Detecting Framing Changes in Topical News

Karthik Sheshadri, *Member, IEEE*, Chaitanya Shivade, and Munindar P. Singh, *Fellow, IEEE*

***Abstract*— Changes in the framing of topical news are known to foreshadow significant public, legislative, and commercial events. Automated detection of framing changes is, therefore, an important problem, which existing research has not considered. Previous approaches are manual surveys that rely on human effort and are consequently limited in scope. This article systematizes the discovery of framing changes through a fully unsupervised computational method that seeks to isolate framing change trends over several years. We demonstrate our approach by isolating framing change periods that correlate with previously known framing changes. We have prepared a new data set, consisting of over 12 000 articles from seven news topics or *domains*, in which earlier surveys have found framing changes. Finally, our work highlights the predictive utility of framing change detection, by identifying two domains in which framing changes foreshadowed substantial legislative activity, or preceded judicial interest.**

***Index Terms*— Framing, news media.**

I. Introduction

TO MOTIVATE the problem and approach of this article, let us investigate the primary causes of obesity in America. Public opinion and behavior on the subject have changed measurably since the late 1990s. As an example, Gunnars [1] compiled a list in 2015 of ten leading causes, six of which suggest that the processed food industry may be primarily responsible. By contrast, in the 1990s and early 2000s, popular opinion appeared to hold [2], [3] that obesity was primarily caused by individual behavior and lifestyle choices. What led to this change in public opinion?

We posit that news publishing on the subject of *obesity* contributed to the change in the public's opinion. Table I shows two representative snippets from news articles on *obesity* published in 1995 and 2015, respectively. While both address the same topic, the 1995 snippet implies responsibility on part of individuals, and the 2015 snippet implies responsibility on part of the processed food industry.

We posit that the above quotes represent a change in how the NYT *framed* the issue of obesity. According to Wikipedia [4], "framing is a schema of **interpretation** that individuals rely on to understand and respond to events. The choices they then make are influenced by their creation of a frame."

Manuscript received August 6, 2020; revised January 6, 2021; accepted February 23, 2021. *(Corresponding author: Karthik Sheshadri.)*

Karthik Sheshadri and Munindar P. Singh are with the Department of Computer Science, North Carolina State University, Raleigh, NC 27606 USA (e-mail: kshesha@ncsu.edu; mpsingh@ncsu.edu).

Chaitanya Shivade is with Amazon Inc., Seattle, WA 98109 USA.

This article has supplementary downloadable material available at https://doi.org/10.1109/TCSS.2021.3063108, provided by the authors.

Digital Object Identifier 10.1109/TCSS.2021.3063108

TABLE I

Representative Snippets From the New York Times on the Domain *Obesity* From 1995 and 2015, Respectively

"For nearly four decades, health and fitness experts have ***prodded and cajoled*** *and used other powers of persuasion in a* ***futile attempt to whip*** *America's youngsters into shape."*
Quote from The New York Times, 1995
"The New York City Health Department has embarked on a new campaign to ***persuade processed food brands to decrease sugar content*** *in a bid to curb obesity."*
Quote from The New York Times, 2015

We hypothesize that the above is an example of an "equivalence frame" [4] in which the information being presented is based on the same facts (the fact of obesity) but the "frame" in which it is presented changes (the frame of who is responsible for obesity), thus creating a reference-dependent perception. Hence, we highlight the word "interpretation" in the previous paragraph. Thus, the "frame" surrounding the issue can change the reader's perception without having to alter the actual facts, as the same information is used as a base. This is done through the media's choice of certain words and images to cover a story (as in the example above). The definitions that we use in this and the previous paragraph quote text from [4].

Thus, framing theory [5], [6] suggests that how a topic is presented to the audience (called "the frame") influences the choices people make about how to process that information.

In general, framing has been shown to have a far-reaching impact on areas, such as Public Policy [7], Political Science [8], International Affairs [9], and Sociology [10].

Existing literature shows that framing may deliberately be used for the purpose of manipulating political processes [11]. This motivates the computational analysis of framing, as a possible means of identifying such manipulation.

An automated method to identify framing changes may, hence, potentially be applied to foreshadow changes in the public interest and opinion [8], legislative interest [7], and political manipulation [11]. These considerations motivate the problem of *framing change detection*, which involves identifying when the dominant frame (or frames) [10] of a topic undergoes a change. We note that our use of the term "dominant frame" is similar in spirit to the term "emphasis frame" defined in [4].

Scholars have noted that all communication necessarily entails framing [12]. Here, we study framing changes from the lens of adjective polarity [13], [14]. We identify framing changes that manifest as trends of changes in polarity [14] over a period of years, such as the one shown in the example on *obesity*. We acknowledge that other interpretations of framing changes are possible. However, we posit that our approach successfully identifies changes, such as the one described above, which is our objective.

A. Related Work

The Media Frames Corpus, compiled by Card *et al.* [15], studies three topics (Immigration, Smoking, and same-sex marriages) and identifies fifteen *framing dimensions* in each. We identify two major limitations of their work. First, Card *et al.* [15] studied framing as a static detection problem, identifying which dimensions appear in a given news article. However, research in sociology [10] shows that most news topics feature a *dominant frame* (or dominant dimension in the terminology of [15]). Furthermore, for a generic news topic, the dominant frame is not necessarily one of fifteen previously chosen dimensions but can instead be an unknown arbitrary frame specific to the topic under consideration. For example, in the example given in Section I, the dominant frame related to the privacy of individuals is not one of the fifteen dimensions described in Card *et al.* [15].

Second, Sheshadri and Singh [7] showed that public and legislative reactions tend to occur only after *changes* in the dominant frame. That finding motivates an approach to framing that focuses on identifying and detecting changes in the dominant frame of a news domain.

Sheshadri and Singh [7] further proposed two simple metrics that they motivate as measures of domain framing: framing polarity and density. They define framing polarity as the average frequency of occurrence in a domain corpus of terms from a benchmark sentiment lexicon. Framing density is measured using an entropic approach that counts the number of terms per article required to distinguish a current corpus from an earlier one.

We identify the following limitations of the aforementioned measures [7]. First, both measures make no effort to associate a given news article with a particular frame. Prior work does not support the inherent assumption that all articles in a given domain belong to a particular frame [10], [15]. We enhance understanding by analyzing each domain using several distinct frames.

Second, framing density relies on estimating changes in the density of occurrence of generic n-grams (both nouns and adjectives) between successive corpora. We illustrate the limitations of this approach by distinguishing between *frame* and *content*. Following the literature [4], [5], [16], we define *content* as the information presented in an article and *frame* as the manner in which it is presented. We posit that highly polarized adjectives communicate *frame* but do not communicate *content*. As an example, consider the following two sentences: 1) “He was murdered” and 2) “His murder was cowardly and reprehensible.” While both sentences convey the same facts, the presence of the adjectives “cowardly” and “reprehensible” serves to *frame* the second sentence without conveying any additional factual information. Therefore, since polarized adjectives are unlikely to communicate factual information, we take them to be artifacts of how an event or topic is framed. Framing density makes no distinction between content and frame.

It is worthwhile to note that our approach is similar in spirit to Tsur *et al.*’s work [17], in which both that work and this article apply a topic modeling strategy to analyze framing as a time series. However, as in the case of framing density, generic n-grams (as used by Tsur *et al.* [17]) do not distinguish between frame and content.

We also distinguish our work from Alashri *et al.* [9] who use standard machine learning tools to classify sentences represented by linguistic features into one of four frames. Such an approach is limited by the need to predefine a frame set, as does Card etal’s approach [15].

B. Contributions

This article contributes an unsupervised, data-driven natural language approach to detecting framing change trends in published news, a problem of significant public and legislative import. Our approach agrees with and extends the results of earlier manual surveys, which required human data collection and were consequently limited in scope. Our method can be run simultaneously over all news domains, limited only by the availability of real-time news data. Furthermore, our approach yields results that foreshadow periods of legislative activity. This motivates the predictive utility of our method for understanding the legislative activity.

Furthermore, we contribute the *Framing Changes* data set (to be publicly released), a collection of over 12 000 news articles from seven news topics or *domains*. In four of these domains, surveys carried out in earlier research have shown framing to change. In two domains, periods with significant legislative activity are considered. Our individual domain data sets within the framing change data set cover the years in which earlier research found framing changes, as well as periods ranging up to ten years before and after the change. Our data set is the first to enable computational modeling of framing change trends.

