

Optimisation des recommandations d'amis sur un réseau social

MATTAN Aishwaria Sania

Numéro de candidat : XXXXXXXX

Cycles et boucles

Travaux d'initiative personnelle encadrés

Table des matières

1 Introduction

2 Stratégies et métriques

3 Simulation

4 Conclusion

Introduction

- Les systèmes de recommandation guident nos connexions sur les réseaux sociaux.
- Mais nous ouvrent-ils vraiment à l'autre, ou nous renvoient-ils toujours aux mêmes profils ?

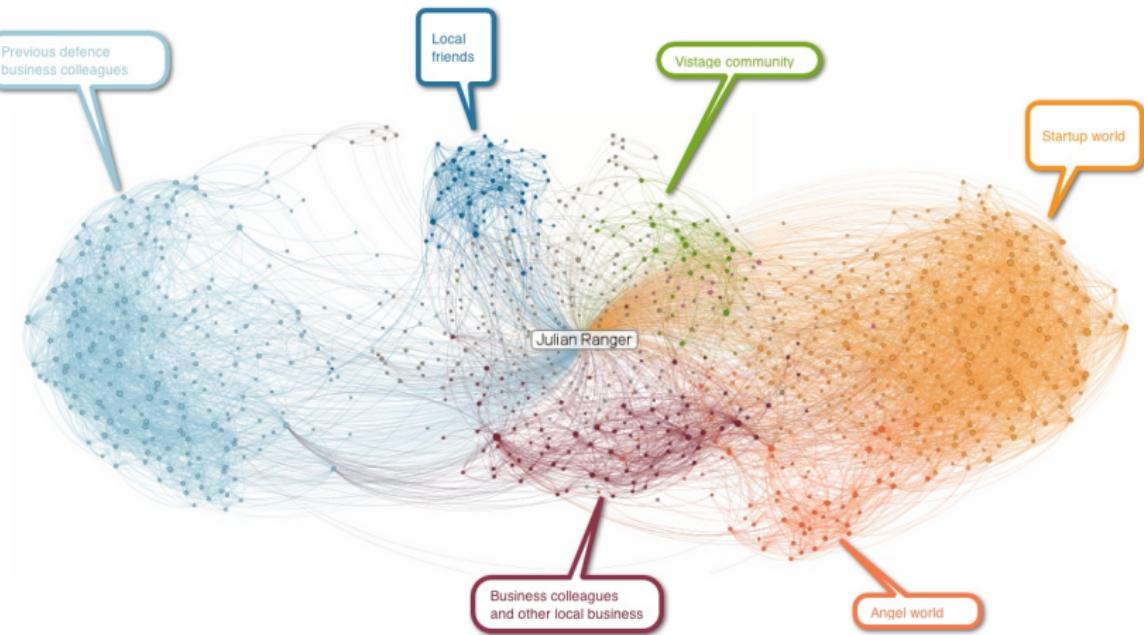
Introduction

- Les systèmes de recommandation guident nos connexions sur les réseaux sociaux.
- Mais nous ouvrent-ils vraiment à l'autre, ou nous renvoient-ils toujours aux mêmes profils ?

Lien avec le thème : *Cycle et boucle*

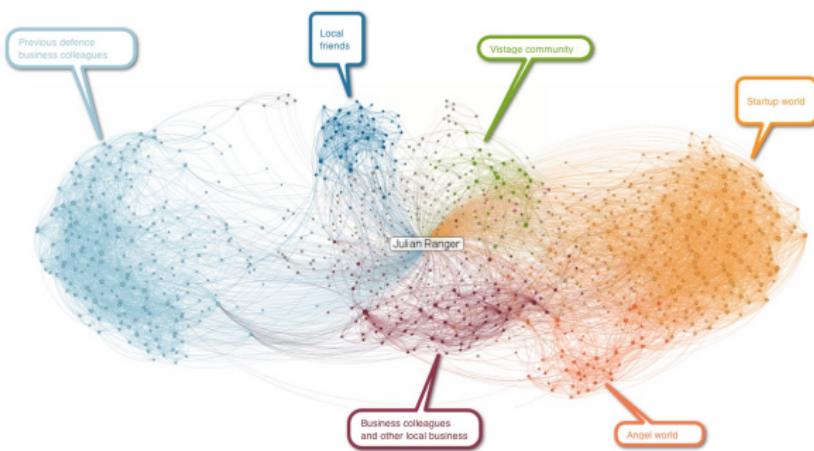
- Dans un graphe social, un **cycle** = un retour vers un profil déjà connu.
- Trop de cycles \Rightarrow peu de diversité, risque de fermeture.

