1 Media-Rich Fake News Detection: A Survey

**Author:** Shivam B. Parikh and Pradeep K. Atrey

Fake News has been around for decades and with the advent of social media and modern day journalism at its peak, detection of media-rich fake news has been a popular topic in the research community. Given the challenges associated with detecting fake news research problem, researchers around the globe are trying to understand the basic characteristics of the problem statement. This paper aims to present an insight on characterization of news story in the modern diaspora combined with the differential content types of news story and its impact on readers. Subsequently, we dive into existing fake news detection approaches that are heavily based on text-based analysis, and also describe popular fake news data-sets. We conclude the paper by identifying 4 key open research challenges that can guide future research.

2 Automatic deception detection: Methods for finding fake news

**Author:** Conroy, N. J., Rubin, V. L., & Chen, Y

This research surveys the current state-of-the-art technologies that are instrumental in the adoption and development of fake news detection. “Fake news detection” is defined as the task of categorizing news along a continuum of veracity, with an associated measure of certainty. Veracity is compromised by the occurrence of intentional deceptions. The nature of online news publication has changed, such that traditional fact checking and vetting from potential deception is impossible against the flood arising from content generators, as well as various formats and genres.

The paper provides a typology of several varieties of veracity assessment methods emerging from two major categories – linguistic cue approaches (with machine learning), and network analysis approaches. We see promise in an innovative hybrid approach that combines linguistic cue and machine learning, with network-based behavioral data. Although designing a fake news detector is not a straightforward problem, we propose operational guidelines for a feasible fake news detecting system.

3 Weakly supervised learning for fake news detection on Twitter

Author: Helmstetter, S., & Paulheim, H

The problem of automatic detection of fake news in social media, e.g., on Twitter, has recently drawn some attention. Although, from a technical perspective, it can be regarded as a straight-forward, binary classification problem, the major challenge is the collection of large enough training corpora, since manual annotation of tweets as fake or non-fake news is an expensive and tedious endeavor. In this paper, we discuss a weakly supervised approach, which automatically collects a large-scale, but very noisy training dataset comprising hundreds of thousands of tweets. During collection, we automatically label tweets by their source, i.e., trustworthy or untrustworthy source, and train a classifier on this dataset. We then use that classifier for a different classification target, i.e., the classification of fake and non-fake tweets. Although the labels are not accurate according to the new classification target (not all tweets by an untrustworthy source need to be fake news, and vice versa), we show that despite this unclean inaccurate dataset, it is possible to detect fake news with an F1 score of up to 0.9.

4 A new benchmark dataset for fake news detection

Author: Wang, W. Y.

Automatic fake news detection is a challenging problem in deception detection, and it has tremendous real-world political and social impacts. However, statistical approaches to combating fake news has been dramatically limited by the lack of labeled benchmark datasets. In this paper, we present liar: a new, publicly available dataset for fake news detection. We collected a decade-long, 12.8K manually labeled short statements in various contexts from [this http URL](http://politifact.com/), which provides detailed analysis report and links to source documents for each case. This dataset can be used for fact-checking research as well. Notably, this new dataset is an order of magnitude larger than previously largest public fake news datasets of similar type. Empirically, we investigate automatic fake news detection based on surface-level linguistic patterns. We have designed a novel, hybrid convolutional neural network to integrate meta-data with text. We show that this hybrid approach can improve a text-only deep learning model.

5 Fake News Detection in Social Media

Author: Stahl, K

How much of what we read on social media and on supposedly “credible” news sites is trustworthy? It is extremely easy for anyone to post what they desire and although that can be acceptable, there is the notion of taking it a step too far, such as posting false information online in order to cause a panic, using lies to manipulate another person’s decision, or essentially anything else that can have lasting repercussions. There is so much information online that it is becoming impossible to decipher the true from the false. Thus, this leads to the problem of fake news.