

Une approche multidimensionnelle basée sur les comportements individuels pour la prédiction de la diffusion de l'information sur Twitter

Adrien Guille*, Hakim Hacid**, Cécile Favre*

*ERIC, Université Lumière Lyon 2, Lyon, France
{prénom.nom}@univ-lyon2.fr

**Bell Labs France, Alcatel-Lucent, Nozay, France
hakim.hacid@alcatel-lucent.com

Résumé

Aujourd'hui, les réseaux sociaux en ligne sont devenus des outils très puissants de propagation de l'information. Ils favorisent la diffusion rapide à grande échelle de contenu et les conséquences d'une information inexacte voire fausse peuvent alors prendre une ampleur considérable. Par conséquent il devient indispensable de proposer des moyens d'analyser le phénomène de diffusion de l'information dans ces réseaux. De nombreuses études récentes ont traité de la modélisation du processus de diffusion de l'information, essentiellement d'un point de vue topologique et dans une perspective théorique, mais les facteurs impliqués sont encore méconnus. Nous proposons ici une solution pratique dont l'objectif est de prédire la dynamique temporelle de la diffusion au sein de Twitter, basée sur des techniques d'apprentissage automatique. Notre approche repose sur l'inférence de probabilités de diffusion tirées d'une analyse multidimensionnelle des comportements individuels. Les expérimentations menées montrent l'intérêt de la modélisation proposée.

Summary

Online social networks facilitate the rapid and large-scale propagation of content and the consequences of an inaccurate or false information can then take considerable proportions. Therefore it is essential to provide means to analyze the phenomenon of information dissemination in such networks. In this paper we propose a practical solution which aims to predict the temporal dynamics of diffusion in Twitter, based on machine learning techniques. Experimental results on real data show the interest of the proposed approach.