

Services et Interactions en Réseau (SIR)

Luigi Lancieri

<http://cristal.univ-lille.fr/~lancieri/>

Luigi.lancieri@univ-lille.fr

<https://www.iut-info.univ-lille1.fr/~lancieri/cours/sir/cours.pdf>

Intelligence collective

Définitions et propriétés

- Capacité émergeant de l'interaction d'un groupe d'individus, dépassant la somme des capacités individuelles.
- Produit : Acquisition de connaissances, décision collective, résolution de problèmes, ..
- L'interaction entre les individus et la décision collective est plus ou moins consciente (e.g. vote ou statistique sur les choix individuels).
- Les interactions collectives ne produisent pas que des effets positifs (paniques, crack boursiers, ..).

Exemples de décisions collectives

- Francis Galton (1906) ; Foire agricole à Plymouth, concours pour deviner le poids d'un bœuf.
 - La moyenne sur 787 estimations s'avéra très proche du poids réel (**543** kilos pour un poids réel de **545**, soit moins de 1% d'erreur).
- Autres expériences similaires
 - Température d'une pièce (**22.2°**, estimée collectivement par une classe à **22.4°**);
 - le classement d'objets par ordre de poids croissant (94% d'exactitude collective);
 - le nombre de bonbons dans un bocal (**850** dragibus pour une estimation collective médiane de **871**);
 - les Oscars 2007, pour lesquels 57 étudiants de l'Université de Columbia ont collectivement deviné **11** vainqueurs sur **12**.
 - Expérience Lumenogic sur la prévision de commercialisation de nouveaux produits : **plus pertinente** que les prévisions basées sur les données dans **plus de 67% des cas**, en réduisant la moyenne d'erreurs d'environ 15% et la portée des erreurs de 40%.

Science de la décision collective

- Le théorème du jury (Condorcet, 1785) dans son essai sur l'application de l'analyse à la probabilité des décisions rendues à la pluralité des voix (1785).
 - Un citoyen moyen à moins d'une chance sur deux de se tromper
 - c'est pourquoi la somme de tous les votes a une faible probabilité d'être inexacte.
- En sciences humaine, sous-discipline de l'économie et des sciences politique Kenneth Arrow (1951) et Duncan Black (1958)
- En informatique : Simulation et systèmes multi-agents en intelligence artificielle

Formation de la décision collective



- 4 éléments :
 - des participants,
 - des options entre lesquelles les participants ont à choisir,
 - les préférences des participants à l'égard de ces options ou leurs manières de les hiérarchiser
 - une règle dite d'agrégation qui associe à toute distribution des préférences une option.
- processus d'agrégation
 - Vote global (majorité, unanimité, ..)
 - Vote réduit (experts, ceux qui participent le plus, ..)
 - Consensus apparent (sans vote) : une proposition (e.g de l'animateur) qui ne suscite pas d'objections
 - Phénomènes d'influence : suivisme, publicité, manipulation, confiance, attention, rôle du temps (spontanéité).



Computer Mediated communication (CMC)

La machine est un lien actif entre les individus

Intelligence collective et intelligence artificielle

- Comment obtenir la décision collective ?
 - Google page rank
 - Facebook, publicité ciblée (scandale analytica)
 - GPS Waze, Wikipedia et open source
 - Open food facts (Open data)
 - Interfaces utilisateurs et scénarii d'usages adaptés aux votes (tumb  )

Traces d'activités et décision collective

- L'ordinateur enregistre des traces qui reflètent les préférences individuelles
- Chaque décision individuelle (achat, consultation, ..) est comparable à un vote.
- Les traces peuvent être agrégées pour produire une décision issue du groupe.