II. Materials and Methods

This section describes our data sets, data sources, and inter-annotator agreement. All data were collected in an anonymous and aggregated manner. All APIs and data used are publicly available, and our data collection complies with the terms and conditions of each API.

A. Data Source: The New York Times API

The New York Times (NYT) Developer’s API [18] provides access to news data from the NYT, Reuters, and Associated Press (AP) newspapers—both print and online versions—beginning in 1985. The NYT has the second largest circulation of any newspaper in the United States [19].

The data object returned by the API includes fields such as the article type (news, reviews, summaries, and so on), the news source (NYT, Reuters, or AP), the article's word count, the date of its publication, and article text (in the form of the abstract, the lead (first) paragraph, and a summary).

B. Data Source: The Guardian API

The Guardian Xplore API [20] provides access to news data from The Guardian, a prominent U.K. newspaper that reaches 23 million readers per month [21]. This API returns full-length articles along with such metadata as the article type (similar to the NYT API) and a general section name (such as sports and politics). Although these section names are manually annotated by humans, we do not use them in our analysis but rely instead on a simple term search procedure (see Section II-C) to annotate our data sets.

C. Domain Data Set Generation

As in earlier work [7], [8], [22], we use a standard term search procedure to create our data sets. Specifically, an article is deemed a *domain positive* if at least a component of the article discusses a topic that is directly relevant to the domain [7] and, otherwise, a *domain negative*. Examples of both domain positives and negatives are presented in the Supplementary Material. We define data set accuracy as the fraction of articles in a data set that is domain positive. For each domain, our APIs were used to extract news data during the time period b (denoting the beginning) to e (denoting the end) of the period of interest.

We analyze framing changes that occurred either in the U.S. or the U.K. We use news articles from the NYT to study news patterns' preceding framing changes that occurred in the U.S. and news articles from The Guardian to study news patterns' preceding changes in the U.K. All but one of our domains are from the NYT. We analyze the following domains from the NYT: *Smoking*, *Surveillance*, *Obesity*, *LGBT Right*, *Abortion*, and *Drones*. We analyze the domain *Immigration* from the Guardian.

D. Interannotator Agreement

To ensure that the articles returned by our term search procedure are indeed relevant to each domain, a random sample of articles from each domain data set was coded by two raters. We obtained median per-domain data set accuracies of 0.8 for coder 1 and 0.75 for coder 2 and a median Kappa of 0.67 over sample domains. A full table is shown in the Supplementary Material.

E. Probability Distribution Over Adjectives

Our approach relies on the key intuition that, during a framing change, the valence of the adjectives describing co-occurring nouns changes significantly. We note that a wealth of existing literature supports this observation [23]–[25]. We further acknowledge that framing changes that do not manifest as substantial changes in tone (for example, changes that occur by a shift from one aspect of the issue to another without changing polarity) can arise. Our approach does not capture such changes, however, and we restrict our analysis to polarized framing. We show that many domains of relevance to public policy and response undergo such polarized framing changes, and we demonstrate that our approach successfully identifies such changes.

To measure this change, we create a reference probability distribution of adjectives based on the frequency of their occurrence in benchmark sentiment data sets.

1) Benchmark Data Sets: We identified three open-source benchmark review data sets (namely, the Trip Advisor data set, the Yelp Challenge data set, and the Amazon Review data set) to create our adjective probability distribution. Together, these data sets provide about 150 million reviews of various restaurants, services, and products, with each review rated from one to five. Given that the large volume of reviews from different sources made available by these data sets, we assume that they provide a sufficiently realistic representation of all adjectives in the English language.

We rely primarily on the Trip Advisor data set to create our adjective probability distribution. Due to the fact that the Yelp and Amazon data sets together comprise about 150 million reviews, it is computationally infeasible for us to include them in our learning procedure. Instead, we learned distributions from these data sets for sample adjectives to serve as a comparison with and as verification of our overall learned distribution. The resulting distributions for these adjectives appeared substantially similar to those of the corresponding adjectives in our learned distribution. We, therefore, conclude that our learned distribution provides a valid representation of all adjectives in the English language. We describe the Trip Advisor data set below and omit a description of the others due to space constraints.

2) Trip Advisor: The Trip Advisor data set consists of 236 000 hotel reviews. Each review provides text, an overall rating, and aspect-specific ratings for the following seven aspects: Rooms, Cleanliness, Value, Service, Location, Checkin, and Business. We limit ourselves to using the overall rating of each review.

F. Polarity of Adjectives

For each adjective in the English language, we are interested in producing a probability distribution that describes the relative likelihood of the adjective appearing in a review whose rating is r, i.e., $P(r|\text{adj}), r \in \{1, 2, 3, 4, 5\}, \text{adj} \in A$, where A is the set of all adjectives in the English language and r ranges from one to five.

We began by compiling a set of reviews from the Trip Advisor data set for each rating from one to five. We used the Stanford CoreNLP parser [14] to parse each of the five sets of reviews so obtained. We, thus, obtained sets of parses corresponding to each review set. From the set of resultant parses, we extracted all words that were assigned a part-of-speech of "JJ" (adjective). Our search identified 454 281 unique adjectives.

For each unique adjective adj, we counted the number of times it occurred in our set of parses corresponding to review ratings one to five. We denote this by N_i, with $1 \le i \le 5$. Our probability vector for adjective adj is then $P(r|\text{adj}) = \{(N^1_{\text{adj}}/S_{\text{adj}}), (N^2_{\text{adj}}/S_{\text{adj}}), \ldots, (N^5_{\text{adj}}/S_{\text{adj}})\}$, where $S_{\text{adj}} = N^1_{\text{adj}} + N^2_{\text{adj}} + N^3_{\text{adj}} + N^4_{\text{adj}} + N^5_{\text{adj}}$.

In addition, we recorded the rarity of each adjective as $(1/S_{\text{adj}})$. This estimates a probability distribution P, with 454 281 rows and six columns.

Table II shows example entries from our learned probability distribution. As can be seen from the table, our learned distribution not only correctly encodes probabilities (the adjective "great" has nearly 80% of its probability mass in the ratings four and five, whereas the adjective "horrible" has nearly 80% of its mass in ratings one and two) but also implicitly learns an adjective ranking, such as the one described by De Melo and Bansal [26]. To illustrate this ranking, consider that the adjective "excellent" has 60% of its probability mass in rating five, whereas the corresponding mass for the adjective "good" is only 38%.

For a visual illustration, we depict our learned probability distribution as a heatmap in the Supplementary Material.

Motivated by our learned probability distribution, we posit that rating 1 represents negativity, ratings 2–4 represent neutrality, and rating 5 represents positivity. We refer to the difference between our highest rating and lowest rating as our polarity.

G. Incorporating Adjective Rarity

Our measure of adjective rarity serves as a method by which uncommon adjectives, which rarely occurs in our benchmark data set and whose learned probability distributions may, therefore, be unreliable, can be excluded.

However, in doing so, we run the risk of excluding relevant adjectives from the analysis. We manually inspect the set of adjectives that describe the nouns in each domain to arrive at a domain-specific threshold.

Section 2.2 in the Supplementary Material details the parameters used in each domain. However, we observe that the trends in our results appeared to be fairly consistent across a reasonable range of threshold values.

H. Domain Period of Interest

We define a period of interest for each domain. Let t_f be a year in which a documented framing change took place in the domain under consideration. Then, our period of interest for this domain is $b = \min(t_f - 10, t_f - l)$ to $e = \max(t_f + 10, t_f + r)$, where the API provides data up to l years before and r years after t_f. All units are in years.

I. Corpus-Specific Representations

A domain corpus is a set of news articles from a given domain. Let a given domain have m years in its period of interest with annual domain corpora $T_1, T_2, \ldots, T_m$.

J. Corpus Clustering

An overall domain corpus is, therefore, $T = T_1 \cup T_2 \cup \ldots \cup T_m$. We assume that a domain corpus addresses k unique topics. We measure the framing of each topic within a domain corpus distinctly from the others. We adopt a standard topic modeling approach to estimate topics within a domain. We use the benchmark latent Dirichlet allocation (LDA) [27] approach to model $k = 5$ topics in each domain corpus. We extract the top $l = 20$ nouns υ from each topic. We also extract the set of all unique nouns in T. We define a cluster as the set of nouns $\upsilon \cap T$. We, thus, generate k clusters, each representing a unique topic. This enables us to measure the framing of each topic using the approach described in Section II-K.

The following example can help understand our approach. In the phrase, "horrible drone strike," the adjective "horrible" frames the phrase "drone strike." We study changes in polarity of adjectives that frame nouns within the same cluster and aggregate them to arrive at an overall estimate of framing within that cluster.

