

Fouille de données hétérogènes pour la caractérisation de 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, 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 fake 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 perfectibles et surtout se fondent trop souvent sur un nombre réduit de jeu de données de référence (Weibo, Twitter15, Twitter16 ...). Le corollaire est qu'il n'y a que 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 Méthodologie

2.1 Données

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 issus des sites web des principaux médias français, et de l'autre la base de données du réseau social Twitter. Nous avons sélectionné plusieurs événements (ie. un ensemble cohérent d'articles et de tweets autour d'un même sujet) 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 événement est d'autant plus intéressante que les résultats de l'analyse seront plus facilement généralisables.

2.2 Modélisation et analyse

Nous avons modélisé les événements 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. Une fois les graphes obtenus nous avons voulu mettre en évidence les marqueurs d'ordre structurel de la propagation d'information fallacieuse également présents dans des analyses similaires faites sur des données anglophones [5]. Nous avons couplé cette étude avec une analyse temporelle dans les deux graphes, afin de mettre en évidence de nouvelles caractéristiques propres aux informations fallacieuses. Cette analyse temporelle se divise en deux parties : une analyse des délais de propagation et une analyse des sous-graphes représentant les différentes phases de la propagation de l'information. Pour générer ces sous-graphes, nous avons analysé les tendances globales des tweets (respectivement des articles) pour déterminer le nombre de périodes d'intérêt. Nous avons alors reconstruit les graphes en isolant les sommets dans ces périodes, et nous avons analysé leurs structure.

3 Premiers résultats

Dans le cadre de l'étude sur les différences structurelles, nous avons dégagé des tendances distinctes entre les événements vrai et fake qui confirment certaines tendances décrites dans l'état de l'art. Dans le cadre des études temporelles, les délais de propagations nous ont permis de mettre en évidence des différences significatives vis à vis de la propagation d'informations, que ce soit pour la presse écrite ou pour Twitter. Nous avons également étudié le cas de la détection précoce, en enlevant tous les articles et les tweets publiés après un certain temps après le début de la rumeur. De manière générale 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.