

Conception d'outils pour analyser et visualiser la diffusion des mèmes Internet

Le réseau social chinois Sina Weibo

Renaud Clément
LTCI SES, ParisTech Telecom

clement.renaud@gmail.com

RÉSUMÉ

Nous proposons de concevoir et développer un outil permettant d'analyser la diffusion d'information sur les services de réseaux sociaux en ligne grâce au traitement et à la visualisation de données. Fruit d'une réflexion méthodologique sur l'étude des actes de communications en ligne, ce dispositif permet d'observer les relations entre leurs dimensions conversationnelles, sémantiques, temporelles et géographiques.

Courts messages se propageant rapidement sur la Toile, les *mèmes Internet* comptent parmi les contenus les plus prisés du web. Sur le service de microblog chinois *Sina Weibo* notamment, discussions personnelles, débats sociétaux et vastes campagnes médiatiques s'articulent quotidiennement dans des *mèmes* aux modèles encore mal connus.

Mobilisant des méthodes issues de l'analyse des réseaux et du traitement automatisé de la langue chinoise, nous procédons à l'analyse d'un vaste corpus de 200 millions de messages représentant l'activité sur *Sina Weibo* durant l'année 2012 (Fu & Chau, 2013).

Notre première tâche consiste à identifier dans cet ensemble de données des mèmes. Leur identification dans un corpus de messages est notamment possible grâce à un algorithme de détection non supervisé (Ferrara, 2013). Néanmoins, la quantité de calculs nécessaires pour obtenir des résultats fiables sur un tel volume de discussions nous amène à abandonner cette approche, montrant par là-même la complexité d'une définition intéressante de l'objet numérique *mème*.

Notre seconde série d'analyse porte sur le volume de conversations entourant les *hashtags* mentionnés. Les résultats montrent que les usages majoritaires de *Sina Weibo* sont similaires à ceux des médias traditionnels (publicité, divertissement, loisirs...). Néanmoins, nous écartons les *hashtags* comme représentatifs des mèmes Internet, artefacts d'usages commerciaux et stratégiques à la diffusion souvent cadrée et planifiée.

L'approche finalement retenue utilise la recherche par mots-clés pour constituer des corpus autour d'une dizaine de mèmes sélectionnés dans la littérature académique et secondaire pour leurs intentions diverses: humour, actualité, scandale politique, faits divers et marketing promotionnel. S'inspirant de la critique des schémas théoriques de la communication, une analyse des mots et des réseaux d'échanges entre utilisateurs met à jour les dynamiques discursives de chaque mème. L'organisation de ces informations selon un axe temporel dans un espace de visualisation interactif offre une lecture détaillée de leur diffusion. La projection de ces réseaux conversationnels et lexicaux sur des cartes géographiques montre également les relations entre leurs aspects textuels et actuels.

Les figures obtenues permettent d'ébaucher une typologie structurelle de la diffusion des mèmes et de formuler ainsi des recommandations pour l'analyse et la conception de stratégies de communication en ligne d'organismes tant privés que publics. Néanmoins, le caractère exploratoire de cette étude et la difficulté de comprendre les actions humaines par une simple analyse de données nous invite à refuser une généralisation a priori des résultats. Nous préférons considérer ce travail comme la première validation d'une méthodologie pouvant être étendue à d'autres formes de conversations en ligne.

À Lyon,
le 29 Juin 2014.