K. Annual Cluster Polarity

For each cluster c, we are interested in arriving at a vector of m annual polarities, i.e., for each year i, $1 \le i \le m$ in the domain period of interest.

Let x_c be the set of all nouns in c. For each noun $\upsilon \in x_c$, we use the Stanford dependence parser [14] to identify all adjectives (without removing duplicates) that describe υ in T_i. We extract the polarity vectors for each of these adjectives from P as the matrix A_i. A_i, therefore, has n rows: one for each adjective so identified, and five columns (see Section II-F). We estimate the annual cluster polarity of c as the vector of columnwise averages of A_i. Let $P_c = \{P_1, P_2, \ldots, P_m\}$ be the set of annual cluster polarities so obtained. Annual polarities for representative clusters from each of our domains are shown in Figs. 3–15.

L. Defining Framing Changes

The measurement of any temporally disparate pair of news corpora using adjective polarity (or any other numerical metric) would result in different representative values of the two corpora. Therefore, we cannot use the fact that two different corpora merely yield different values in order to isolate a framing change since this is always the case.

Furthermore, individual metrics are susceptible to noisy readings due to imprecise data and measurement. In particular, such an effect may cause sudden isolated spikes between successive measurements. For example, in Fig. 4, during the period between 2005 and 2006, while ratings 1, 3, 4, and 5 changed little, rating 2 showed a substantial change.

This motivates the question of how a framing change is defined in the context of our computational measurements. The usual social science definition [16] is that a framing change is a shift in the way that a specific topic is presented to an audience. To isolate such changes computationally, we use the following key observations from ground-truth framing changes: 1) framing changes take place as trends that are consistent over at least k years and 2) framing changes must be consistent across multiple measurements.

Our aim in this article is to begin from a set of time series, such as the ones in Fig. 4, and isolate such trends.

TABLE II
SAMPLE ENTRIES FROM OUR LEARNED PROBABILITY DISTRIBUTION FOR POSITIVE AND NEGATIVE SENTIMENT ADJECTIVES

Adjective	Rating 1	Rating 2	Rating 3	Rating 4	Rating 5	Rarity (Inverse Scale)
Great	0.039	0.048	0.093	0.274	0.545	4.495e-07
Excellent	0.019	0.028	0.070	0.269	0.612	2.739e-06
Attractive	0.095	0.125	0.192	0.296	0.292	0.0001
Cute	0.039	0.068	0.155	0.330	0.407	1.499e-05
Compassionate	0.068	0.020	0.010	0.038	0.864	0.0004
Good	0.076	0.095	0.185	0.336	0.308	3.459e-07
Horrible	0.682	0.143	0.076	0.042	0.057	7.453e-06
Ridiculous	0.461	0.180	0.125	0.116	0.118	2.033e-05
Angry	0.546	0.138	0.092	0.098	0.126	6.955e-05
Stupid	0.484	0.136	0.099	0.117	0.164	5.364e-05
Beautiful	0.043	0.049	0.085	0.222	0.599	6.233e-06

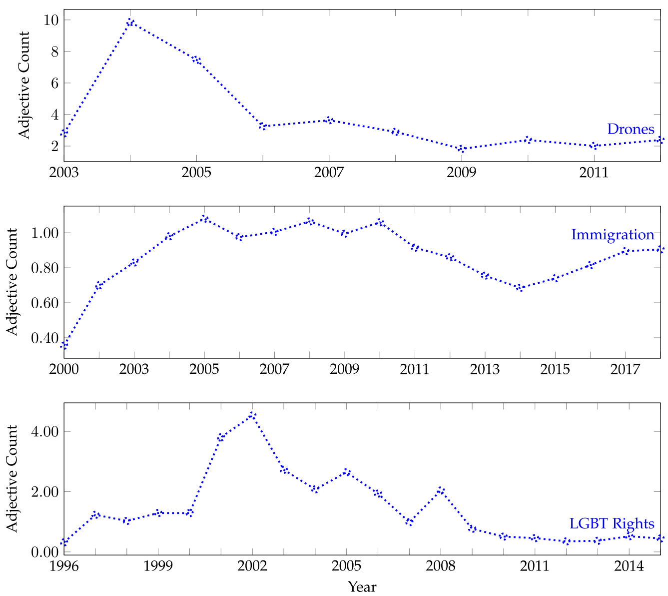


Fig. 1. Average number of adjectives per article, shown for our domains over their respective periods of interest. This metric serves as a measure of the subjectivity of news in a domain.

The requirement motivated by our first condition, namely, that framing changes must last at least k years, is easy to satisfy by imposing such a numerical threshold.

To satisfy the requirement motivated by our second observation, we rely on *correlations* between different measurements, as described in Section II-M.

M. Detecting Framing Changes Using Periods of Maximum Correlation

Our five polarity ratings serve as measures of framing within a domain. We conceive of a framing change as a trend, consistent across our five polarity ratings, over a period of some years.

We describe our intuition and approach to detecting framing changes in the following. First, we show that the frequency with which adjectives occur in articles varies both by domain and in different years within a domain.

Fig. 1 depicts the average number of adjectives per article for a subset of our domains over the years in their respective periods of interest. We note that this count serves also as a measure of how subjective news publishing in a domain is since adjectives are indicative of how events are framed. An equivalent figure for the remaining domains is given in the Supplementary Material.

Despite the fact that the volume of adjectives used per article varies dramatically (by up to 30%), we find that the variation in our annual cluster polarity between successive years is generally on the order of less than 1%. However, through a consistent trend lasting multiple years, our measure of annual polarity can change (increase or decrease) cumulatively by, for example, up to 5% (see Fig. 4 for an example). We identify a framing change based on such a cumulative trend.

We now consider the problem of fusing estimates from our five measures of annual cluster polarity. Consider the change in polarity of ratings 1, 3, 4, and 5 between 2005 and 2006 in Fig. 4, as against the change in rating 2. As mentioned earlier, ratings 1, 3, 4, and 5 changed little, whereas rating 2 showed a substantial change.

In contrast, in the period 2002–2004, a consistent trend was observable across all five ratings, with substantial reductions in ratings 2 and 3, and a notable corresponding increase in rating 5. We exploit *correlations* between the changes in our five ratings to identify framing changes.

Accordingly, we measure trend consistency via the Pearson correlation [28] between our ratings. Suppose that a domain has m years in its period of interest. We generate all possible contiguous subsets of T_m, namely, T_{i-j}, where $i \leq j \leq m$, and T_{i-j} denotes the domain corpus from year i to year j.

Let $C = \{C_1, \ldots, C_5\}$ be the set of rating vectors for this domain subset, where

$$C_1 = \begin{bmatrix} C_1^i \\ C_1^{i+1} \\ \cdots \\ C_1^j \end{bmatrix}, \quad C_1^i$$

is the value of rating 1 for year i and similarly for $C_2, \ldots, C_5$.

To measure the correlation of subset T_{i-j}, we compute its matrix of correlation coefficients [29] K. We reshape K into a vector of size $f \times 1$, where $f = i * j$, and evaluate its median, l. We now introduce two variants of our approach below. A block diagram depicting our overall approach is shown in Fig. 2.

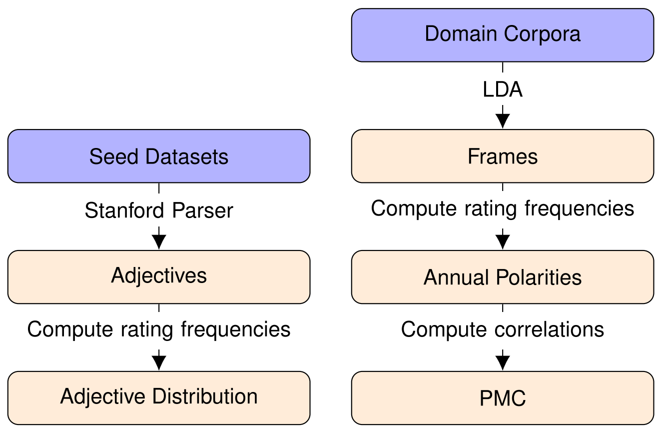


Fig. 2. Block diagram illustrating our approach. Our adjective distribution is computed using per rating frequencies of occurrence for each adjective in the seed data set(s). We use this distribution to compute annual cluster polarities of frames obtained using LDA from our domain corpora.

