

LES BIG DATA AU SECOURS DES GRANDES CATASTROPHES

Publié le 10 février 2016



par Elise Dubuisson

Qu'il s'agisse de la grippe, d'Ebola ou de rumeurs, pour réussir à les endiguer, il faut commencer par savoir à quelle vitesse ils se propagent. Mais de quoi faut-il tenir compte ? De la structure de la population concernée ? De ses habitudes ? De la fréquence des contacts entre les membres d'une même famille ? De tous ces paramètres à la fois ?

Voilà la question à laquelle a répondu [Jean-Charles Delvenne](#), professeur à l'[École Polytechnique de Louvain](#). À la clé : une meilleure analyse de la transmission des virus, rumeurs, opinions politiques, etc.

Pléthore de données

« Les nouveaux moyens de stockage et d'acquisition de données ont engendré les [Big Data](#) et avec elles une multitude d'informations détaillées sur une grande partie de l'activité des gens et donc d'une population. C'est bien entendu très intéressant car auparavant les analyses de transmission de données se voyaient bloquées par la disponibilité d'une quantité très limitée d'informations, ce qui n'est plus le cas. Mais cette masse d'informations a toutefois son talon d'Achille : comment analyser ces données rapidement et de manière efficace ? », détaille le chercheur.

Et c'est là qu'il intervient avec ses collègues [Luis Rocha](#) et [Renaud Lambiotte](#), chercheurs à [Namur](#).

Le réseau social et la dynamique des individus

Jusqu'à présent, lorsque une épidémie se déclarait, les mathématiciens chargés de prédire la propagation du virus ou de la bactérie se basaient essentiellement sur un modèle simple : au sein d'une ville ou d'un hôpital, la personne malade va rencontrer une personne et la contaminer, personne qui à son tour rencontrera une autre personne et ainsi de suite.

« Sauf que la manière dont se rencontrent les gens n'est pas identique pour tous les individus. Il faut

donc tenir compte de deux éléments essentiels : leur réseau social et leur dynamique sociale. Le réseau social qui indique la manière dont les individus sont liés et de nombreux moyens permettent de l'estimer, les données téléphoniques par exemple. [Cette information](#) est étudiée depuis une dizaine d'années dans les cercles académiques et permet notamment de savoir quels segments d'une population seront prioritairement touchés en cas d'épidémie. »

Ce réseau social peut être analysé à différentes échelles : une classe, une ville, un hôpital, un pays, un continent, etc. Mais savoir comment sont connectés les individus ne suffit pas, il faut également tenir compte de la fréquence à laquelle ils se rencontrent mais aussi de la régularité de ces rencontres : le temps entre deux contacts successifs est-il toujours le même ou bien imprédictible ?

« *Ce paramètre influence directement l'ordre et le timing de la transmission* », insiste le Pr Delvenne.

Une analyse rapide et efficace

Le problème lorsqu'on s'intéresse à un nombre plus important de paramètres et donc de données, c'est que le temps nécessaire à leur analyse est également plus important. Ce qui n'est pas forcément une avancée : face à une épidémie, il faut agir le plus vite possible ! [Raison pour laquelle Jean-Charles Delvenne et ses collègues ont mis au point une formule qui permet de tenir compte de ces deux paramètres de manière pertinente et rapide.](#)

« *Le réseau et la dynamique sociale apportent tous les deux des informations intéressantes mais nos expériences nous ont montré que très souvent l'un des paramètres écrase l'autre. C'est-à-dire que dans une situation donnée, c'est le réseau social qui fournit l'information la plus intéressante et dans une autre, c'est la dynamique sociale. Notre idée a donc été de mettre au point un algorithme mathématique permettant d'analyser les deux et de déterminer rapidement lequel des deux utiliser.* »

Ainsi, face à la transmission d'un virus au sein d'un hôpital on sait dorénavant qu'il vaut mieux tenir compte de la dynamique et donc de la fréquence des contacts entre individus. À l'inverse, si on étudie la transmission d'une maladie vénérienne, c'est plutôt le réseau social qui est déterminant. Deux cas qui ont pu être vérifiés sur le terrain par Jean-Charles Delvenne.

« *Nous avons étudié un scénario de transmission des maladies nosocomiales dans les hôpitaux et d'épidémie au sein d'un réseau de contacts sexuels au Brésil. Dans le premier cas, le réseau social intra-hospitalier est très régulier mais les temps de contact ne le sont pas. Dans le second, l'épidémie est complètement dominée par le découpage social très net du pays entre différentes villes et quartiers. L'avantage de notre approche est qu'elle combine simplicité et précision* », conclut le chercheur.