



HAL
open science

Fouille de données hétérogènes pour la caractérisation de rumeurs et fake news : analyse empirique croisée de la presse en ligne et Twitter

Matthieu Bachelot, Inna Lyubareva, Thomas Epalle, Romain Billot,
Raphaël-David Lasserri

► To cite this version:

Matthieu Bachelot, Inna Lyubareva, Thomas Epalle, Romain Billot, Raphaël-David Lasserri. Fouille de données hétérogènes pour la caractérisation de rumeurs et fake news : analyse empirique croisée de la presse en ligne et Twitter. ROADEF 2023, Feb 2023, Rennes, France. hal-04221824

HAL Id: hal-04221824

<https://hal.science/hal-04221824v1>

Submitted on 28 Sep 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Fouille de données hétérogènes pour la caractérisation de rumeurs et fake news : analyse empirique croisée de la presse en ligne et Twitter

Matthieu Bachelot¹, Inna Lyubareva¹, Thomas Epalle², Romain Billot¹,
Raphaël-David Lasserri²

¹ IMT Atlantique, DECIDE, France

{prenom.nom}@imt-atlantique.fr

² Magic LEMP, France

{nom}@magic-lemp.com

Mots-clés : *informations fallacieuses, rumeurs, analyse des réseaux sociaux, graphes, propagation, fouille de données, visualisation*

1 Introduction

Les réseaux sociaux jouent un rôle significatif dans la transmission de l'information, notamment de par leur facilité d'accès sur plusieurs supports et leur ergonomie. Cependant, l'abondance et la diversité des points de vues permises par le contexte numérique ne sont pas toujours synonymes d'une meilleure qualité de l'information en ligne. Parmi les problèmes récents, la prolifération des rumeurs ou fausses informations (fake news) constitue un enjeu majeur pour l'ensemble des acteurs économiques du secteur, mais aussi plus largement en matière de sérénité du débat public voire de sécurité nationale. Le problème de la détection des informations fallacieuses a été étudié depuis longtemps [1], mais les approches se basant uniquement sur les contenus deviennent insuffisantes dans le contexte actuel [2]. Des techniques utilisant d'autres leviers ont été développées pour essayer de détecter ces fakes news de manière précoce, en y joignant par exemple des profilages d'utilisateurs [3] ou bien des analyses temporelles [4]. Cependant ces méthodes sont encore perfectible et surtout se fondent trop souvent sur un nombre réduits de jeu de données de références (Weibo, Twitter15, Twitter16 ...). Le corollaire est qu'il n'y a très peu de travaux de recherches dans la détection d'informations fallacieuses sur un jeu de données en français, quel que soit le support (liste d'articles, base de données de tweets). Notre travail vise ainsi à combler ces lacunes et à proposer une des premières analyses empiriques en langue française, en croisant plusieurs sources de données.

2 Méthode

L'originalité de notre approche réside dans la proposition d'une méthodologie de modélisation d'informations fallacieuses qui utilise plusieurs support de propagations : d'un côté, des articles de presse issues des sites web des principaux médias français, et de l'autre la base de données du réseau social Twitter. Cette approche inédite présente plusieurs avantages, à commencer par une prise en compte de différents acteurs présents lors de la propagation de l'information en général, c'est-à-dire à la fois les producteurs d'information (éditeurs, journalistes) ainsi que les consommateurs d'information (utilisateurs des réseaux sociaux). Plus précisément, nous avons sélectionné plusieurs événements pouvant s'apparenter à des rumeurs ou fake news ainsi que d'autres correspondant à des événements normaux (vraies informations)

dans une base de données contenant plus d'un millions d'articles et un demi million de tweet récoltés depuis Novembre 2021. Cette approche par évènement est d'autant plus intéressante car les résultats de l'analyse seront plus facilement généralisables à différents types d'information à risque quelque soit le type de producteurs d'informations. Après deux modélisations parallèles par graphes, ce double niveau de granularité nous a permis de mettre en évidence les marqueurs d'ordre structurel de la propagation d'information fallacieuse également présent dans l'analyse de données anglophones [5], mais aussi, grâce à une analyse temporelle de la propagation de l'information dans les deux graphes, de mettre en évidence de nouvelles caractéristiques propres aux informations fallacieuses.

3 Premiers résultats

Nous avons modélisé les évènements par deux graphes, un par support étudié (Twitter et articles). Pour les données Twitter, le graphe a été reconstitué à partir des retweets tandis que les arêtes du graphe de la base de données des articles s'appuient sur la notion de proximité sémantique des contenus. Après avoir sélectionné un échantillon d'articles traitant de l'évènement, nous avons testé plusieurs méthodes d'embedding des contenus dans l'optique de calculer leurs similarité deux à deux dans l'espace latent. Ces calculs de similarité ont permis de générer une matrice d'adjacence pour pouvoir créer le graphe de propagation de l'information dans la presse écrite. Dans le cadre des études temporelles, nous avons analysé les délais de propagation ainsi que la décomposition de graphes en snapshots. Les délais de propagations nous ont permis de mettre en évidence des tendances globales et individuelles vis à vis de la propagation d'informations, que ce soit pour la presse écrite ou pour Twitter. En ce qui concerne les snapshots, nous avons analysé les tendances globales des tweets (nombre de tweets en fonction du temps) pour choisir automatiquement le nombre de périodes d'intérêt via des algorithmes de classification non supervisée. Nous avons alors reconstruit les graphes de la même manière mais en isolant les sommets dans les périodes temporelles d'intérêt pour mettre en évidence les différences structurelles de la propagation de l'information. Les premiers résultats empiriques présentés dans ce travail sont prometteurs et permettent de décliner un certain nombre de propriétés de graphe inhérentes aux informations fallacieuses, en se fondant sur les métriques classiques de l'analyse des réseaux sociaux et des analyses statistiques uni- et multidimensionnelles.

Références

- [1] Tiantian Qin, J.K. Burgoon, J.P. Blair, and J.F. Nunamaker. Modality Effects in Deception Detection and Applications in Automatic-Deception-Detection. In *Proceedings of the 38th Annual Hawaii International Conference on System Sciences*, pages 23b–23b, Big Island, HI, USA, 2005. IEEE.
- [2] Kai Shu, Suhang Wang, and Huan Liu. Beyond News Contents : The Role of Social Context for Fake News Detection. In *Proceedings of the Twelfth ACM International Conference on Web Search and Data Mining*, pages 312–320, Melbourne VIC Australia, January 2019. ACM.
- [3] Yi-Ju Lu and Cheng-Te Li. GCAN : Graph-aware Co-Attention Networks for Explainable Fake News Detection on Social Media, April 2020. Number : arXiv :2004.11648 arXiv :2004.11648 [cs, stat].
- [4] Natali Ruchansky, Sungyong Seo, and Yan Liu. CSI : A Hybrid Deep Model for Fake News Detection. In *Proceedings of the 2017 ACM on Conference on Information and Knowledge Management*, pages 797–806, Singapore Singapore, November 2017. ACM.
- [5] Amirhosein Bodaghi and Jonice Oliveira. The theater of fake news spreading, who plays which role? A study on real graphs of spreading on Twitter. *Expert Systems with Applications*, 189 :116110, March 2022.