1) Unweighted Optimization: We find the maximum value of l, $l_{\max}$, over all possible values of i and j. We denote the values of i and j corresponding to $l_{\max}$ as $i_{\max}$ and $j_{\max}$. We return $T_{i_{\max}-j_{\max}}$ as our *period of maximum correlation (PMC)*.

We note that the smaller the duration of a PMC, the greater the possibility that our rating vectors may have a high correlation in the period due to random chance. To compensate for this effect, we employ a threshold, whereby a period is not considered as a candidate for the domain PMC unless it lasts at least y years. We uniformly employ a value of $y = 3$ in this article.

Our approach, thus, identifies polarity drifts that are both correlated (quantitatively measured by correlations between different measures of polarity) and sustained (by the imposition of a threshold of duration). Note that our approach filters out isolated drifts in individual polarity measures since such drifts are uncorrelated across multiple measures.

2) Weighted Optimization: In this variant, we maximize $l_{\max} * \text{dist}(C_i, C_j)$ over all possible values of i and j, where $\text{dist}(C_1, C_j)$ represents the Euclidean distance between the rating vectors for years i and j. The intuition behind this variant is to incorporate the magnitude of the drift into our optimization so that a larger drift, if consistent across multiple polarity measures, will be weighted higher than a smaller drift that is also correlated. Please refer to the Appendix for an example.

While, in the case of unweighted optimization, larger values of y automatically yield lower correlations, the threshold of y years is presented as a minimum, not an exact period. In the case of weighted optimization, this consideration does not hold. Therefore, we take y to refer to an exact period of years rather than a minimum threshold in this variant.

III. Quantitative Evaluation

We now discuss a partial quantitative evaluation of our approach using a precision–recall analysis. Our analysis relies on the ground-truth annotation of framing changes, as detailed in Section III-A. We are unable to conduct a full precision–recall analysis over all domains due to the challenges discussed in the following. However, we expect that our partial analysis is representative of the general performance of the approach.

A. Ground-Truth Annotation

We label a ground truth for each domain, marking years corresponding to framing changes as positives and other years as negatives. We primarily obtain our positives using the findings of large-scale surveys from earlier research.

In order to do so, we study the literature pertaining to framing changes in the domains we examine. We identify large-scale studies conducted by reputed organizations, such as the National Cancer Institute (NCI) [30], the Columbia Journalism Review (CJR) [31], and Pew Research [32]. These studies examine news and media publishing in a particular domain over a period of time, as we do, and manually identify changes in the framing of domain news during these periods.

The studies that we rely on for ground truth sometimes provide quantitative justification for their findings. For example, the NCI monograph on the framing of *smoking* news identifies the number of pro and antitobacco control frames before and after a framing change [30]. These studies, therefore, provide an expert annotation of framing changes in our domains, for the periods that we examine. Details of each study that we used and their findings are reported in Section IV.

By demonstrating substantial agreement between the results of our approach and those of earlier ground-truth surveys, we establish our claim that our approach may be used to automatically identify framing changes in domain news publishing.

B. Precision–Recall Analysis

To gain confidence that our approach successfully identifies framing changes, we conduct a precision–recall analysis on our data. We consider each year in each domain as a data point in our analysis. We calculate overall precision and recall over all data points in our domains. Within a specific domain, let $L = e - b + 1$ be the number of years in our period of interest. Hence, for each year $1 \leq i \leq L$, we define a binary label corresponding to whether this year was part of a ground-truth framing change or not. Hence, the problem of framing change detection involves predicting a binary label corresponding to change/no change, for each value of i within a domain.

Thus, we consider a data point a true positive or true negative if both a ground-truth study and our approach are labeled as corresponding to a framing change, or otherwise, respectively. We refer to a data point that was labeled as a positive (or negative) by our approach but a negative (or positive) according to the relevant ground-truth survey as a false positive or false negative, respectively.

We calculate precision as $P = (\text{tp}/\text{tp} + \text{fp})$ and recall as $R = (\text{tp}/\text{tp} + \text{fn})$, where tp, fp, and fn are the numbers of true positives, false positives, and false negatives, respectively.

In the *Obesity* domain, the earlier survey that we found did not state a precise framing change period. We, hence, exclude this domain from our precision–recall analysis.

Fig. 3. Estimated clusters for the *smoking* domain. Each cluster represents a unique *frame*. The frame of cluster 3, characterized by the terms "cancer" and "smoke," discusses the health risks associated with smoking. We analyze this cluster and estimate a PMC of 2002–2004 (see Fig. 4). Our PMC coincides with an earlier monograph from the NCI that describes a progression toward tobacco control frames in American media between 2001 and 2003.

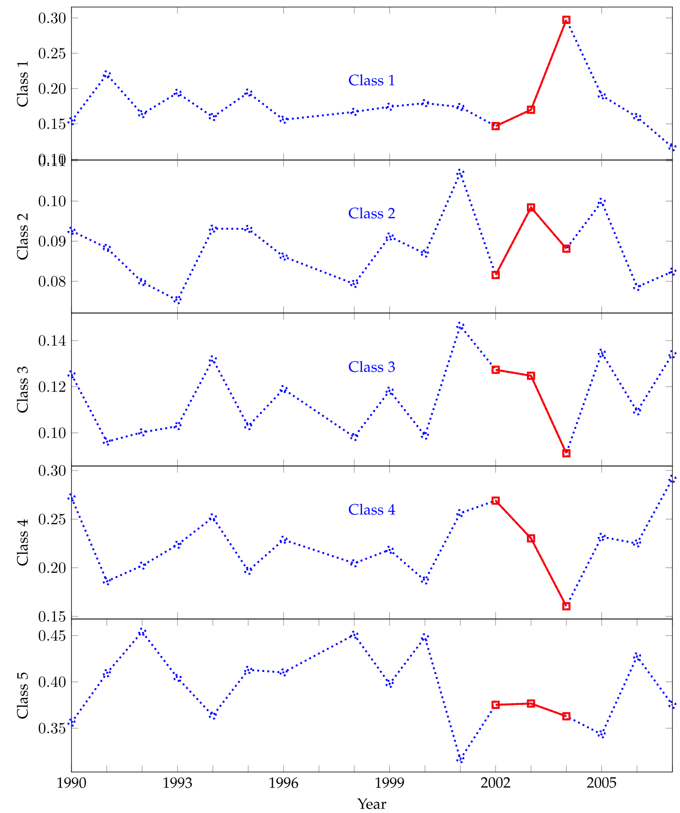


Fig. 4. Annual polarities for cluster 3 (characterized by the terms "cancer" and "smoke") from Fig. 3 from the *smoking* domain for ratings 1–5. The PMC is shown with solid lines in square markers and agrees with a framing change described in an earlier NCI monograph.

IV. Results

We find that our periods of maximum correlation correlate substantially with framing changes described in earlier surveys [2], [31], [33], [34] and may also foreshadow legislation.

Our computed rating vectors are depicted in Figs. 4–15. We discuss each domain in the following.

A. Smoking

The NCI published a monograph discussing the influence of the news media on tobacco use [30]. On page 337, the monograph describes how, during the period 2001–2003, American news media had progressed toward tobacco control frames. It states that 55% of articles in this period reported progress on tobacco control, whereas only 23% reported setbacks.

In contrast, the monograph finds (also on page 337) that, between 1985 and 1996, tobacco control frames (11) were fairly well balanced with pro-tobacco frames (10). We extracted a data set of over 2000 articles from 1990 to 2007.

Our approach returns a PMC of 2002–2004 (see Fig. 4) for this domain. Since no studies cover the period 1997–2000 [30], we interpret the findings described in the monograph to imply that the change toward tobacco control frames predominantly began in 2001 and ended in 2003. This domain, therefore, contributes two true positives (2002 and 2003), one false negative (2001), and one false positive (2004) to our precision–recall analysis.

B. Surveillance

The CJR [31] found that following the Snowden revelations, news coverage of Surveillance in the U.S. changed to a narrative focusing on individual rights and digital privacy [7]. We compiled a data set consisting of approximately 2000 *surveillance* articles from the NYT for the period 2010–2016.

The frame of cluster 3 in Fig. 5, characterized by the terms "national," "security," and "agency," discusses the Snowden revelations of 2013. We analyze this cluster. Our rating vectors for this domain are shown in Fig. 6. We obtain a PMC of 2012–2014 for this period corresponding closely to the ground-truth framing change.

The trends in our rating vectors are indicative of the change. As can be seen from Fig. 6, positivity (measured by rating 5) drops markedly, together with a simultaneous increase in negativity (rating 1) and neutrality (ratings 2 and 3). Rating 4 remains close to constant during this period and, thus, does not affect our hypothesis.

