26th International Congress on Modelling and Simulation, Adelaide, SA, Australia, 30 Nov–4 Dec 2025
mssanz.org.au/modsim2025

Modelling the Evolution of Topics in Social Media Conversations

Scott Carnie-Bronca[a]

[a]School of Computer and Mathematical Sciences, University of Adelaide, 5005, Australia
Email: scott.carnie-bronca@adelaide.edu.au

Abstract: With the ever-growing popularity of social media over the last decades, mis/dis-information has become common and widespread online. In addition, the advent of non-human actors on social media such as bots and LLM-based agents, has allowed misinformation campaigns to reach an unprecedented scale. While a majority of misinformation may be relatively benign, it can also be exploited in a contemporary warfare scenario such as those in Ukraine and the Middle East. Hence, it is critical that disinformation campaigns are identified in real time, in such rapidly evolving environments to allow for effective countermeasures. However, traditional narrative detection models may be less effective in these scenarios due to a lack of ground truth.

In order to efficiently investigate the prevalence of misinformation and coordinated activity, we take an unsupervised approach to extract potential conversations of interest, irrespective of truth. Topic modelling is a natural language processing technique that extracts unique *topics* - or collections of related ideas - from a set of documents. We hypothesise that such modelling can extract and cluster groups of social media discussions in a meaningful way, allowing us to aggregate and investigate posts in a rigorous and efficient way.

Many topic modelling pipelines exist, from the traditional LDA and LSA techniques (e.g. Blei et al. (2003)) to LLM-based topic modelling. We choose to use BERTopic, introduced by Grootendorst (2022), due to its scalability and high performance compared to LDA.

Topic modelling is not an exact science, and depending on the model results can vary significantly, even between runs. To reduce the impact of this and provide more reliable outcomes, we introduce a pair of coherence measures based broadly on information entropy, that allow us to measure both the quality of the resulting topics and the variance in topics between runs. While we present these in the context of BERTopic, the same measures can be used for other topic models due to their independence from the intermediate results.

While Grootendorst introduces a method of incorporating temporal information into BERTopic, this method has a major drawback in that it assumes that the set of underlying topics present in the text does not significantly change over time. This is not well-suited to data from social media, as conversations can rapidly evolve depending on external events, and this evolution process is critical to tracking misinformation. Hence, we also use the coherence metrics to measure the similarity of topics through discrete time, to track the evolution and spread of discussions and potential mis/dis-information campaigns.

We demonstrate the aforementioned topic modelling process using a novel dataset collected from Twitter, related to the Russian invasion of Ukraine. This dataset relates to two key events in the recent invasion - first, the coastal city of Mariupol, which was captured early in the invasion. Secondly, posts related to the hydroelectric power plant in Nova Kakhovka, which was destroyed over a year into the invasion, were collected from the months leading up to and weeks after the explosion. These data provide an opportunity to analyse the effectiveness of our model in tracking the spread of misinformation and coordinated activity in a real-world application.

Keywords: *Topic modelling, natural language processing, social media analysis, large language models, machine learning*

REFERENCES

Blei David M, Ng Andrew Y and Jordan Michael I (2003), Latent dirichlet allocation.
Grootendorst Maarten (2022), Bertopic: Neural topic modeling with a class-based tf-idf procedure.
URL: *https://arxiv.org/abs/2203.05794*