Introduction



www.reseau-canope.fr

Introduction



www.reseau-canope.fr

Problématique

Comment concevoir une stratégie qui favorise l'ouverture du réseau et un système de recommandation qui étend le réseau personnel d'un utilisateur en dehors de son voisinage immédiat, en se plaçant sur le graphe du réseau social ?

Table des matières

1 Introduction

2 Stratégies et métriques

3 Simulation

4 Conclusion

Stratégies

Stratégie par liens

- Basée sur les amitiés communes
- Arête ajoutée si ≥ 3 amis en commun
- Stratégie locale sur les voisins proches

Stratégie par intérêts

- Basée sur un vecteur de préférences
- Utilisation de la similarité cosinus
- Permet des connexions éloignées

Métriques

- **Nombre de cycles** : effet de fermeture

Métriques

- **Nombre de cycles** : effet de fermeture
- **Clustering coefficient** : densité locale

Métriques

- **Nombre de cycles** : effet de fermeture
- **Clustering coefficient** : densité locale
- **Composantes connexes** : fragmentation locale

Métriques

- **Nombre de cycles** : effet de fermeture
- **Clustering coefficient** : densité locale
- **Composantes connexes** : fragmentation locale
- **Modularité** : fragmentation globale

Métriques

- **Nombre de cycles** : effet de fermeture
- **Clustering coefficient** : densité locale
- **Composantes connexes** : fragmentation locale
- **Modularité** : fragmentation globale

Objectif

Mesurer comment chaque stratégie influence la structure du graphe.

Table des matières

1 Introduction

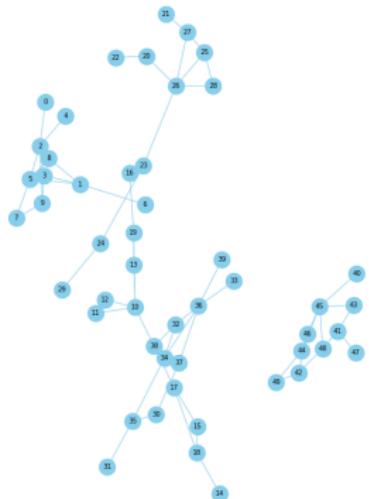
2 Stratégies et métriques

3 Simulation

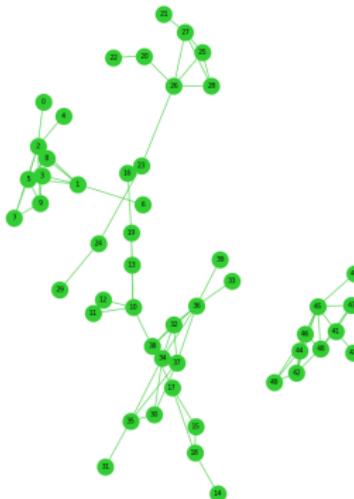
4 Conclusion

Simulation

Graph initial



Recommandation par liens



Recommandation par intérêts



Moyenne sur 30 essais

Graphe initial

- ↳ Nombre d'arêtes : 48.0
- ↳ Clustering coefficient : 0.820
- ↳ Composantes connexes : 13.6
- ↳ Modularity : 0.914

Graphe initial

Recommandation par liens

- ↳ Nombre d'arêtes : 57.6
- ↳ Clustering coefficient : 0.940
- ↳ Composantes connexes : 13.6
- ↳ Modularity : 0.883

Recommandation par liens

Recommandation par intérêts

- ↳ Nombre d'arêtes : 278.3
- ↳ Clustering coefficient : 0.437
- ↳ Composantes connexes : 1.0
- ↳ Modularity : 0.234