We interpret the findings of [31] to refer primarily to 2013, the year in which the revelations were made (hence, a change between 2012 and 2013), and the following year, 2014. While other interpretations may conclude a longer framing change, they must necessarily include this period. This domain, therefore, contributes three true positives (2012, 2013, and 2014) with no false positives or negatives to our quantitative evaluation.

C. Obesity

Kim and Willis [2] found that the framing of *obesity* news underwent changes between the years 1997 and 2004. During

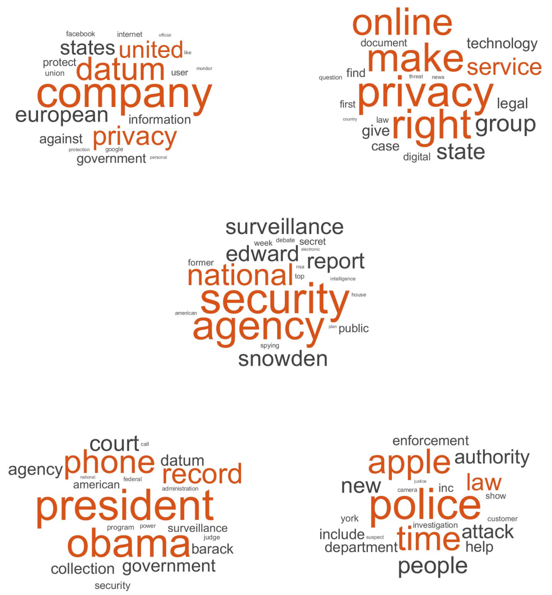


Fig. 5. Our estimated clusters for the domain *surveillance*. Each cluster is said to represent a unique *frame*. The frame of cluster 3, characterized by the terms "national," "security," and "agency," discusses the Snowden revelations of 2013. We analyze this cluster and estimate a PMC of 2013–2014 (see Fig. 6). Our PMC coincides exactly with the period following the Snowden revelations. In addition, we note that the CJR [31] found that following the Snowden revelations, news coverage of Surveillance changed to a narrative focusing on individual rights, and digital privacy [7].

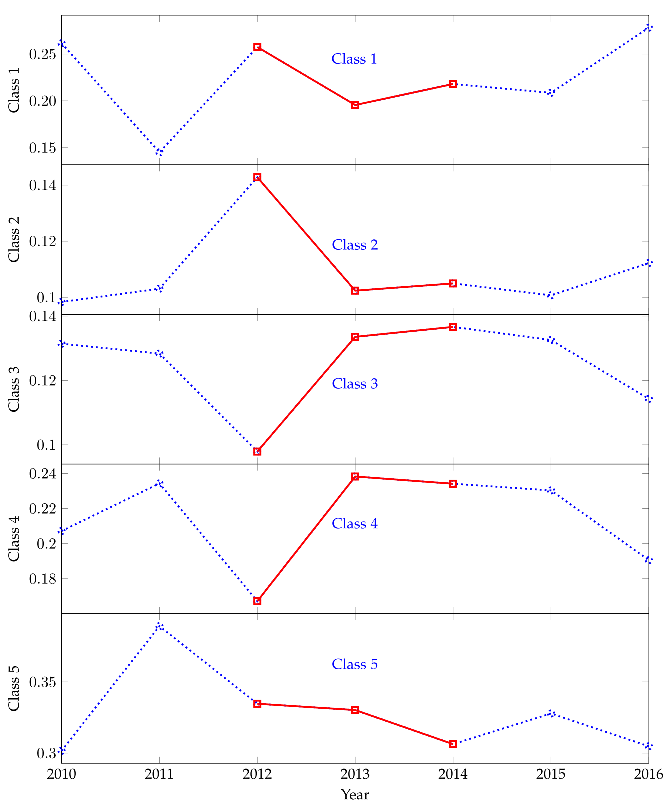


Fig. 6. Annual polarities for a representative cluster (characterized by the terms "national," "security," and "agency") from the domain *surveillance* for ratings 1–5. The PMC is shown with solid lines in square markers.

Fig. 7. Our estimated clusters for the domain *obesity*. Each cluster is said to represent a unique *frame*. We posit that cluster 2 (characterized by the terms "food," "diet," and "make") represents societal causes of obesity (see Section IV-C). We analyze this cluster and estimate a PMC of 2005–2007 (see Fig. 8). Our PMC agrees with the findings of an earlier human survey [2].

this period, Kim and Willis [2] found that the fraction of news frames attributing responsibility for obesity to social causes increased significantly. Prior to this period, *obesity* tended to be framed as an issue of individual responsibility. For example, *obesity* news after the year 2000 has often criticized food chains for their excessive use of sugar in fast food, as shown in the NYT snippet in Section I. We compiled a data set of over 3000 articles from the NYT (since Kim and Willis [2] restrict their study to Americans) from 1990 to 2009.

The clusters that we estimate for this domain are shown in Fig. 7. Cluster 2 addresses possible causes of obesity, with a particular focus on dietary habits. We posit that this cluster represents societal causes more than individual ones (since individual causes, as shown in the NYT snippet of Section I, tend to discuss topics, such as fitness and sedentary lifestyles, as opposed to food content). We observe that the PMC for this domain (2005–2007) is characterized by increased negativity, shown by ratings 1 and 2, and decreased positivity (rating 5). Our results for this domain agree with the findings of Kim and Willis [2].

We were unable to use this domain in our precision–recall analysis since Kim and Willis, to the best of our knowledge, do not specify a precise period during which the framing change took place.

However, since Kim and Willis [2, Figs. 2 and 3] show a dramatic increase of social causes in 2004 and a corresponding marked decline of individual causes, we conclude a substantial agreement between their findings and our results.

D. LGBT Rights

We compiled a data set of over 3000 articles from the period 1996–2015 in this domain. Fig. 9 depicts our estimated clusters. Cluster 3 represents a frame that discusses the subject of same-sex marriage and its legality. We note that the Supreme Court ruled to legalize same-sex marriages in the

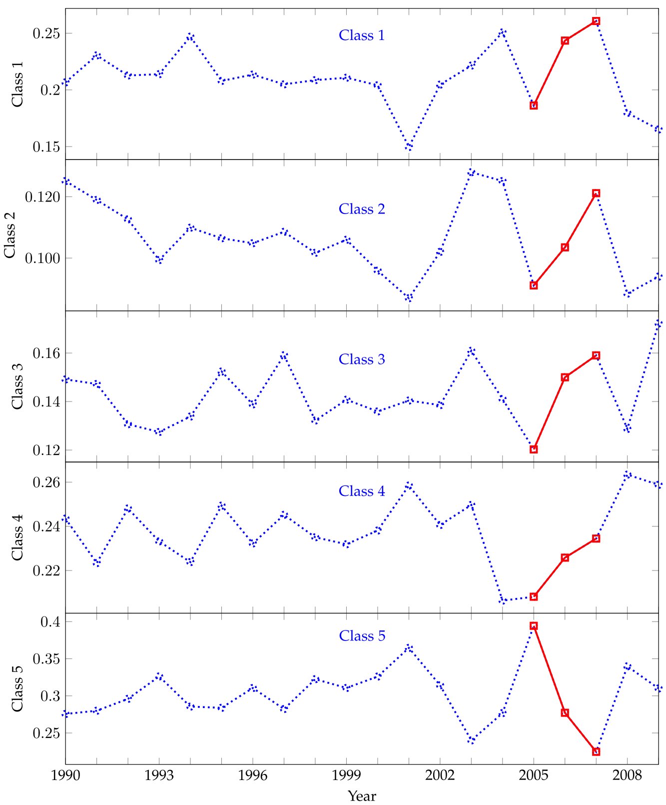


Fig. 8. Annual polarities for cluster 2 (characterized by the terms "diet," "food," and "make") from Fig. 7 from the domain *obesity* for ratings 1–5. The PMC is shown with solid lines in square markers. We posit that this cluster represents societal causes of obesity (see Section IV-C). We observe that the PMC for this cluster (2005–2007) agrees with the findings of Kim and Willis [2].

U.S. in the year 2015. Our rating vectors for this domain are shown in Fig. 10. At $y = 4$, our weighted approach captures and foreshadows this period (2012–2015), as shown in Fig. 10.