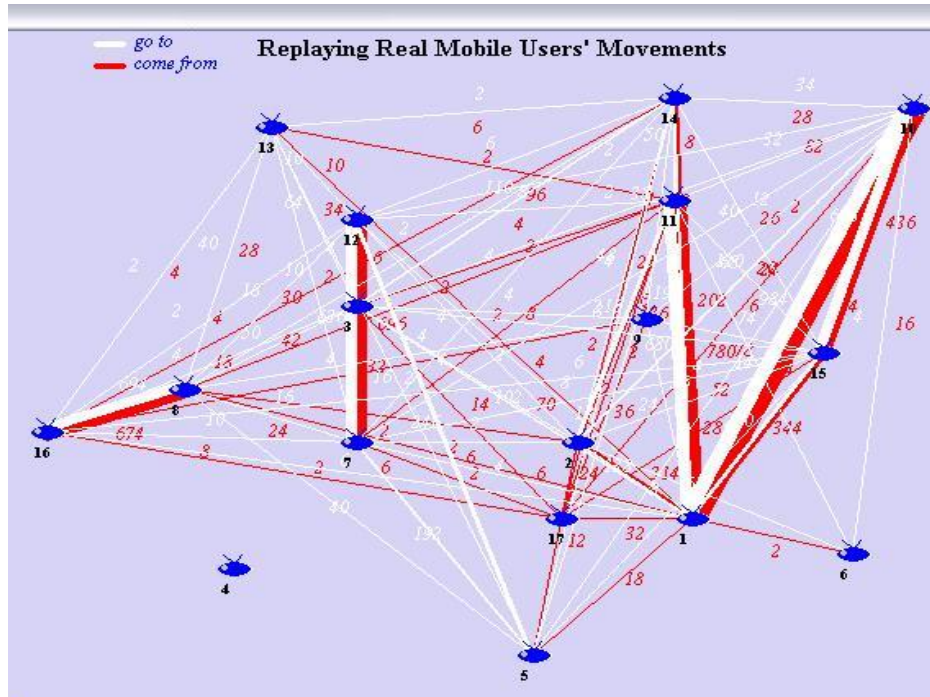
Illustration

Cas du e-commerce



Capture des traces de mobilité dans une entreprise

- 17 WIFI Spots
- 125 days period
- 240 mobile users.



(Benayoune et Lancieri, 2006)

Les systèmes de recommandations :

Exploiter les choix des usagers pour anticiper leurs centres d'intérêts

Quelques données

- Amazon déclare qu' 1/3 de ses ventes est généré par son moteur de recommandation.
- Yankee group (2008) indique qu'un moteur de recommandation permet d'augmenter de 5% le revenu des sites en ligne.
- Certaines chaînes publicitaires interactives ont vu la consommation de leurs programmes multiplié par 5.

Types de recommandations

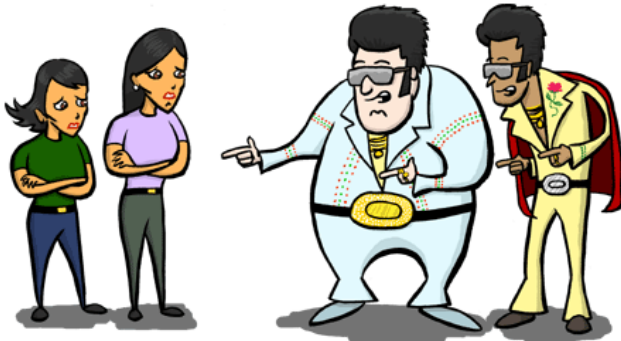
- Le choix collectif résulte de plus ou moins d'interactions
 - Recommandation **éditoriale** : correspond à une vision globale statique (peu ou pas d'interactions) : Choix dominants, best-of, ..
 - Recommandation **sociale** : Filtrage collaboratif et leaders d'opinions, Marketing viral, avis réseaux sociaux
- Personnalisation : mesure de similarités entre le profil d'un usager et les caractéristiques d'un produit.
- Quelques acteurs et logiciels
 - Media unbound (musique), Orca (compass), Epiphany (crm), Kxen
 - **En open source**: Taste , Racofi, lenskit, easyrec, ..

Distances et similarités

- Mesurer la similarité entre deux vecteurs correspondant aux profils ou aux caractéristiques des produits.
 - Distance Euclidienne, Cosinus, Jaccard, etc

