

# A primer on network analysis for business

Clément Levallois

2017-12-08

# Table of Contents

|  |   |
|--|---|
| 1. Definitions .....   | 1 |
| a. Social networks .....   | 2 |
| b. Other networks .....  | 2 |
| c. How big can networks be? .....  | 4 |
| d. How to discuss networks? Some vocabulary .....                            | 4 |
| 2. Networks: what use for business? .....                                    | 5 |
| a. Segmentation .....  | 5 |
| b. Finding key players .....   | 6 |
| c. Understanding how information spreads .....                               | 7 |
| d. Identifying patterns - for fraud detection, control or intelligence. .... | 8 |
| 3. To go further .....   | 9 |
| The end .....  | 9 |



# 1. Definitions

A network is a dataset made of entities and their relations

Scientists use the term "graph" to discuss networks.



*Figure 1. This is a network*

## a. Social networks

As users, we are very familiar with one type of networks - social networks:



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

## b. Other networks

It is important to realize that networks cover more than relations between humans:

For example, it is possible to imagine a network made out of cooking recipes. 2 ingredients are

Scanning all recipes and their ingredients from a website of cooking recipes, this gives:



The general idea is the same as in cooking recipes: 2 terms of a text will be connected in the network if they frequently appeared in same paragraphs.



### c. How big can networks be?

With a surge in computing power in the age of big data, and the adequate NOSQL databases (such as [Neo4J](#) or [OrientDB](#)), we can deal with huge networks:

For example, “[The Anatomy of the Facebook Social Graph](#)” (2011)

→ study of 721 million active Facebook users and the 69 billion (!) friendship links connecting them.

A limit is quickly reached in terms of visualization: it is hard to fit millions of nodes on a screen. In the next visualization, we can see a network of 90,000 Swedish speakers and their relations on Twitter. The view is very cluttered.

(open the source for an interactive version)

