InfoTrace: A System for Information Campaign Source Tracing and Analysis on Social Media

Linus Xin Wei Cheng
Singapore Uni. of Tech. and Design
Singapore

Daniel Wai Kit Chin
Singapore Uni. of Tech. and Design
Singapore

Shaun Toh
Singapore Uni. of Tech. and Design
Singapore

Wenchuan Mu
Singapore Uni. of Tech. and Design
Singapore

James Kay Liang Ong
Singapore Uni. of Tech. and Design
Singapore

Kenny Tsu Wei Choo
Singapore Uni. of Tech. and Design
Singapore

Roy Ka-Wei Lee
Singapore Uni. of Tech. and Design
Singapore

Kwan Hui Lim
Singapore Uni. of Tech. and Design
Singapore

Abstract

Social media platforms have become integral to daily life and serve as powerful channels for influencing public opinion, spanning applications from viral marketing to disinformation campaigns. While these platforms amplify the reach of marketing efforts, they also present significant risks when used for spreading misinformation via campaigns. Thus, it is crucial to understand such information campaigns by identifying the source, the different discussion topics within this campaign and how they change over time. Towards addressing this problem, we develop and present an interactive system for information campaign source tracing and analysis. This includes a demonstration and visualization of the main components of a information campaign, including the source of the campaign via explicit and implicit links, the discussion topics/clusters, their content, and how these evolve over different time periods.

CCS Concepts

• **Computing methodologies → Natural language processing**; **Neural networks**; • **Information systems → Clustering and classification**; **Data encoding and canonicalization**.

Keywords

Information Campaigns, Social Computing, Disinformation Detection, Twitter, Social Media

ACM Reference Format:
Linus Xin Wei Cheng, Daniel Wai Kit Chin, Shaun Toh, Wenchuan Mu, James Kay Liang Ong, Kenny Tsu Wei Choo, Roy Ka-Wei Lee, and Kwan Hui Lim. 2024. InfoTrace: A System for Information Campaign Source Tracing and Analysis on Social Media. In *The 2024 ACM/IEEE Joint Conference on Digital Libraries (JCDL '24), December 16–20, 2024, Hong Kong, China.* ACM, New York, NY, USA, 3 pages. https://doi.org/10.1145/3677389.3702583

JCDL '24, December 16–20, 2024, Hong Kong, China

ACM ISBN 979-8-4007-1093-3/24/12
https://doi.org/10.1145/3677389.3702583

1 Introduction

Online social media platforms, such as Twitter, Facebook, TikTok, etc, are now heavily embedded in our daily lives, with an average of 2 hours 24 minutes spent on these platforms daily [6]. Social media has also spread across more than half of the world's population with 4.9 billion users in 2023 and is further projected to grow to 5.85 billion users by 2027. Social media has also emerged as a major source of information, with a wide coverage of topics ranging from products/services recommendations to geopolitical news. This effect is evident in how 76% of users have reportedly purchased a product or service that they initially saw on social media [6].

Due to the effectiveness of social media, it is also highly susceptible to issues such as information campaigns to further the agenda of an individual or group [8, 15]. For example, anti-government organizations may utilize disinformation campaigns to influence public opinions and cause social unrest. Another aggravating factor is how people are easily misled in such information campaigns, with research highlighting that "fake news headlines fool American adults about 75% of the time" [5].

Contributions. To address these challenges, we present InfoTrace, a system designed to facilitate the understanding of information campaigns on social media. InfoTrace enables: (i) tracing the sources of an information campaign by identifying its initiator(s); (ii) unsupervised clustering of campaign content to detect discussion topics based on textual and temporal patterns; (iii) text-based classification of individual social media posts to assess attributes, such as the type of information; and (iv) visualization tools that illustrate these facets of a campaign across clusters and time intervals. While InfoTrace is broadly applicable to diverse types of information campaigns, including fake news, marketing promotions and viral complaints on social media, we demonstrate its capabilities through a case study on tracing disinformation campaigns.

2 InfoTrace System Architecture

Our InfoTrace system comprises four main components, namely:

Data Collection and Pre-processing Component: This component collects and loads social media posts, including those belonging to information campaigns, along with general posts of routine conversations. The dataset is pre-processed through steps, such

