Uncovering Information Operations on Twitter Using Natural Language Processing and the Dynamic Wavelet Fingerprint

Spencer Lee Kirn

William & Mary, Applied Science Department, 2021
Field: Nondestructive Evaluation, Degree: PhD

Advisor: Mark Hinders, Professor of Applied Science

Abstract

Information Operations (IO) are campaigns waged by covert, powerful entities to distort public discourse in a direction that is advantageous for them. It is the behaviors of the underlying networks that signal these campaigns in action, not the specific content they are posting. In this dissertation we introduce a social media analysis system that uncovers these behaviors by analyzing the specific post timings of underlying accounts and networks. The presented method first clusters tweets based on content using Natural Language Processing (NLP). Each of these clusters - referred to as topics - are plotted in time using the attached metadata for each tweet. These topic signals are then analyzed using the Dynamic Wavelet Fingerprint (DWFP), which creates binary images of each topic that describe localized behaviors in the topic's propagation through Twitter. The features extracted from the DWFP and the underlying tweet metadata can be applied to various analyses. In this dissertation we present four applications of the presented method. First, we break down seven culturally significant tweet storms to identify characteristic, localized behavior that are common among and unique to each tweet storm. Next, we use the DWFP signal processing to identify bot accounts. Then this method is applied to a large dataset of tweets from the early weeks of the Covid-19 pandemic to identify densely connected communities, many of which display potential IO behaviors. Finally, this method is applied to a live-stream of Turkish tweets to identify coordinated networks working to push various agendas through a volatile time in Turkish politics.