

Content-based features predict social media influence operations | Science Advances

By Meysam Alizadeh, Jacob N. Shapiro, Cody Buntain & Joshua A. Tucker

July 24, 2020

We study how easy it is to distinguish influence operations from organic social media activity by assessing the performance of a platform-agnostic machine learning approach. Our method uses public activity to detect content that is part of coordinated influence operations based on human-interpretable features derived solely from content. We test this method on publicly available Twitter data on Chinese, Russian, and Venezuelan troll activity targeting the United States, as well as the Reddit dataset of Russian influence efforts. To assess how well content-based features distinguish these influence operations from random samples of general and political American users, we train and test classifiers on a monthly basis for each campaign across five prediction tasks. Content-based features perform well across period, country, platform, and prediction task. Industrialized production of influence campaign content leaves a distinctive signal in user-generated content that allows tracking of campaigns from month to month and across different accounts.

[...]

Source: [Content-based features predict social media influence operations | Science Advances](#)