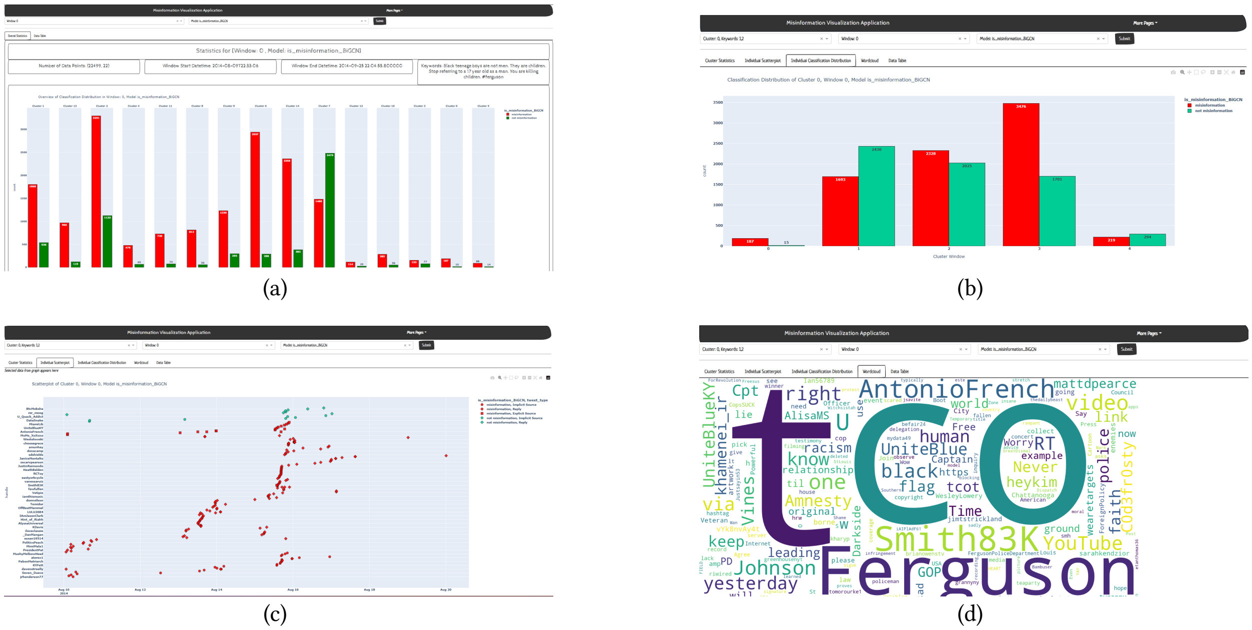

Figure 1: Screenshots of our proposed InfoTrace system: (a) Overview of Detected Clusters; (b) Cluster Change over Time; (c) Link Types and Post Types over Time; (d) Wordcloud

as data cleaning and normalization, preparing it for subsequent campaign source tracing and information type detection.

Campaign Source Tracing: This component aims to identify the source, i.e., an initiator or set of initiators, of an information campaign, starting with an initial query post of interest. Using textual similarity and temporal proximity derived from embeddings [11, 12], it clusters the social media posts into groups to reveal the main discussion topics.

Social Post Classification: This component classifies individual social media posts into different information categories. We evaluated various text-based models, including BERT [3] and its variants S-BERT and IS-BERT, which enhance sentence representation [14]. In addition, we experimented with network-based models such as BiGCN [1] and EBGCN [16], which utilize conversation structure to improve classification accuracy.

Visualisation: This component provides an interactive interface for visualizing aspects such as campaign sources, topic clusters, and classifications over time. We further illustrate specific visualizations through a user scenario, showcasing its application in analyzing information campaigns on platforms like X/Twitter.

3 Use Case: Information Campaign Analysis

After collecting and loading a specific dataset, the user can assess an overview of the different detected clusters (Figure 1a), along with its summary statistics. On the x-axis, clusters are displayed according to our dual-clustering algorithm, which groups posts based on textual similarity and temporal proximity, reflecting distinct discussion topics. Key information includes the social media post distribution across clusters, selected time windows, and the proportion of disinformation versus non-disinformation posts, as identified by various algorithms, alongside other cluster statistics.

To investigate a specific cluster, users can view a detailed visualization of disinformation and regular posts within the cluster over time (Figure 1c). Node shapes represent different types of links used in posts (e.g., mentions, replies, etc), while node colors indicate individual posts and whether they have been classified as disinformation based on the implemented classification algorithms.

For temporal analysis, the system provides a chart showing the volume of posts and the ratio of disinformation to regular posts within the cluster across different time periods (Figure 1b). Additionally, a word cloud offers a high-level summary of the discussion content within each cluster, as shown in Figure 1d.

The user can then focus on specific subsets of posts within a cluster, such as those originating from a query post of interest, to trace the source of a disinformation campaign. Relevant posts and their details are displayed in a data table. The user is able to identify source posts using either explicit links (based on reply or mention links) or implicit links (based on an adjacent thread that is similar in terms of topic and time).

4 Conclusion

This paper proposes InfoTrace, an interactive system for campaign source tracing and analysis, using a series of text and temporal similarity clustering for source tracing, text-based classification for disinformation detection, and data analytics and visualization techniques to better understand these aspects. While there has been various systems developed for studying social media from the perspectives of disinformation detection [7, 13], sentiment analysis [2, 9], health interventions [4, 10] and other useful applications, there are limited systems on campaign source tracing and understanding, which our proposed InfoTrace aims to address.

