

A primer on network analysis for business

Table of Contents

1. Définitions	1
a. Réseaux sociaux	2
b. Autres réseaux	3
c. Quelle peut être la taille des réseaux ?	5
d. Comment discuter des réseaux ? Un peu de vocabulaire	5
2. Réseaux : pour quelle utilisation ?	6
a. Segmentation et catégorisation	6
b. Trouver des acteurs clés	7
c. comprendre comment l'information se propage	8
d. Identification des modèles - pour la détection, le contrôle ou le renseignement des fraudes... ..	9
3. Pour aller plus loin	10
Contact	11

last modified: 2023-05-15



1. Définitions

Un réseau est un ensemble de données constitué d'entités et de leurs relations

Les scientifiques utilisent le terme "graphe" pour parler des réseaux.



Figure 1. Ceci est un réseau

a. Réseaux sociaux

En tant qu'utilisateurs, nous connaissons très bien un type de réseaux - les réseaux sociaux :





LinkedIn



Figure 2. source : <http://www.minanacheva.com/getting-visual-with-facebook-data/>

b. Autres réseaux

Il est important de réaliser que les réseaux couvrent plus que les relations entre humains. Par exemple, il est possible d'imaginer un réseau constitué de recettes de cuisine. 2 ingrédients sont liés s'ils apparaissent fréquemment dans les mêmes recettes.

En scannant toutes les recettes et leurs ingrédients depuis un site de recettes de cuisine, cela donne :

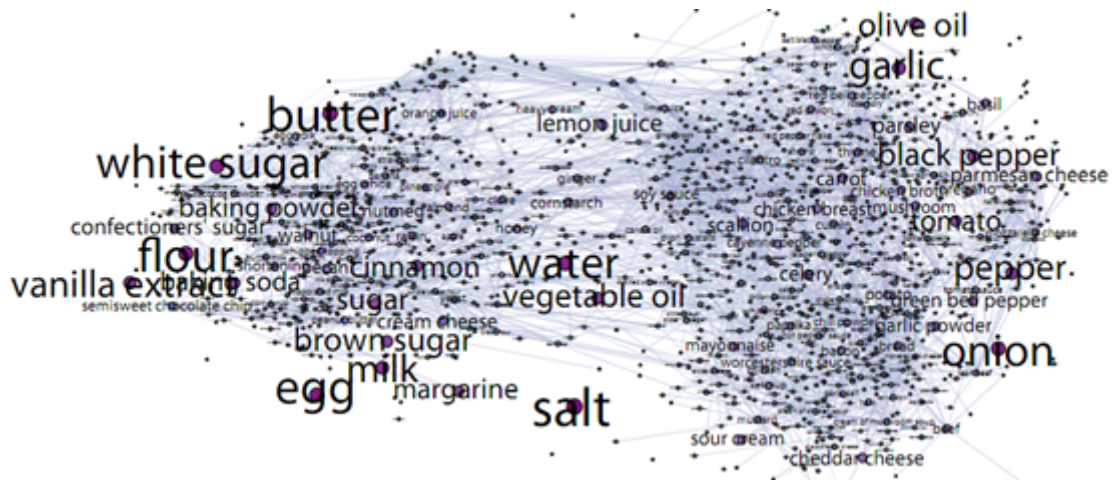


Figure 3. source : <http://arxiv.org/abs/1111.3919>

Les réseaux sémantiques sont une autre grande catégorie de réseaux. La méthode est la même : il faut trouver un moyen de « relier » les mots dans un texte, puis on obtient un réseau.

L'idée générale est la même que dans les recettes de cuisine : 2 termes d'un texte seront connectés dans le réseau s'ils apparaissent fréquemment dans les mêmes paragraphes.



Figure 4. source : <http://www.nature.com/naturejournal/v463/n7278/full/463157a.html>

c. Quelle peut être la taille des réseaux ?

Avec une augmentation de la puissance de calcul à l'ère du big data et des bases de données NOSQL adéquates (telles que [Neo4J](#) ou [OrientDB](#)), nous pouvons faire face à d'énormes réseaux:

Par exemple, "[The Anatomy of the Facebook Social Graph](#)" (2011)

→ étude de 721 millions d'utilisateurs Facebook actifs et des 69 milliards (!) de liens d'amitié qui les relie.

