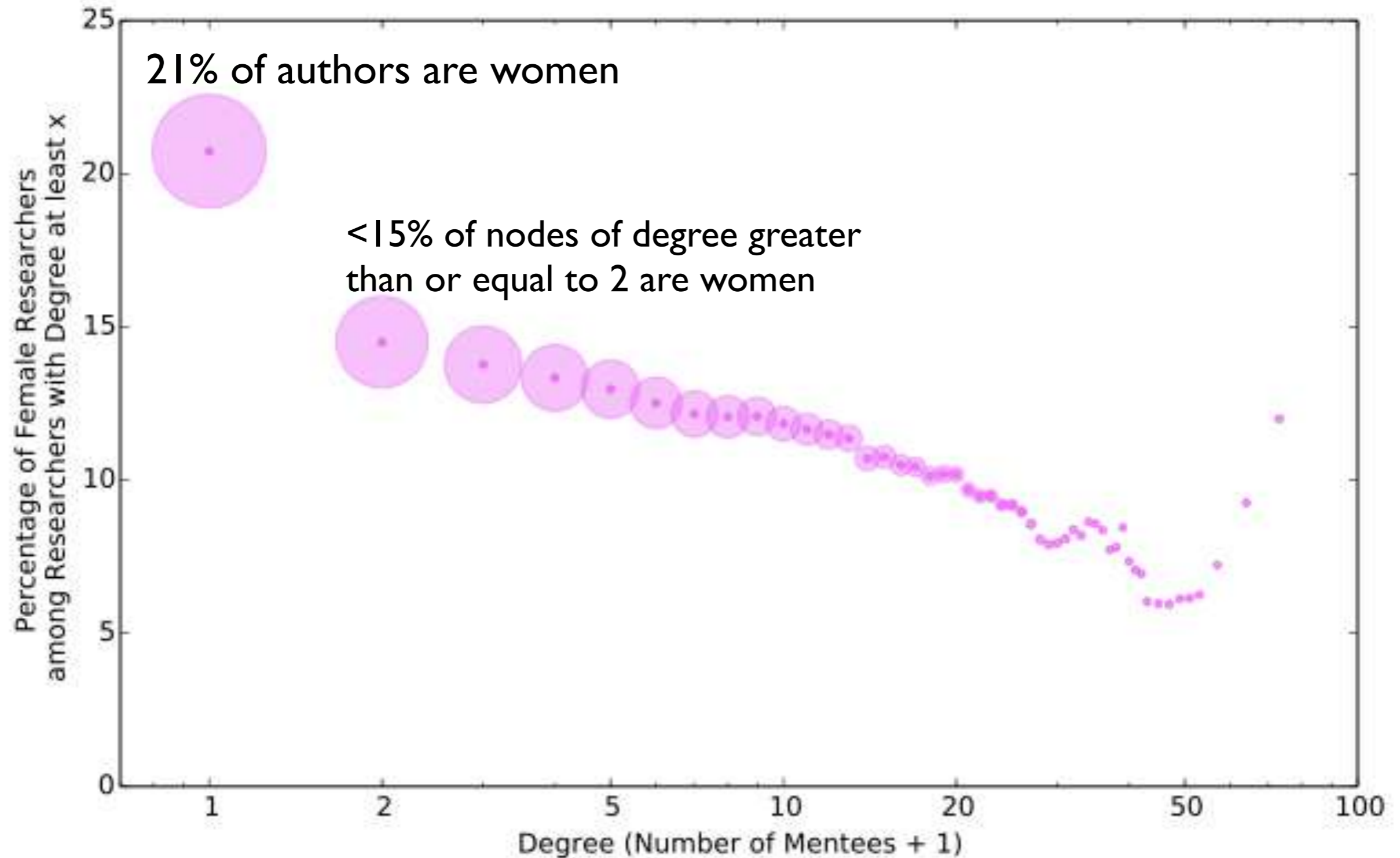
Homophily: unconditional or conditional on total degree

Is there preferential attachment in DBLP?

Distribution of degrees



Is there a glass ceiling in DBLP?



Conclusion

Modéliser, analyser, construire
Impact potentiel: guider politiques publiques
Difficulté mathématique ?
Pas de révolution