Furthermore, Gainous and Rhodebeck [34] studied the framing of LGBT-related publishing in the NYT over the period 1988–2012 and found a dramatic increase in equality frames between approximately 25 in 2008 and approximately 110 in 2012. Correspondingly, our findings of Fig. 10 show that, between 2008 and 2012, there was a dramatic increase in the measures of ratings 4 and 5 (representing positivity) and a marked reduction in the measures of ratings 1 and 2 (representing negativity). At a threshold of $y = 5$, this period has the second highest correlation and weighted by distance (0.0842 versus 0.0859, which is the highest) and is shown in green in the same figure.

We rely on Gainous and Rhodebeck [34] for ground truth in this domain and use this period as our PMC. We, therefore, conclude that our measures return four true positives and no false positive for this domain.

E. Abortion

The Partial-Birth Abortion Ban Act was enacted in 2003. We obtained 248 articles for the period 2000–2003, for this domain. We obtain a PMC of 2001–2003 for this domain, corresponding to three true positives and no false positives or negatives. Due to space constraints, we present the corresponding figure in the Supplementary Material.

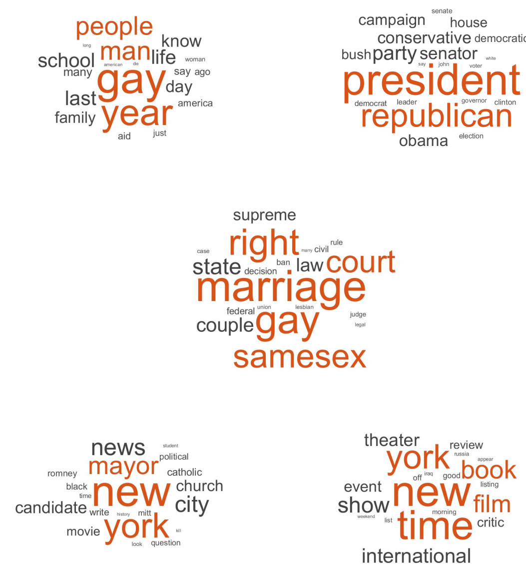


Fig. 9. Our estimated clusters for the domain *LGBT Rights*. Each cluster is said to represent a unique *frame*. The frame discussed in cluster 3 discusses the subject of same-sex marriage and, in particular, judicial interest in this topic. We analyze this cluster and estimate two PMCs of nearly identical correlation score (2006–2008 and 2013–2015 Fig. 10). The PMC of 2013–2015 coincides exactly with the Supreme Court judgment of 2015 that legalized same-sex marriage in the entire U.S.

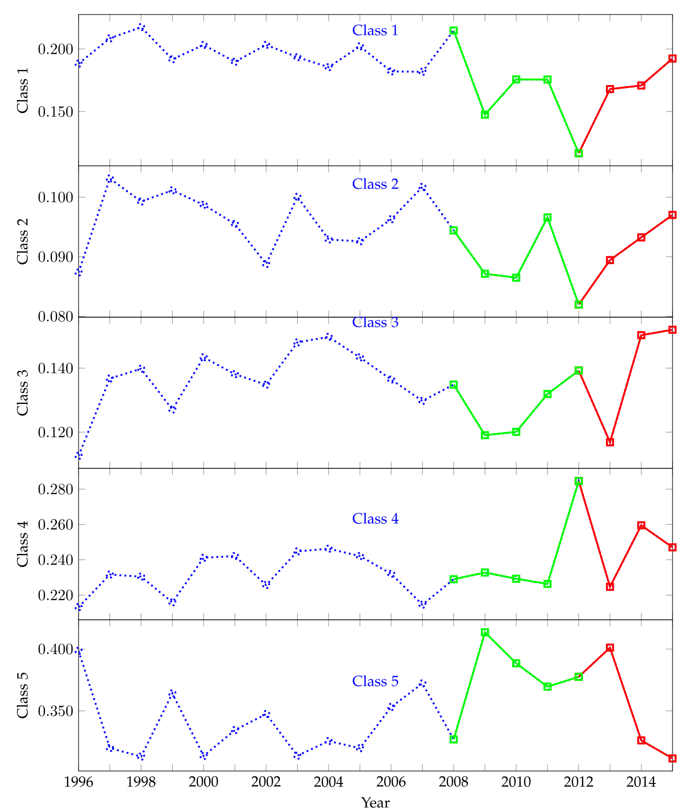


Fig. 10. Annual polarities for cluster 3, characterized by the terms "gay," "rights," and "marriage," in Fig. 9 from the domain *LGBT Rights* for ratings 1–5. Our PMC of 2012–2015 immediately precedes the judicial interest of 2015.

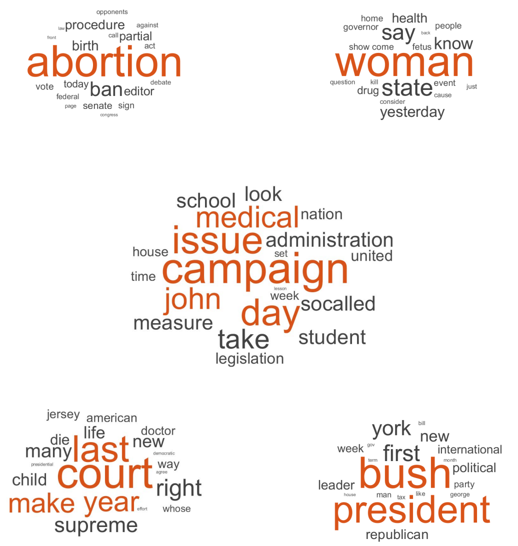


Fig. 11. Our estimated clusters for the domain *abortion*. Each cluster is said to represent a unique *frame*. The frame discussed in cluster 1 (characterized by the terms "abortion" and "ban") concerns a proposed ban on abortion. We analyze this cluster and find that our estimated PMC coincides with the period immediately preceding the Partial Birth Abortion Act of 2003.

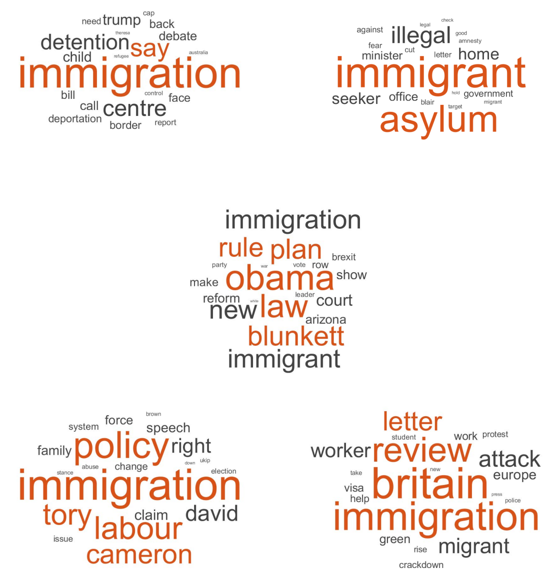


Fig. 12. Our estimated clusters for the domain *Immigration*. Each cluster is said to represent a unique *frame*. The frame of cluster 2 discusses the waning of asylum grants, increased border refusals, and the final 2002 white paper on "Secure Borders, Safe Haven." We analyze this cluster and estimate a PMC of 2000–2002 (see Fig. 13). Our PMC coincides exactly with the period immediately foreshadowing the government white paper.

F. Immigration

We study the framing of *immigration* news in the U.K. We obtained about 3600 articles on the subject of Immigration from the Guardian API for the period 2000–2017. For this domain, we carried out our analysis on the article titles (rather than the full text). Since the Guardian returns full-length articles, we found that this design choice allows us to produce a more focused domain corpus than the one generated by the full article text. We depict our estimated rating vectors and PMC in Fig. 13.

We analyze the frame of cluster 2 in Fig. 12. This cluster deals with the issue of asylum seekers to the U.K. In the period beginning immediately before the year 2000, a new peak in asylum claims to the U.K. of 76 040 had been reached [35]. This event coincided with a high-profile terrorist act by a set of Afghan asylum seekers [35].

These events resulted in increased border refusals and the final 2002 white paper on "Secure Borders, Safe Haven." We estimate a PMC of 2000–2002 (see Fig. 13). Our PMC coincides exactly with the period immediately foreshadowing the government white paper. This corresponds to three true positives with no false positives or negatives.