Une limite est vite atteinte en termes de visualisation : il est difficile de faire tenir des millions de nœuds sur un écran. Dans la visualisation suivante, nous pouvons voir un réseau de 90 000 suédophones et leurs relations sur Twitter. La vue est très encombrée.

(ouvrir la source pour une version interactive)

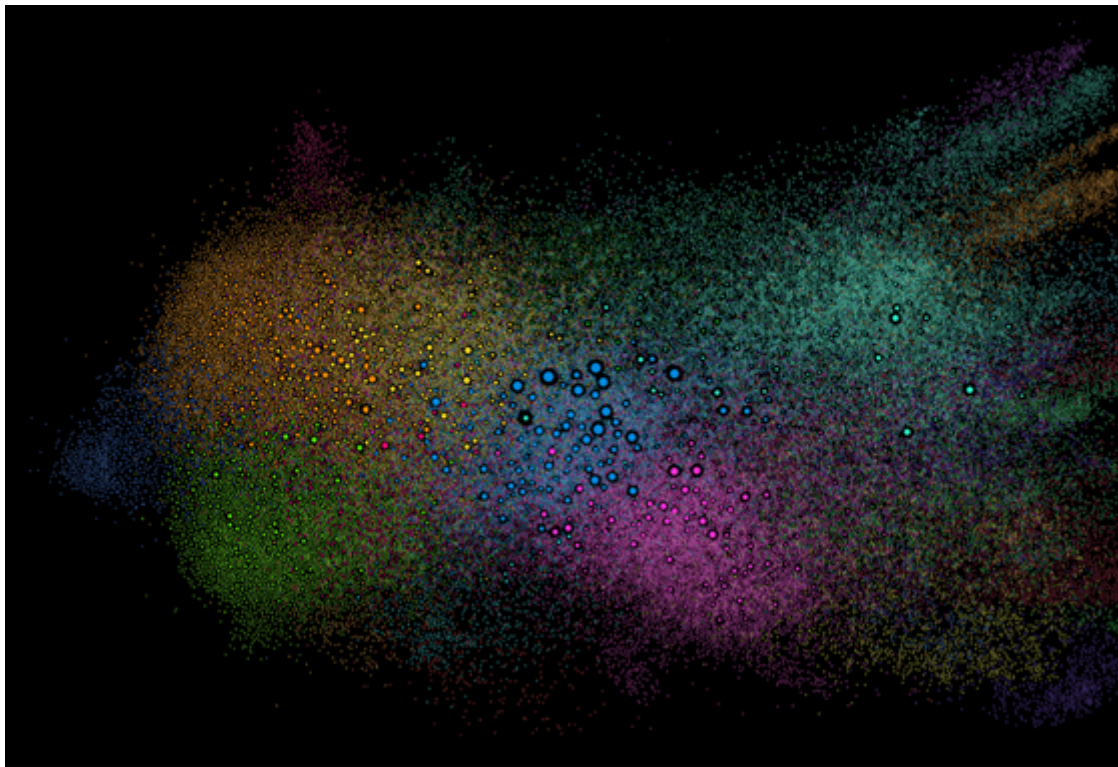


Figure 5. source : <http://twittercensus.se/graph2015/>

d. Comment discuter des réseaux ? Un peu de vocabulaire

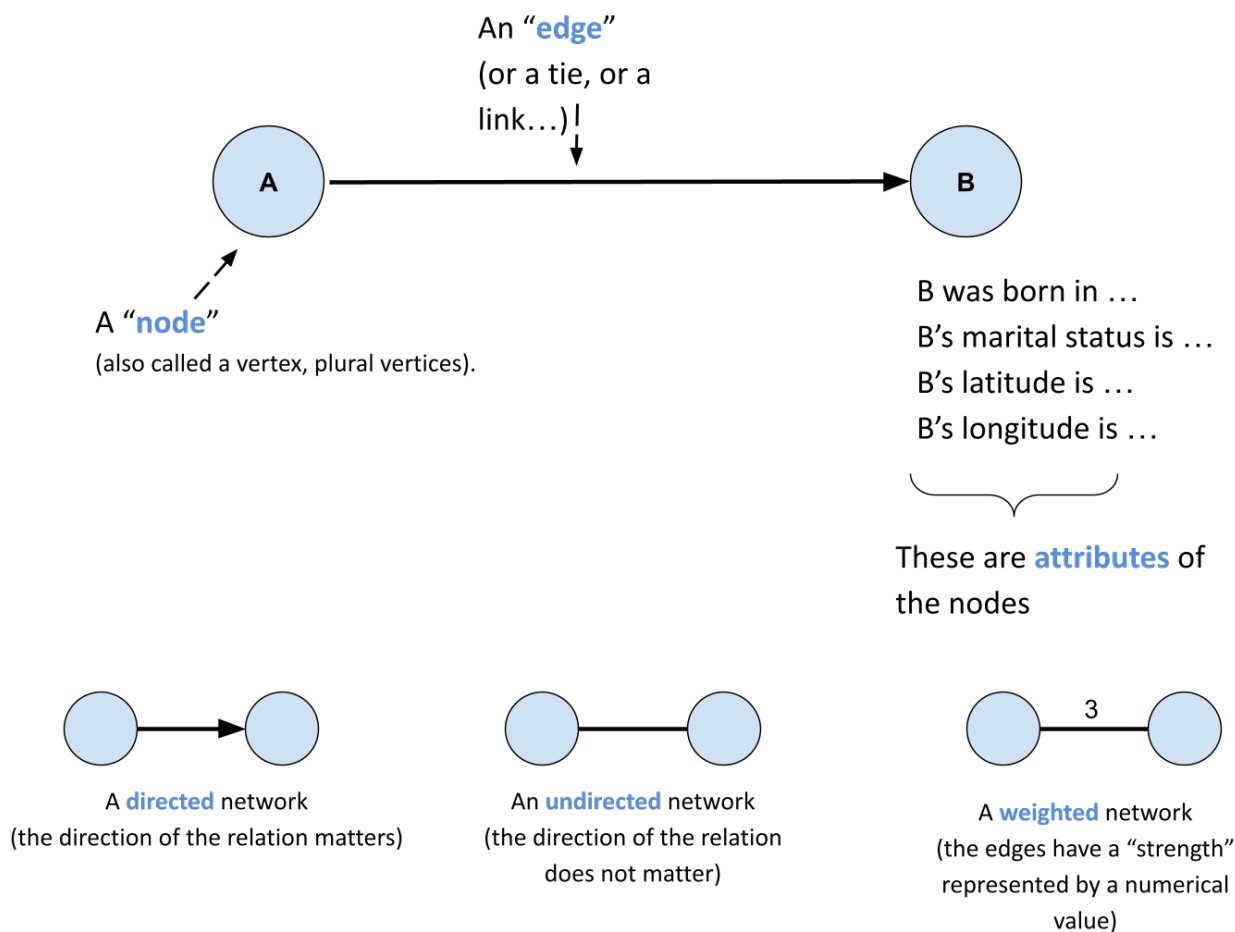


Figure 6. Terminology

2. Réseaux : pour quelle utilisation ?

a. Segmentation et catégorisation

Si un réseau est constitué d'entités et de leurs relations, alors un segment est un sous-groupe d'entités du réseau, qui a une certaine cohésion ou quelque chose en commun.

Ce sous-groupe de nœuds du réseau est souvent appelé une « **communauté** ».

La détection de communautés dans un réseau, aussi appelée "clustering", consiste à trouver des nœuds qui ont de nombreuses connexions en commun.

