Abstract

Many changes in our digital corpus have been brought about by the interplay between rapid advances in digital communication and the current environment characterized by pandemics, political polarization, and social unrest. One such change is the pace with which new words enter the mass vocabulary and the frequency at which meanings, perceptions, and interpretations of existing expressions change. The current state-of-the-art algorithms do not allow for an intuitive and rigorous detection of these changes in word meanings over time. We propose a dynamic graph-theoretic approach to inferring the semantics of words and phrases ("terms") and detecting temporal shifts. Our approach represents each term as a stochastic time-evolving set of contextual words and is a count-based distributional semantic model in nature. We use local clustering techniques to assess the structural changes in a given word's contextual words. We demonstrate the efficacy of our method by investigating the changes in the semantics of the phrase "Chinavirus". We conclude that the term took on a much more pejorative meaning when the White House used the term in the second half of March 2020, although the effect appears to have been temporary. We make both the dataset and the code used to generate this paper's results available.