• *Similar due to inheritance*

*Similar due to...
uh...other factors*

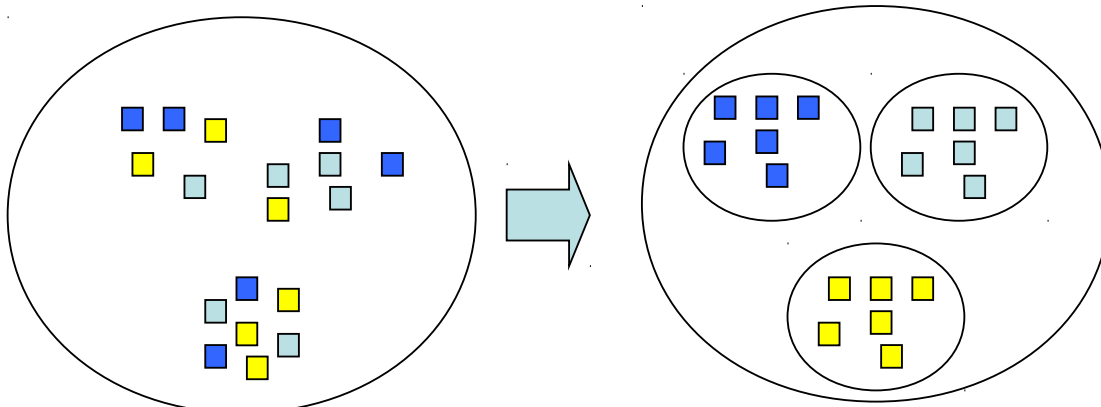


Qu'est ce que la similarité ?

Source evolution.Berkeley.edu

Classification

- Utilisation des distances pour constituer des groupes (eg. communautés d'intérêts)
- Recherche de similarités (ressemblance / dissemblance) entre les items (caractéristiques plus ou moins proches).
- Classification hiérarchique, type K-means, réseaux de neurones, etc



Cas de la distance de Jaccard

Principe : 1- Le rapport entre **taille de l'intersection** sur la **taille de l'union** entre deux ensembles. $\Rightarrow D = 1 - \frac{(U1 \cap U2)}{(U1 \cup U2)}$

Exemple : 4 usagers indiquent les films qu'ils ont apprécié

	SpiderMan	StarWars	Colombo	Maigret
User 1	1	1	0	1
User 2	0	1	0	1
User 3	1	1	1	1
User 4	1	0	1	0

- $D(U1, U2) = 1 - 2/3 = 0.33$ (starwars, maigret) / (spiderman, starwars, maigret)
- $D(U1, U3) = 1 - 3/4 = 0.25$
- $D(U1, U4) = 1 - 1/4 = 0.75$

User1 est plus proche de User3 (0,25) que de U.2 ou de U.4

Règles d'associations

- Mettre en évidence les items liés
 - Produits achetés ensembles
 - Films correspondant à un même « goût »
- Règles de la forme
 - si produit 1 alors produit 2
 - Si produit 1 et produit 2 alors produit 3
- On mesure la fiabilité (support) et la précision (confiance) des règles

Un exemple simple

Objectif : **Prédire** les cases vides de la Matrice d'utilité et surtout les cases avec des **notes** élevées.

Question : User 2 appréciera t'il Spiderman ?

- 1 **Matrice d'utilité :** Combien de fois un usager a t'il vu un film (ou acheté un produit) ?

	SpiderMan	StarWars	Colombo	Maigret
User 1	5	4		1
User 2	???	4		1
User 3	1	2	4	5
User 4	5		1	

Un exemple simple

Objectif : **Prédire** les cases vides de la Matrice d'utilité et surtout les cases avec des **notes** élevées.

Question : User 2 appréciera t'il Spiderman ?

- 1 **Matrice d'utilité :** Combien de fois un usager a t'il vu un film (ou acheté un produit) ?

	SpiderMan	StarWars	Colombo	Maigret
User 1	5	4		1
User 2	???	4		1
User 3	1	2	4	5
User 4	5		1	

- 2 **Matrice de cooccurrences ou de distances**

- le nombre de fois où deux éléments sont apparus ensembles avec **note élevée**
- **ou distance entre films**

	SM	SW	C	M
SM		2	1	0
SW			2	0
C				1
M				

Un exemple simple

Objectif : **Prédire** les cases vides de la Matrice d'utilité et surtout les cases avec des **notes** élevées.

Question : User 2 appréciera t'il Spiderman ?

1 **Matrice d'utilité** : Combien de fois un usager a t'il vu un film (ou acheté un produit) ?

	SpiderMan	StarWars	Colombo	Maigret
User 1	5	4		1
User 2	???	4		1
User 3	1	2	4	5
User 4	5		1	

2 **Matrice de cooccurrences ou de distances**

- le nombre de fois ou deux éléments sont apparus ensembles avec **note élevée**
- **ou** distance entre films

	SM	SW	C	M
SM	2	1	0	0
SW		2	0	0
C			1	1
M				1

3 **Règles de recommandation**

Si C alors M
Si SW alors SM