Il s'agit d'une procédure mathématique et algorithmique, mais elle est très simple à comprendre visuellement :

SEGMENTATION

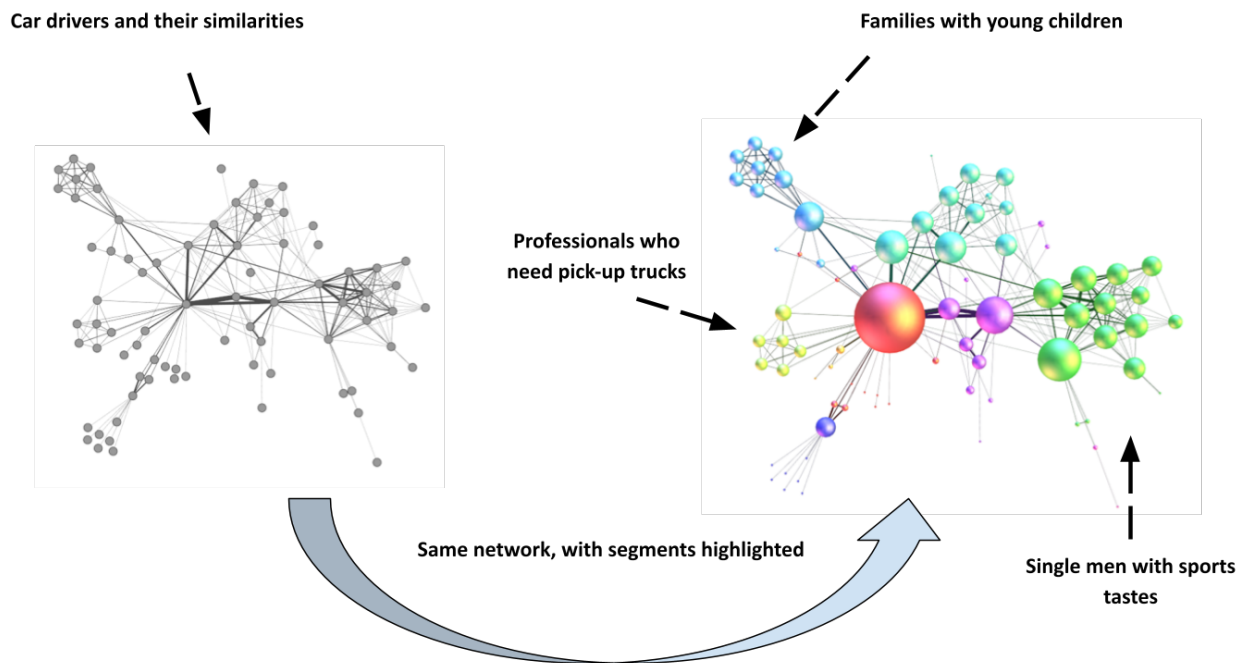


Figure 7. segmentation with community detection in networks

b. Trouver des acteurs clés

FINDING KEY PLAYERS

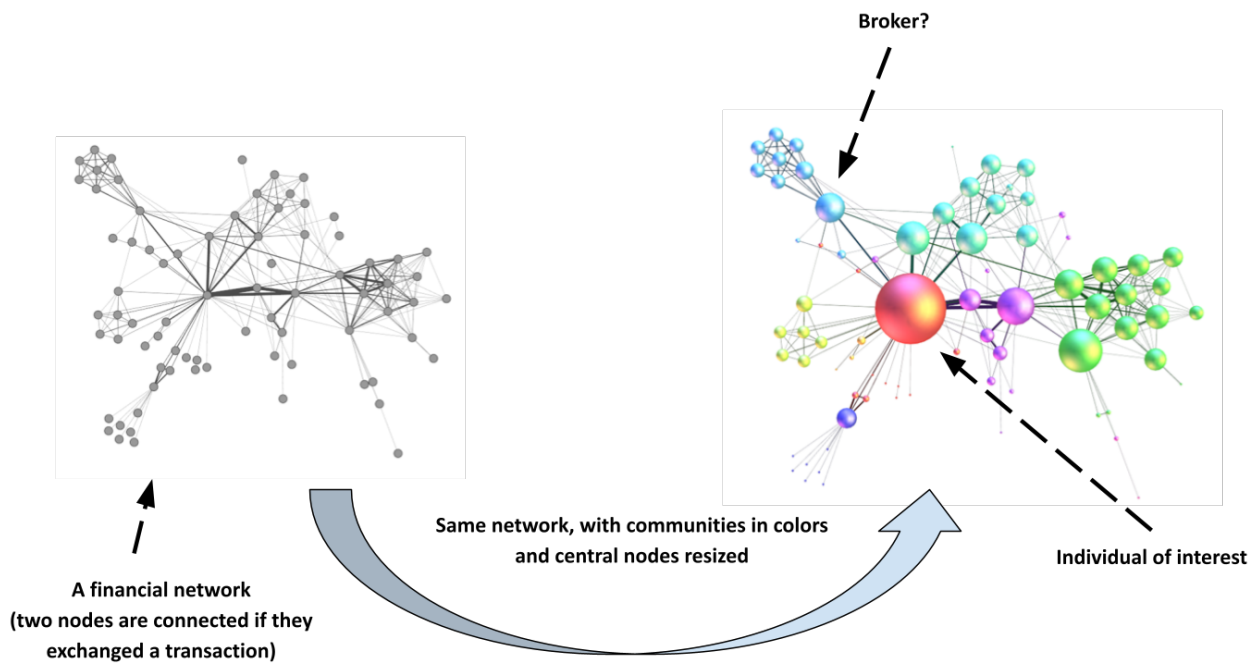


Figure 8. Key players visualized by resizing nodes