Acknowledgments

This research is supported by the Ministry of Education, Singapore, under its Academic Research Fund Tier 2 (Award No. MOE-T2EP20123-0015). Any opinions, findings and conclusions, or recommendations expressed in this material are those of the authors and do not reflect the views of the Ministry of Education, Singapore. W. Mu and K. H. Lim are also supported by a CoT R&D Programme (Award No. COT-V2-2020-1).

References

[1] Tian Bian, Xi Xiao, Tingyang Xu, Peilin Zhao, Wenbing Huang, Yu Rong, and Junzhou Huang. 2020. Rumor detection on social media with bi-directional graph convolutional networks. In *Proceedings of the AAAI conference on artificial intelligence*, Vol. 34. 549–556.

[2] Jose Camacho-Collados, Kiamehr Rezaee, Talayeh Riahi, Asahi Ushio, Daniel Loureiro, Dimosthenis Antypas, Joanne Boisson, Luis Espinosa Anke, Fangyu Liu, and Eugenio Martínez-Cámara. 2022. TweetNLP: Cutting-Edge Natural Language Processing for Social Media. In *Proceedings of the 2022 Conference on Empirical Methods in Natural Language Processing: System Demonstrations*. 38–49.

[3] Jacob Devlin, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova. 2019. BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. In *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers)*. 4171–4186.

[4] Sarah M Edney, Tim S Olds, Jillian C Ryan, Corneel Vandelanotte, Ronald C Plotnikoff, Rachel G Curtis, and Carol A Maher. 2020. A social networking and gamified app to increase physical activity: cluster RCT. *American journal of preventive medicine* 58, 2 (2020), e51–e62.

[5] Emilio Ferrara, Onur Varol, Clayton Davis, Filippo Menczer, and Alessandro Flammini. 2016. The rise of social bots. *Commun. ACM* 59, 7 (2016), 96–104.

[6] Forbes. 2023. Top Social Media Statistics And Trends Of 2024. Internet. Available from: https://www.forbes.com/advisor/business/social-media-statistics/.

[7] Dongqi Fu, Yikun Ban, Hanghang Tong, Ross Maciejewski, and Jingrui He. 2022. DISCO: Comprehensive and explainable disinformation detection. In *Proceedings of the 31st ACM International Conference on Information & Knowledge Management*. 4848–4852.

[8] Christian Janze and Marten Risius. 2017. Automatic Detection of Fake News on Social Media Platforms. In *Proceedings of the 21st Pacific Asia Conference on Information Systems*.

[9] Jolin Shaynn-Ly Kwan and Kwan Hui Lim. 2021. Tweetcovid: a system for analyzing public sentiments and discussions about covid-19 via twitter activities. In *Companion Proceedings of the 26th International Conference on Intelligent User Interfaces*. 58–60.

[10] Chenhao Lin, Pengwei Hu, Hui Su, Shaochun Li, Jing Mei, Jie Zhou, and Henry Leung. 2020. Sensemood: depression detection on social media. In *Proceedings of the 2020 international conference on multimedia retrieval*. 407–411.

[11] Peter Mathews, Caitlin Gray, Lewis Mitchell, Giang T Nguyen, and Nigel G Bean. 2018. SMERC: Social media event response clustering using textual and temporal information. In *Proceedings of the 2018 IEEE International Conference on Big Data (BigData'18)*. 3695–3700.

[12] Wenchuan Mu and Kwan Hui Lim. 2023. Modelling Text Similarity: A Survey. In *Proceedings of the 2023 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM'23)*. 698–705.

[13] Dan S Nielsen and Ryan McConville. 2022. Mumin: A large-scale multilingual multimodal fact-checked misinformation social network dataset. In *Proceedings of the 45th international ACM SIGIR conference on research and development in information retrieval*. 3141–3153.

[14] Nils Reimers and Iryna Gurevych. 2019. Sentence-BERT: Sentence Embeddings using Siamese BERT-Networks. In *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing*. Association for Computational Linguistics. https://arxiv.org/abs/1908.10084

[15] Kai Shu, Amy Sliva, Suhang Wang, Jiliang Tang, and Huan Liu. 2017. Fake news detection on social media: A data mining perspective. *ACM SIGKDD Explorations Newsletter* 19, 1 (2017), 22–36.

[16] Lingwei Wei, Dou Hu, Wei Zhou, Zhaojuan Yue, and Songlin Hu. 2021. Towards Propagation Uncertainty: Edge-enhanced Bayesian Graph Convolutional Networks for Rumor Detection. In *Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing (Volume 1: Long Papers)*. 3845–3854.