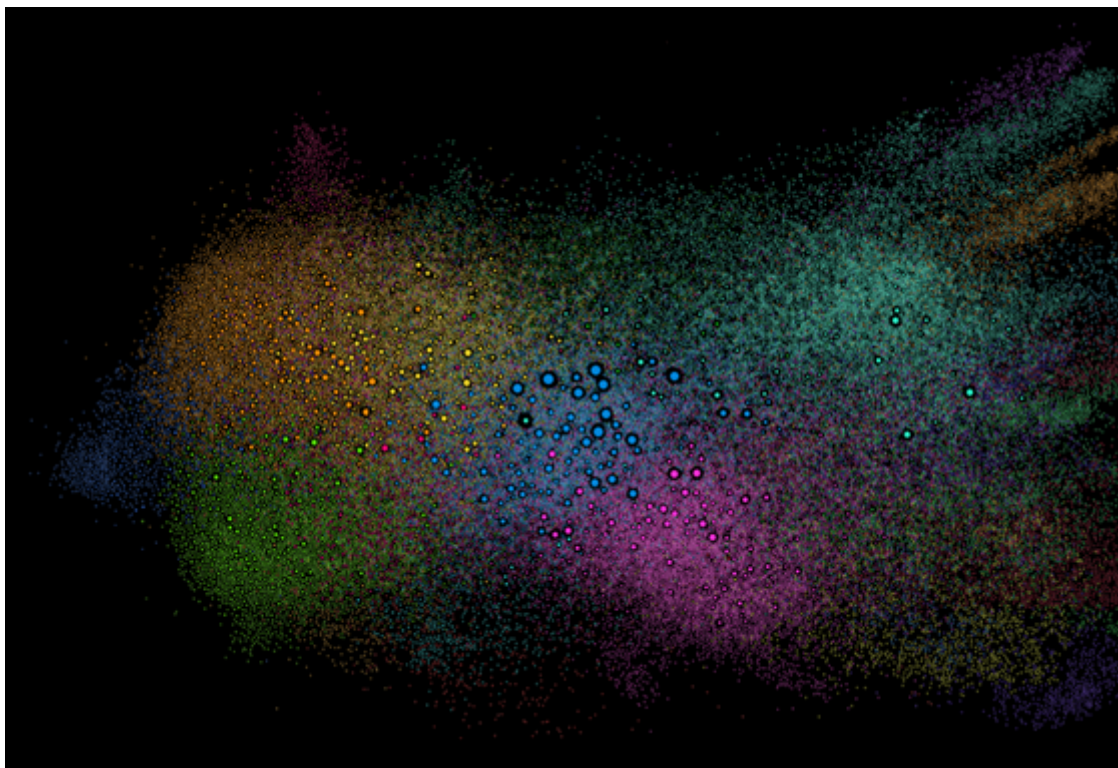


Figure 5. source: <http://twittercensus.se/graph2015/>

### d. How to discuss networks? Some vocabulary



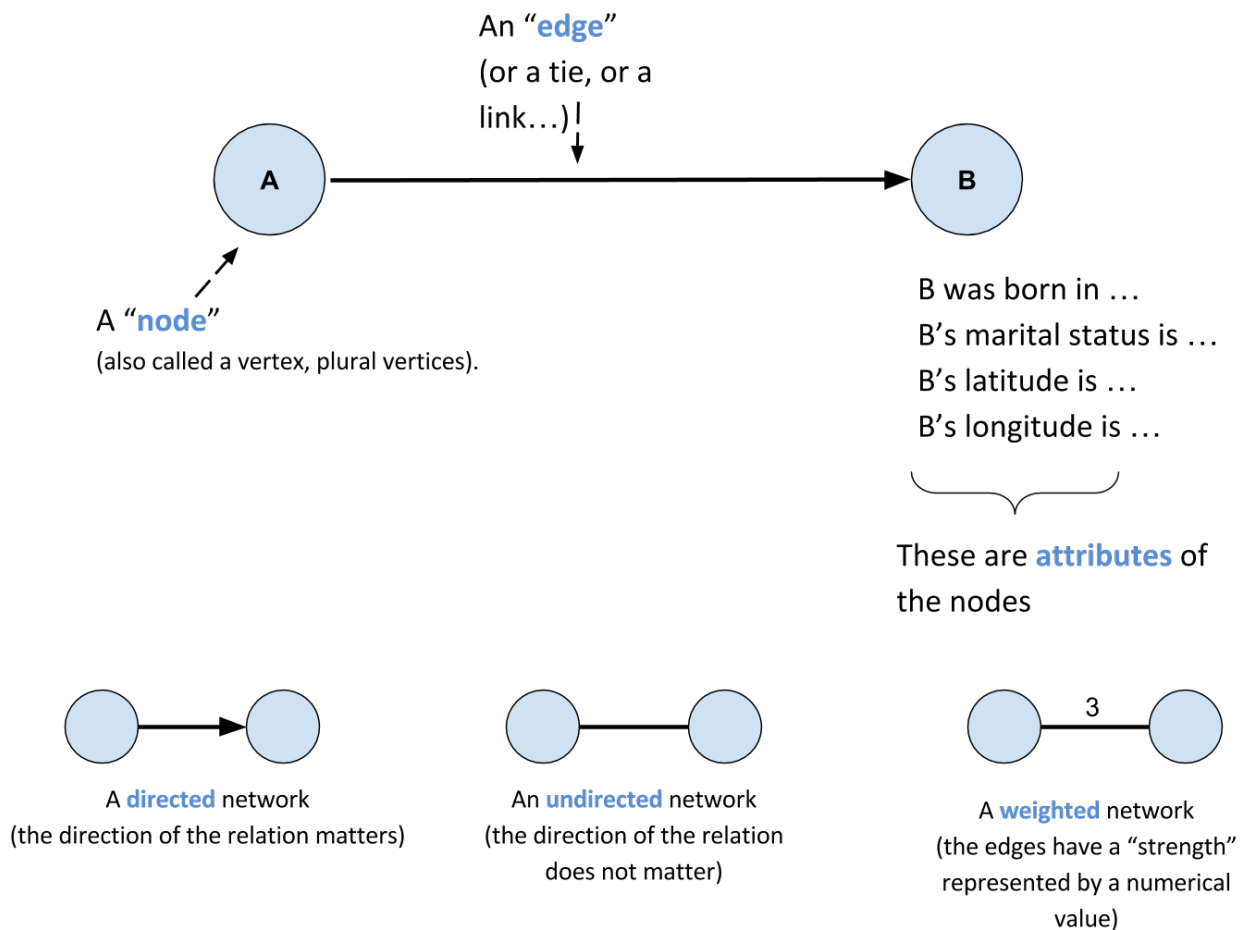


Figure 6. Terminology

## 2. Networks: what use for business?

### a. Segmentation

If a network is made of entities and their relations, then a segment is a subgroup of entities in the network, which has some cohesion or something in common.

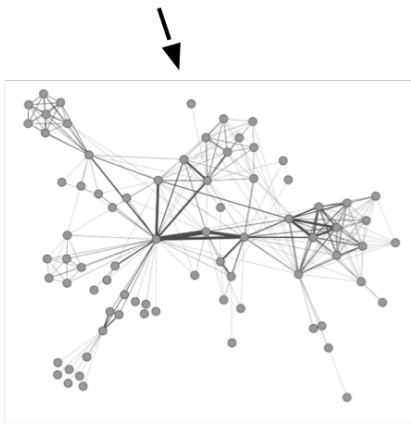
This subgroup of nodes in the network is often called a **"community"**.

Detecting communities in a network, also called "clustering", consists in finding nodes that have many connections in common.

This is a mathematical and algorithmic procedure, but it is very simple to understand visually:

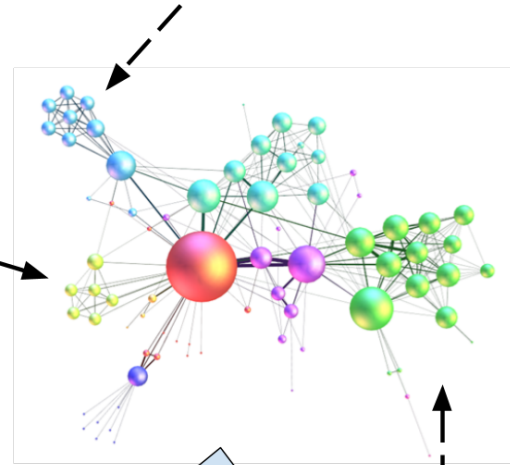
# SEGMENTATION

Car drivers and their similarities



Families with young children

Professionals who  
need pick-up trucks



Single men with sports  
tastes

Same network, with segments highlighted

Figure 7. segmentation with community detection in networks

## b. Finding key players



## FINDING KEY PLAYERS

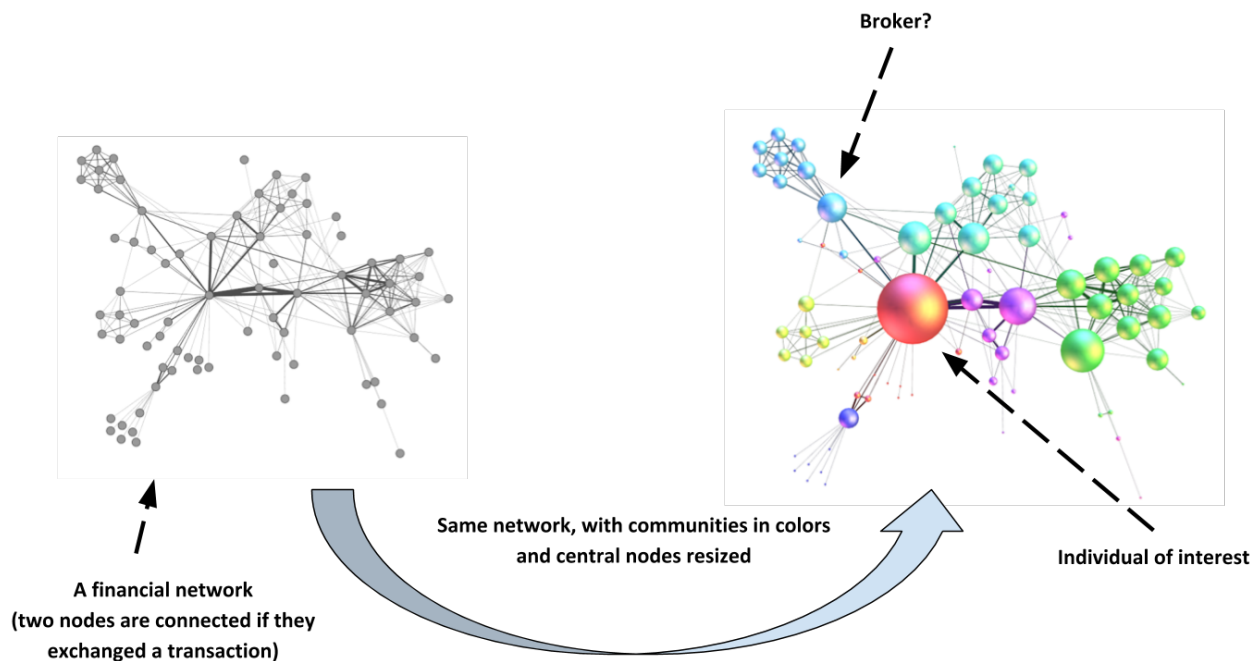


Figure 8. Key players visualized by resizing nodes

### c. Understanding how information spreads

A data science company created "Where does my tweet go", which traces how a given tweet spreads through retweets.

The service is now discontinued (Twitter data was too expensive to buy) but the mechanism can be explained:

# Understanding info spread

Initial tweet (appears at the center)

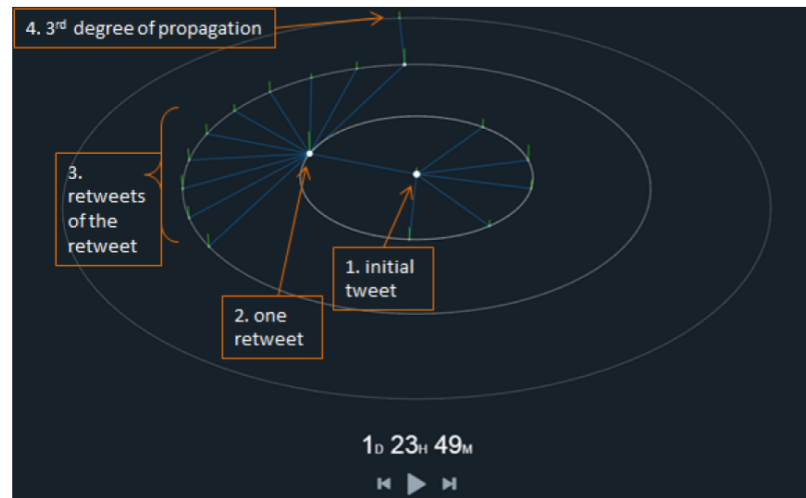
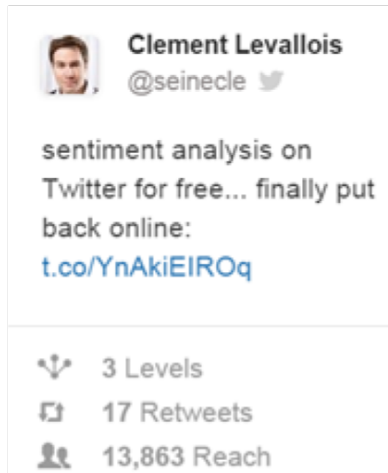


Figure 9. WDMTG by MFGLabs

## d. Identifying patterns - for fraud detection, control or intelligence.

In the following video, we see participants in the money market (short term loans between banks) in Europe.