c. comprendre comment l'information se propage

Une société de science des données a créé "Où va mon tweet", qui retrace la propagation d'un tweet donné à travers les retweets. Le service est désormais discontinué mais le mécanisme s'explique :

Understanding info spread

Initial tweet (appears at the center)

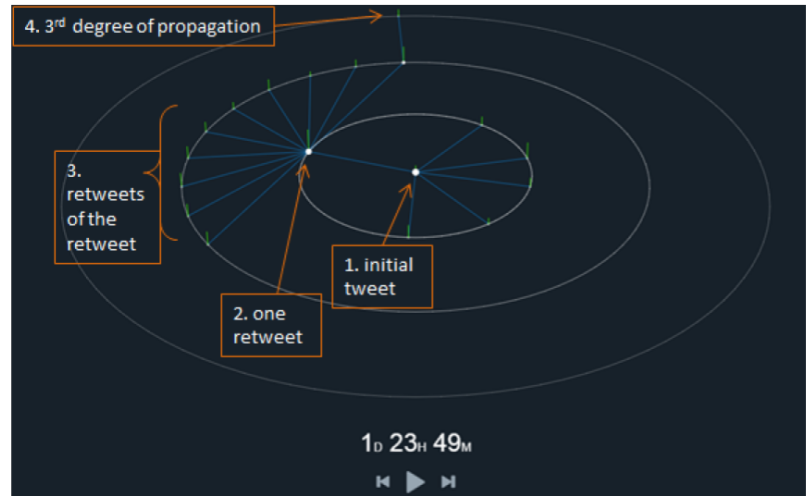
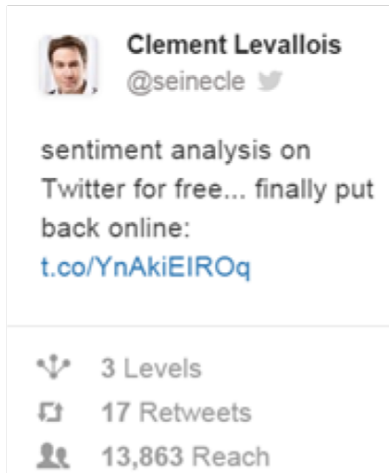