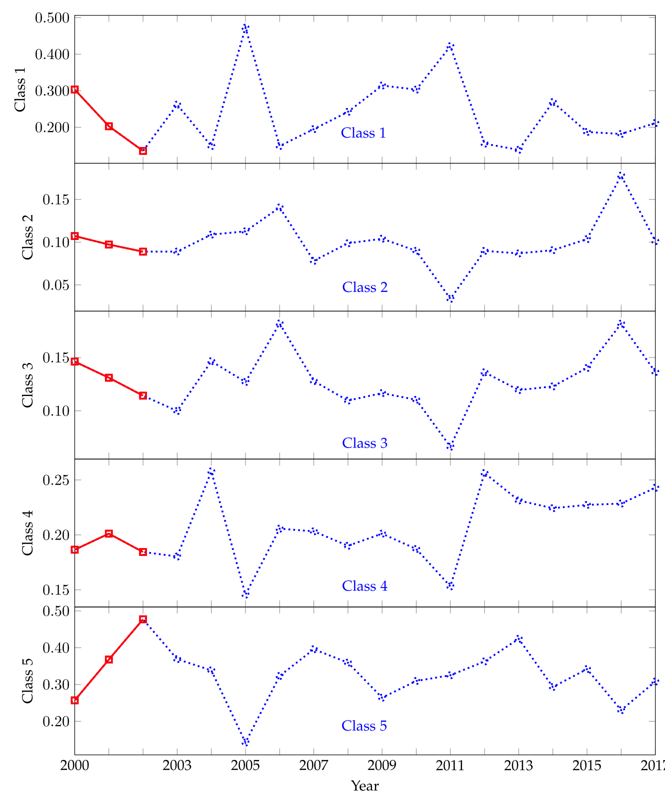


Fig. 13. Annual polarities for cluster 2 (discussing asylum grants) from Fig. 12 from the domain *immigration* for ratings 1–5. The PMC is shown with solid lines in square markers and foreshadows the "Secure Borders, Safe Haven" white paper of 2002.

G. Drones

We obtained nearly 4000 articles on this domain for the period 2003–2012. We obtain a PMC of 2009–2011 for this domain, as shown in Fig. 15.

Our PMC immediately foreshadows the Federal Aviation Administration's Modernization and Reform Act of 2012. This corresponds to three true positives with no false positives or negatives.

H. Predictive Utility

The aforementioned two domains (*immigration* and *drones*) highlight the predictive utility of news framing. While we did not find earlier surveys that coincide with our PMCs for these domains, we note that these PMCs foreshadowed substantial legislative activity. This observation suggests that

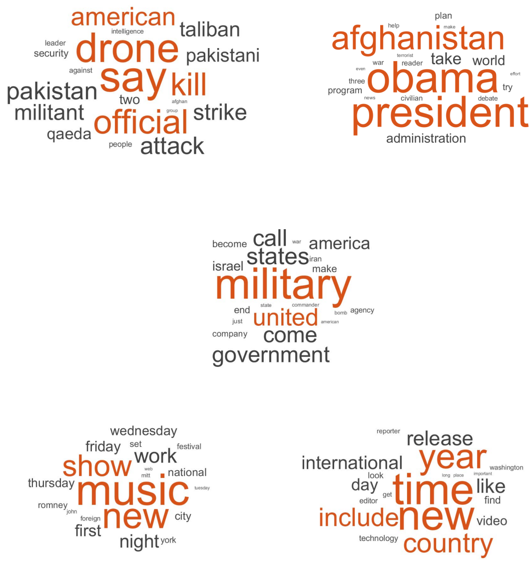


Fig. 14. Our estimated clusters for the domain *drones*. Each cluster is said to represent a unique *frame*. The frame discussed in cluster 1 concerns the use of drones against terrorist targets. Our analysis of this cluster returns a PMC of 2009–2011 (see Fig. 15). Our PMC immediately foreshadows the Federal Aviation Administration's Modernization and Reform Act of 2012.

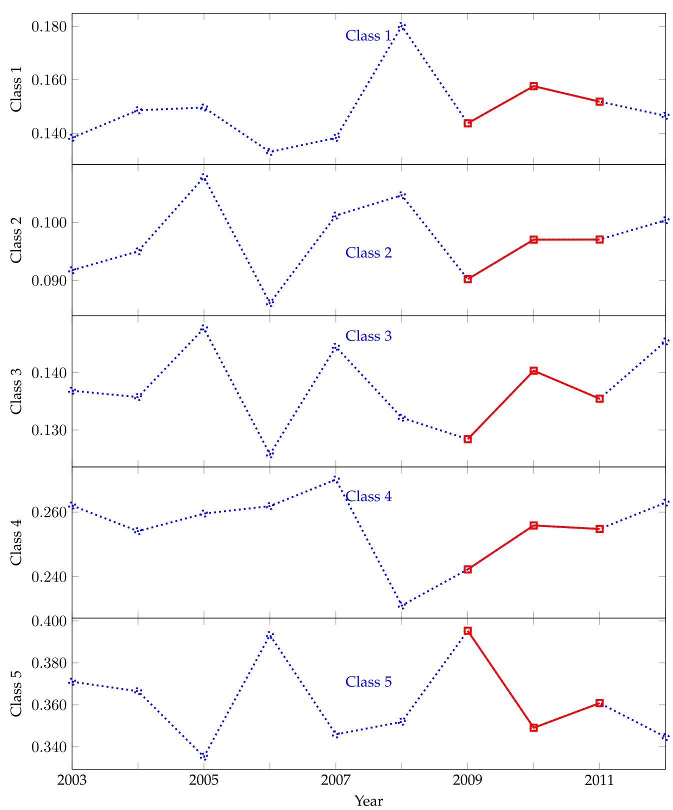


Fig. 15. Annual polarities for cluster 1 (discussing drone strikes) from Fig. 14 from the domain *drones* for ratings 1–5. The PMC is shown with solid lines in square markers and immediately foreshadows the Federal Aviation Administration's Modernization and Reform Act of 2012. This finding suggests the predictive utility of framing change detection for legislative activity.

PMCs estimated through real-time monitoring of domain news may yield predictive utility for legislative and commercial activity.

I. Overall Precision and Recall

We obtain an overall precision of 0.95 and a recall of 0.95. Our results demonstrate that we successfully identify 95% of true positives, whereas only 5% of the positives that we identify are false positives.

Furthermore, we point out that our false positives generally either precede or succeed a ground-truth framing change. Therefore, we posit that such false positives may be due to imprecision in measurement rather than any considerable failure of our approach.

Our results demonstrate substantial agreement with ground truth in domains for which prior surveys have studied framing changes. In domains for which we did not find such surveys, we demonstrate that our PMCs foreshadow periods of substantial public and legislative import. We posit, therefore, that our approach successfully identifies framing changes.

V. Discussion and Limitations

We describe certain limitations of our work and point out some suggestions for future work. First, as it is well-known [14], the English language is complex, and noun–adjective pairs are limited in their ability to capture semantic variations, such as sarcasm and context-specific polarities. Future work could attempt to estimate framing from the rhetorical structure and other similar approaches to attempt to capture more linguistic variation.

Furthermore, our approach does require several parameters, such as the number of frames, the number of nouns in each frame, and the threshold on adjective rarity. Future work could attempt to automate these thresholds using machine learning on a larger data set gathered from a wider temporal and geographical range. We note, however, that doing so would require enormous effort for manual labeling.

Finally, our work is restricted to polarized framing. While it is true that framing changes are usually associated with polarity drifts, such drifts are not a necessary condition for a framing change to occur. Changes that do not manifest as drifts in polarity would not be detected by our approach. However, based on the existing literature [7], [8], [16], [36], we conclude that such framing changes are the exception rather than the norm.

VI. Conclusion

We highlight a problem of significant public and legislative importance, and framing change detection. We contribute an unsupervised natural language processing approach that detects framing change trends over several years in domain news publishing. We identify a key characteristic of such changes; namely, during frame changes, the polarity of adjectives describing co-occurring nouns changes cumulatively over multiple years. Our approach agrees with and extends the results of earlier manual surveys. While such surveys depend on human effort and are, therefore, limited in scope, our approach is fully automated and can simultaneously run over all news domains. We contribute the *Framing Changes Data Set*, a collection of over 12 000 news articles from seven *domains* in which framing has been shown to change by earlier surveys. We will release the data set with our paper. Our work suggests the predictive utility of automated news monitoring, as a means to foreshadow events of commercial and legislative import.

Our work represents one of the first attempts at computational modeling of framing and framing changes. We, therefore, claim that our approach produces promising results, and it will serve as a baseline for more sophisticated analysis over wider temporal and geographical data.

References

[1] K. Gunnars. (2015). *Ten Causes of Weight Gain in America*. [Online]. Available: https://www.healthline.com/nutrition/10-causes-of-weight-gain#section12

[2] S.-H. Kim and L. Anne Willis, "Talking about obesity: News framing of who is responsible for causing and fixing the problem," *J. Health Commun.*, vol. 12, no. 4, pp. 359–376, Jun. 2007.