2 banks are connected if one lends to the other. The pattern of exchanges shifts through years - banks withdraw from the market.

► <https://www.youtube.com/watch?v=YvauCrHGWYc> (YouTube video)

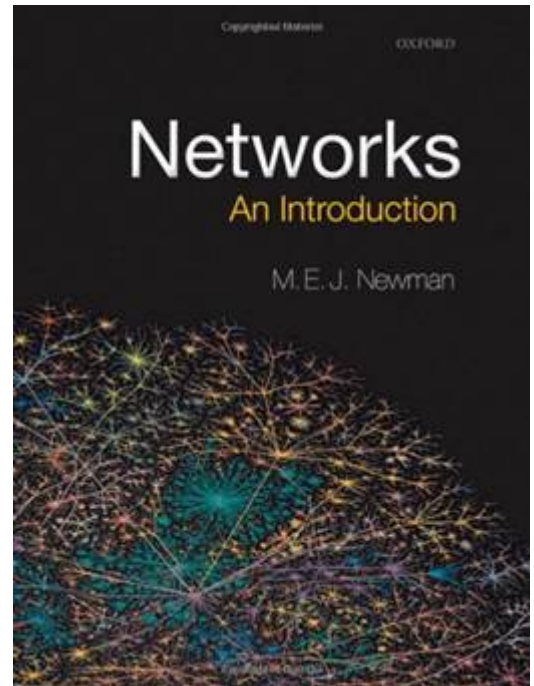
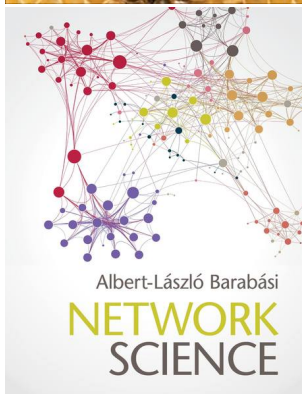
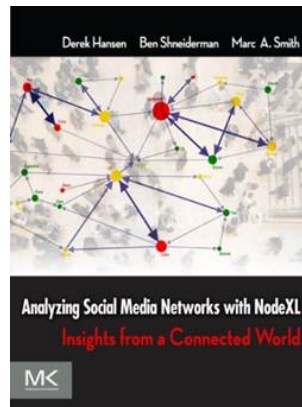
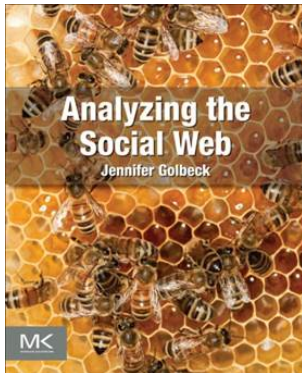
(the full study is available here: [https://www.dnb.nl/en/binaries/Working%20Paper%20418\\_tcm47-305800.pdf](https://www.dnb.nl/en/binaries/Working%20Paper%20418_tcm47-305800.pdf))

Another example: connecting seemingly unrelated measures of business performance with [Oracle BI](#) and [Linkurious](#):

► <https://www.youtube.com/watch?v=KBIZoUikfwo> (YouTube video)

## 3. To go further

(if viewing from a screen you can click on the covers to get to the Amazon page)



You can also visit my tutorials on Gephi, the leading software to visualize large graphs:

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

## The end

Find references for this lesson, and other lessons, [here](#).



This course is made by Clement Levallois.

Discover my other courses in data / tech for business: <http://www.clementlevallois.net>

Or get in touch via Twitter: [@seinecle](#)