TIPE : Propagation de rumeurs dans un réseau social
Rapport

Hugo LEVY-FALK

2017

Table des matières

1 Préambule

Ce rapport synthétise le travail réalisé sur la propagation de rumeurs dans les réseaux sociaux. Les objectifs fixés lors de la mise en cohérence ont étés atteints, sauf en ce qui concerne l'élaboration d'un algorithme optimisant le choix des conditions initiales pour une meilleure propagation de la rumeur. Une approche théorique non prévue initialement a également été réalisée.

2 Introduction

Faisant suite à la recherche documentaire, une première approche théorique de la propagation des rumeurs, s'inspirant du livre de Easley et Kleinberg, a été réalisée. Cette approche permet de fixer le cadre des simulations ainsi que d'identifier les paramètres à faire varier lors des expériences. Dans un deuxième temps trois critères de choix des éléments propageant initialement la rumeurs ont été expérimentés en faisant varier les paramètres mis en évidence précédemment.

3 Propagation de rumeurs

3.1 Modalités d'action

3.1.1 Étude théorique

L'étude théorique a consisté en la formalisation du concept de rumeur et de propagation. On modélise naturellement un réseau social par un graphe $G = (V, E)$ avec V un ensemble de nœuds et $E \subset V^2$ que l'on supposera connexe par la suite. Chaque membre du réseaux est modélisés par un nœud du graphe et ses liens sociaux par une arrête.

Dans un premier temps, on cherche à caractériser le comportement d'un nœud du graphe. Pour cela on va considérer que le nœud peut se trouver dans deux états distincts : un premier dans lequel le nœud transmet la rumeur à tous ses voisins (A) et un autre dans lequel il ne relaie pas la rumeur (B). Pour chaque nœud $i \in V$ du graphe, si i est dans l'état A, alors pour chacun des ses voisins dans l'état A, i réalise un gain a. De même si i est dans l'état B il réalise un gain b pour chacun de ses voisins dans l'état B. On procède ensuite à une simulation dans laquelle à chaque étape, tous les nœuds du graphe tentent de maximiser leur gain en choisissant ou non de passer à l'état A (on n'autorise pas les passages de A vers B).

Lors d'une étape de simulation, si l'on note p la proportion de voisins dans l'état A du nœud i, alors le nœud i maximise son gain en passant à l'état A si $p \times a > (1 - p) \times b$, ou encore

$$p > \frac{b}{a + b} = q$$

Proposition 3.1. *Si $q > 1$, alors la rumeur ne peut pas se propager.*

Proposition 3.2. *Si l'on pose $(V_k)_{k\in\mathbb{N}}$ la suite des nœuds dans l'état A à l'étape k de la simulation, s'il existe $n \in \mathbb{N}$ tel que $V_n = V_{n+1}$ alors la suite est stationnaire à partir du rang n.*

Démonstration. En effet si $V_n = V_{n+1}$ alors chaque nœud maximise déjà son gain à l'étape n, la situation étant identique à l'étape $n + 1$, on a $V_{n+1} = V_{n+2}$. □

Proposition 3.3. *On pose $n = |V|$. La suite $(V_k)_{k\in\mathbb{N}}$ converge en n étape au plus.*

Démonstration. La suite $(V_k)_{k\in\mathbb{N}}$ étant croissante pour l'inclusion et majorée par V, elle converge. D'autre part, d'après la proposition 3.2, cette suite est stationnaire à partir du rang $k \in \mathbb{N}$ si $V_k = V_{k+1}$. Or il n'y a que n nœuds dans le graphe, donc la suite est stationnaire à partir du rang n au plus. □

La proposition 3.3 permet d'assurer que toutes les simulations pourront être menées à bout en un temps fini. La proposition 3.2 permet de raccourcir éventuellement la durée d'une simulation en comptant le nombre d'éléments qui passent à l'état A à chaque étape de simulation. S'il est nul, alors la simulation est terminée.

On peut ensuite chercher un critère plus complet que la proposition 3.1 pour qualifier la possibilité de propager une rumeur.

Définition 3.1. On appelle p-cluster tout sous-ensemble $C \subset V$ tel que pour tout $i \in C$ il existe un p-uplet $(v_k)_{k\in[\![1,p]\!]} \in C^p$ deux à deux distincts et tel que pour tout $k \in [\![1,p]\!]$, i et v_k soient voisins.

Théorème 3.1. *Une rumeur de note q et non initialement propagée à tout le graphe ne se propagera pas à l'ensemble du réseau si et seulement si il existe un p-cluster avec $p > q$.*

Démonstration. On pose $n = |V|$. S'il existe un p-cluster C avec $p > q$, alors tout nœud de C possède au moins une proportion p de voisins non informés. Ceci valant pour tous les nœuds de C, aucun nœud de C ne sera informé au bout de n étapes. *Les clusters sont donc des obstacles aux rumeurs.*

S'il existe un nœud $i \in V$ tel qu'au bout de n étapes i ne soit pas dans l'état informé, alors la proportion p de voisins de i dans l'état informé vérifie $p \leq q$ ou encore $(1-p) > q \leq 0$. Il existe donc des voisins de i vérifiant cette propriété, *on a un z-cluster avec $z > q$.* □

3.1.2 Génération des graphes

Les graphes sont générés avec l'algorithme de Watts-Strogatz.

```
Données : N ∈ ℕ, K ∈ ⟦1, ⌊N/2⌋⟧ (N ≫ K ≫ ln N), β ∈ [0,1]
Résultat : Matrice d'adjacence d'un graphe aléatoire.
M ← matrice avec pour i ∈ ⟦0, N−1⟧, j ∈ ⟦1, K⟧, M_{i,i+j[N]} = M_{i,i−j[N]} = Vrai, Faux
  pour les autres ;
pour i ∈ ⟦0, N−1⟧ faire
  pour j ∈ ⟦1, K⟧ faire
    r ← Nombre aléatoire sur [0,1] ;
    si r < β alors
      M_{i,i+j[N]} ← Faux ;
      M_{i+j[N],i} ← Faux ;
      Choisir au hasard k tel que M_{i,k} = Faux ;
      M_{i,k} ← Vrai ;
      M_{k,i} ← Vrai ;
    finsi
  finpour
finpour
retourner M
```

Algorithm 1 : Algorithme de Watts-Strogatz

3.2 Résultats

3.3 Analyse, exploitation et discussion

4 Conclusion