[3] K. Flegal, M. Carroll, B. Kit, and C. Ogden, "Prevalence of obesity and trends in the distribution of body mass index among us adults, 1999–2010," *J. Amer. Med. Assoc.*, vol. 307, no. 5, pp. 491–497, 2012.

[4] Wikipedia. (2020). *Wikipedia Article on Framing*. [Online]. Available: https://en.wikipedia.org/wiki/Framing_(social_sciences)

[5] D. Chong and J. Druckman, "Framing theory," *Annu. Rev. Political Sci.*, vol. 10, pp. 103–126, 2007.

[6] C. H. Vreese, "News framing: Theory and typology," *Identifying Inf. Tenor texts*, vol. 13, no. 1, pp. 51–62, Apr. 2005.

[7] K. Sheshadri and M. P. Singh, "The public and legislative impact of hyperconcentrated topic news," *Sci. Adv.*, vol. 5, no. 8, Aug. 2019, Art. no. eaat8296.

[8] G. King, B. Schneer, and A. White, "How the news media activate public expression and influence national agendas," *Science*, vol. 358, no. 6364, pp. 776–780, Nov. 2017.

[9] S. Alashri, J.-Y. Tsai, S. Alzahrani, S. Corman, and H. Davulcu, "'Climate Change' Frames Detection and Categorization Based on Generalized Concepts," in *Proc. 10th IEEE Int. Conf. Semantic Comput. (ICSC)*, Feb. 2016, pp. 277–284.

[10] R. D. Benford and D. A. Snow, "Framing processes and social movements: An overview and assessment," *Annu. Rev. Sociol.*, vol. 26, no. 1, pp. 611–639, Aug. 2000.

[11] G. Lakoff, *The All New Don't Think Of An Elephant!: Know Your Values And Frame The Debate*. Victoria, BC, Canada, Chelsea Green, 2014.

[12] D. Schon and M. Rein, *Frame Reflection: Toward Resolution Intractrable Policy Controversies*. New York, NY, USA: Basic Books, 1995.

[13] S. Baccianella, A. Esuli, and F. Sebastiani, "SentiWordNet 3.0: An enhanced lexical resource for sentiment analysis and opinion mining," in *Proc. 7th Int. Conf. Lang. Resources Eval.* Valletta, Malta: European Language Resources Association, May 2010, pp. 2200–2204.

[14] R. Socher *et al.*, "Recursive deep models for semantic compositionality over a sentiment treebank," in *Proc. Conf. Empirical Methods Natural Lang. Process.* Seattle, WA, USA: Association for Computational Linguistics, Oct. 2013, pp. 1631–1642.

[15] D. Card, A. E. Boydstun, J. H. Gross, P. Resnik, and N. A. Smith, "The media frames corpus: Annotations of frames across issues," in *Proc. 53rd Annu. Meeting Assoc. for Comput. Linguistics*, 2015, pp. 438–444.

[16] R. M. Entman, "Framing: Toward clarification of a fractured paradigm," *J. Commun.*, vol. 43, no. 4, pp. 51–58, Dec. 1993.

[17] O. Tsur, D. Calacci, and D. Lazer, "A frame of mind: Using statistical models for detection of framing and agenda setting campaigns," in *Proc. 53rd Annu. Meeting Assoc. Comput. Linguistics*, 2015, pp. 1629–1638.

[18] NYT. (2016). *Developer APIs*. [Online]. Available: http://developer.nytimes.com/

[19] Wikipedia. (2001). *The New York Times*. [Online]. Available: https://en.wikipedia.org/wiki/The_New_York_Times

[20] T. Guardian. (2016). *Guardian Open Platform*. Accessed: Mar. 3, 2016. [Online]. Available: http://open-platform.theguardian.com/

[21] Wikipedia. (2002).*The Guardian*. [Online]. Available: https://en.wikipedia.org/wiki/The_Guardian

[22] K. Sheshadri, N. Ajmeri, and J. Staddon, "No (Privacy) news is good news: An analysis of new york times and guardian privacy news from 2010–2016," in *Proc. 15th Annu. Conf. Privacy, Secur. Trust (PST)*, Calgary, AB, Canada, Aug. 2017, pp. 159–167.

[23] D. N. Hopmann, R. Vliegenthart, C. De Vreese, and E. Albák, "Effects of election news coverage: How visibility and tone influence party choice," *Political Commun.*, vol. 27, no. 4, pp. 389–405, Oct. 2010.

[24] F. Baumgartner, S. D. Boef, and A. Boydstun, *The Decline Death Penalty Discovery Innocence*. Cambridge, U.K.: Cambridge Univ. Press, 2008.

[25] M. Abrajano and Z. Hajnal, *White Backlash: Immigration, Race, and American politics*. Princeton, NJ, USA: Princeton Univ. Press, 2017.

[26] G. de Melo and M. Bansal, "Good, great, excellent: Global inference of semantic intensities," *Trans. Assoc. Comput. Linguistics*, vol. 1, pp. 279–290, Dec. 2013.

[27] D. M. Blei, A. Y. Ng, and M. I. Jordan, "Latent Dirichlet allocation," *J. Mach. Learn. Res.*, vol. 3, pp. 993–1022, Mar. 2003.

[28] J. Benesty, J. Chen, Y. Huang, C. Yiteng, and I. Cohen, "Pearson correlation coefficient," in *Noise Reduction in Speech Processing*. New York, NY, USA: Springer, 2009, pp. 1–4.

[29] Mathworks. (2019). *Correlation Coefficients*. [Online]. Available: https://www.mathworks.com/help/matlab/ref/corrcoef.html

[30] National Cancer Institute. (2019). *How the News Media Influence Tobacco Use*. [Online]. Available: https://cancercontrol.cancer.gov/brp/tcrb/monographs/19/m19_9.pdf

[31] P. Vernon. (2018). *Five Years Ago, Edward Snowden Changed Journalism*. [Online]. Available: https://www.cjr.org/the_media_today/snowden-5-years.php

[32] Pew Research. (2016). *The State of Privacy in Post-Snowden America*. [Online]. Available: http://www.pewresearch.org/fact-tank/2016/09/21/the-state-of-privacy-in-america/

[33] K. M. Cummings and R. N. Proctor, "The changing public image of smoking in the United States: 1964–2014," *Cancer Epidemiology Biomarkers Prevention*, vol. 23, no. 1, pp. 32–36, Jan. 2014.

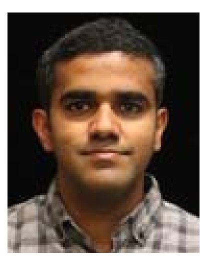
[34] J. Gainous and L. Rhodebeck, "Is same-sex marriage an equality issue? Framing effects among African Americans," *J. Black Stud.*, vol. 47, no. 7, pp. 682–700, Oct. 2016.

[35] Wikipedia. (2019). *History of UK Immigration control*. [Online]. Available: https://en.wikipedia.org/wiki/History_of_UK_immigration_control

[36] F. Baumgartner, B. Jones, and P. Mortensen, "Punctuated equilibrium theory: Explaining stability and change in public policymaking," *Theories Policy Process*, vol. 8, pp. 59–103, 2014.

Karthik Sheshadri (Member, IEEE) received the Ph.D. degree from North Carolina State University, Raleigh, NC, USA, in 2019.

He is currently a Senior Research Engineer with Cisco Systems, San Jose, CA, USA, where he works on applying NLP and graph database techniques to problems in networking. His research interests include computational social science, natural language understanding, and privacy.



Chaitanya Shivade received the Ph.D. degree from Ohio State University, Columbus, OH, USA, in 2016.

He is currently an Applied Scientist with Amazon, Seattle, WA USA, where he works on applying machine learning and natural language processing to biomedical data. Prior to joining Amazon, he served as a Research Staff Member with IBM Research, Almaden, CA, USA.

Munindar P. Singh (Fellow, IEEE) is currently a Professor in computer science with NC State University, Raleigh, NC, USA, and the Co-Director of the Science of Security Lablet. His research interests include social computing and AI ethics.

Prof. Singh is a fellow of AAAI and AAAS. He was a recipient of the ACM SIGAI Autonomous Agents Research Award and the IEEE TCSVC Research Innovation Award. He is also a former Editor-in-Chief of IEEE Internet Computing and *ACM Transactions on Internet Technology*.