Recommandation par intérêts

Moyenne sur 30 essais

● Graphe initial
↳ Nombre d'arêtes : 48.0
↳ Clustering coefficient : 0.820
↳ Composantes connexes : 13.6
↳ Modularity : 0.914

Graphe initial

Par liens

- + aretes
- + clustering
- + composante connexe
- + modularité

● Recommandation par liens
↳ Nombre d'arêtes : 57.6
↳ Clustering coefficient : 0.940
↳ Composantes connexes : 13.6
↳ Modularity : 0.883

Recommandation par liens

Par intérêts

- ++ arêtes
- – clustering
- – composante connexe
- – modularité

● Recommandation par intérêts
↳ Nombre d'arêtes : 278.3
↳ Clustering coefficient : 0.437
↳ Composantes connexes : 1.0
↳ Modularity : 0.234

Recommandation par intérêts

Moyenne sur 30 essais

Graphe initial

- ↳ Nombre d'arêtes : 48.0
- ↳ Clustering coefficient : 0.820
- ↳ Composantes connexes : 13.6
- ↳ Modularity : 0.914

Graphe initial

Par liens

- + aretes
- + clustering
- + composante connexe
- + modularité

Recommandation par liens

- ↳ Nombre d'arêtes : 57.6
- ↳ Clustering coefficient : 0.940
- ↳ Composantes connexes : 13.6
- ↳ Modularity : 0.883

Recommandation par liens

Par intérêts

- ++ arêtes
- – clustering
- – composante connexe
- – modularité

Recommandation par intérêts

- ↳ Nombre d'arêtes : 278.3
- ↳ Clustering coefficient : 0.437
- ↳ Composantes connexes : 1.0
- ↳ Modularity : 0.234

Recommandation par intérêts

Interprétation

L'ouverture du graphe est meilleure avec les recommandations par intérêts.

Table des matières

1 Introduction

2 Stratégies et métriques

3 Simulation

4 Conclusion

Conclusion

- Le type de recommandation influence la structure du graphe.
- Recommander par liens d'amitié renforce les cycles et l'enfermement.
- Recommander par intérêts permet plus de diversité.

Conclusion

- Le type de recommandation influence la structure du graphe.
- Recommander par liens d'amitié renforce les cycles et l'enfermement.
- Recommander par intérêts permet plus de diversité.

Ouverture

- Se placer sur un graphe orienté

Conclusion

- Le type de recommandation influence la structure du graphe.
- Recommander par liens d'amitié renforce les cycles et l'enfermement.
- Recommander par intérêts permet plus de diversité.

Ouverture

- Se placer sur un graphe orienté
- Combiner les deux stratégies

Conclusion

- Le type de recommandation influence la structure du graphe.
- Recommander par liens d'amitié renforce les cycles et l'enfermement.
- Recommander par intérêts permet plus de diversité.

Ouverture

- Se placer sur un graphe orienté
- Combiner les deux stratégies
- Tester sur des données réelles

Annexes

Coefficient de clustering

Définition — Le coefficient de *clustering*.

$$C(p) = \frac{1}{n} \sum_{i=1}^n C_i \quad (1)$$

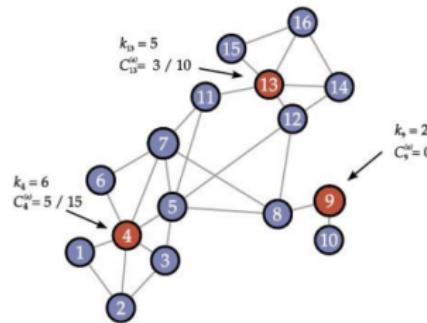
avec

$$C_i = \frac{\text{nombre de triangles dont un des trois sommets est le nœud } i}{\binom{k}{2}} \quad (2)$$

où k est le coefficient local ou le degré du nœud et n le nombre de nœuds du graphe.

Coefficient de clustering

Exemple Avec le graphe présenté ci-dessous,



$$C_4 = \frac{5}{\binom{6}{2}} = \frac{5}{15} = \frac{1}{3}$$

et le coefficient de *clustering* du réseau est :

$$C(p) = \frac{1}{n} \sum_{i=1}^n C_i = 0,5208.$$