Figure 9. Where Does my Tweet Go by MFGLabs

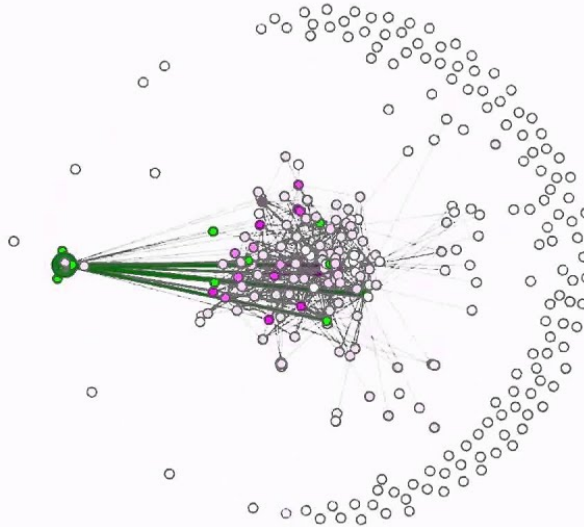
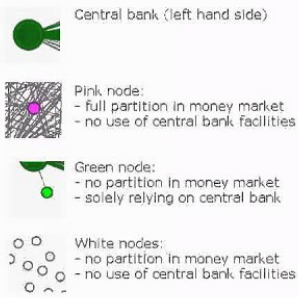
d. Identification des modèles - pour la détection, le contrôle ou le renseignement des fraudes.

Dans la vidéo suivante, on voit les [les banques participant aux prêts inter bancaires en Europe](#). 2 banques sont connectées si l'une prête à l'autre. Le schéma des échanges évolue au fil des ans - les banques se retirent du marché.

The money market network

This network graph shows the central bank and all banks in the payment system that were at any time active in the money market or made use of central bank facilities.

Money market transactions are represented by pink lines, central bank facilities are represented by green lines and consist of monetary loans, standing facilities (marginal lending and overnight deposits) and fixed term deposits.



The size and type of payments will cause a node to be attracted to:

- the money market (right hand side, centre)
- the central bank (left hand side)
- the periphery.



Autre exemple : connecter des mesures de performances commerciales apparemment sans rapport avec [Oracle BI](#) et [Linkurious](#) :

PEAK indicators

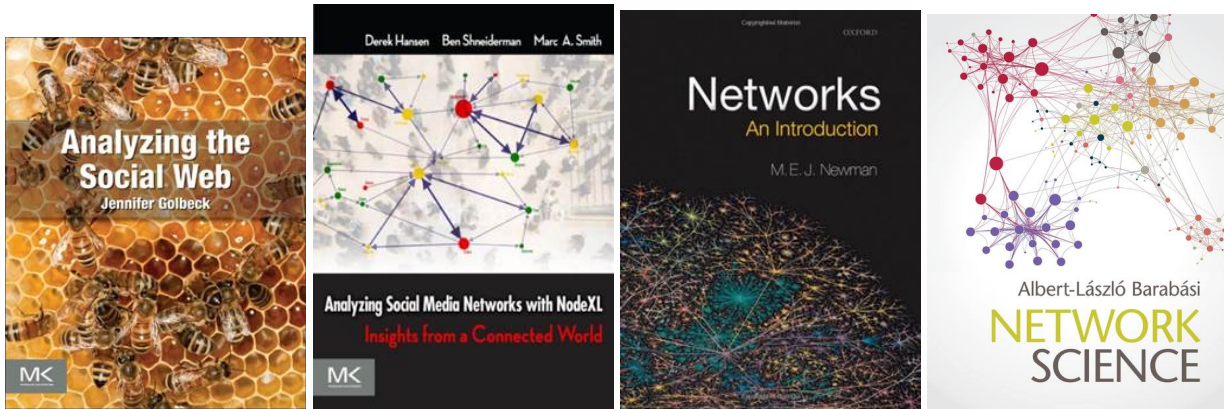
The image shows a dashboard with multiple screens displaying various data visualizations, including charts, maps, and abstract patterns. Below the dashboard is a row of eight green circular icons representing different aspects of data and analytics: gears, a world map, a cloud with arrows, server racks, a bar chart, a person pointing at a screen, stacked coins, and a line graph with an upward arrow.

Big Data and Analytics

Oracle BI and Graph Visualizations

Antony Heljula

3. Pour aller plus loin



Vous pouvez aussi visiter mes tutos sur Gephi, le logiciel leader pour visualiser de grands graphes :

<https://seinecle.github.io/gephi-tutorials/>

Contact

Trouvez des références pour cette leçon, et d'autres leçons, [ici](#).



Ce cours est réalisé par Clément Levallois.

Découvrez mes autres formations en data/tech for business : <https://www.clementlevallois.net>

Ou contactez-moi via Twitter : [@seinecle](#)