Exploring Double Subjectivity Through News Frames In Online News Sources: *A Network Approach*

Loretta H. Cheeks
CIDSE, Computer Science
Arizona State University
Tempe, AZ U.S.A.
loretta.cheeks@asu.edu

Ashraf Gaffar
CIDSE, Software Engineering
Arizona State University
Mesa, AZ U.S.A.
ashraf.gaffar@asu.edu

Abstract**—***The Internet is a major source of online news content. Current efforts are made to unlock latent meaning in online news content using advanced language processing tools and machine intelligence. This necessitates exploring the internal structure of news narratives to cope with the challenges posed by limitations of existing tools. This article explores the conceptualization of Double Subjectivity in news frames as deployed by online news sources. We propose a new perspective by exploring a) a News Frame Issues Network that is useful for describing the structure of online news media and b) formulating an influence model for understanding the dynamics of bias that underpins Double Subjectivity. This research has the potential to inform more intelligent conclusions about narrative text meaning (or semantics) to address real-world socio-environmental issues. We use water insecurity in the Southwestern United States as our contextual case. Our experimental evaluation shows the proposed network and model is an effective approach for advancing what we know about the production of language in narrative text where subjectivity exist.*

Keywords *–Subjectivity; sentiment; sentiment analysis; social influence; bias; subjectivity; framing; news source; natural language processing*

I. Introduction

Approximately 80% of all data today exists in digital form as unstructured text (e.g., news, e-mails, social media feeds, contracts, memos, clinical notes, and legal briefs) [14]. The Internet is the premier digital platform for the dissemination of online news content and unstructured text. The ability to unlock hidden structure and latent meanings in unstructured text is an important area of research because text is a fundamental device of communication and human interaction for expressing real-world issues.

The value of knowledge depends on communication among other things. Data is one common medium of capturing and communicating valuable information and hence knowledge. While computers have been used to leverage our ability to communicate and acquire knowledge, it's mostly an automated approach. An approach of the generation and dissemination of data with less ability to delve deeper into the hidden meanings of data. Humans still have an unprecedented ability to get beyond the syntax of data.

The automation of data processing can be generally categorized under two main areas: formal data and informal data.

Formal data represented by mathematical formulas and computer languages have little room for subjective interpretation, and is typically concise and predictable. A correctly developed source code can compile and run on a computing machine and will always produce the same output for a specific input. The combination of the formal set of instructions resulting from successful compilation and run of source code will transition the computer through a set of predefines states, depending exclusively on the compiled code and input data. This leaves no room for interpretation or subjectivity. In other words, the meaning and interpretations must be fully represented in the source code. A good programming language must therefore have sufficient expressiveness to fully represent a feasible solution in a specific problem space. If the source code arrived to an undefined state, it will simply freeze or crash, and is unable to make any independent decision. Some areas of computer science like fuzzy logic and AI attempt to address these inherent limitations. However, it's not the focus of our work.

While source code is a good medium of communication with computers, natural languages are the medium of communication among humans in an informal way. Such informal data can be further divided into two subcategories depending on the source of origination: direct and indirect.

Direct informal data arises when a person encounters an event and perceives, in part, its details firsthand. Human perception of events leads to their personal interpretation of the event. This interpretation depends greatly on the viewer perception as different people can view different prospects of the same event. Perception is often associated with the viewer's state of mind, prejudice, background, and beliefs. Two persons witnessing the same event or being involved in the same situation could assign different interpretations and end with separate conclusions. We believe that while the external situation was the same (we define it as the Absolute Baseline), it could be internalized differently by different people. We define this case as first level subjectivity. Even if different

978-1-5090-4284-5/17 $31.00 © 2017 IEEE
DOI 10.1109/ICSC.2017.50

people were asked to report an incident or a situation (the same Absolute Baseline), we often end up with different views. Many factors further affect this subjectivity.

- Complex situations, such as watching and reporting a movie, are harder to perceive and interpret compared to simple or atomic events, like an incident of a rocket launch.
- Unlike formal languages, the medium of communication, a natural language, does not fully express the contents in its syntax. Natural languages report "about" something, rather than fully describes the event.

Beyond natural language and direct communication, the invention of printing, and multimedia was followed by the Internet and the WWW which led to an explosion of indirect informal data. While we can still directly perceive some events and build our own subjective interpretation of them. We can then receive a much larger number of events indirectly. We do not witness the events; hence we do not have an Absolute Baseline. We rather "hear about the events" through written or spoken reports. Our baseline will hence be relative to the reporter subjectivity or simply "float".

While we still perceive the indirect reports and interpret them differently (adding our subjectivity), we lost the Absolute Baseline, and we start with a reporter-relative or a floating baseline. The difference here is that if the reporter subjectivity is strongly biased towards certain beliefs, we already start with a completely different baseline that could strongly affect our own beliefs. We refer to it as "Double Subjectivity".

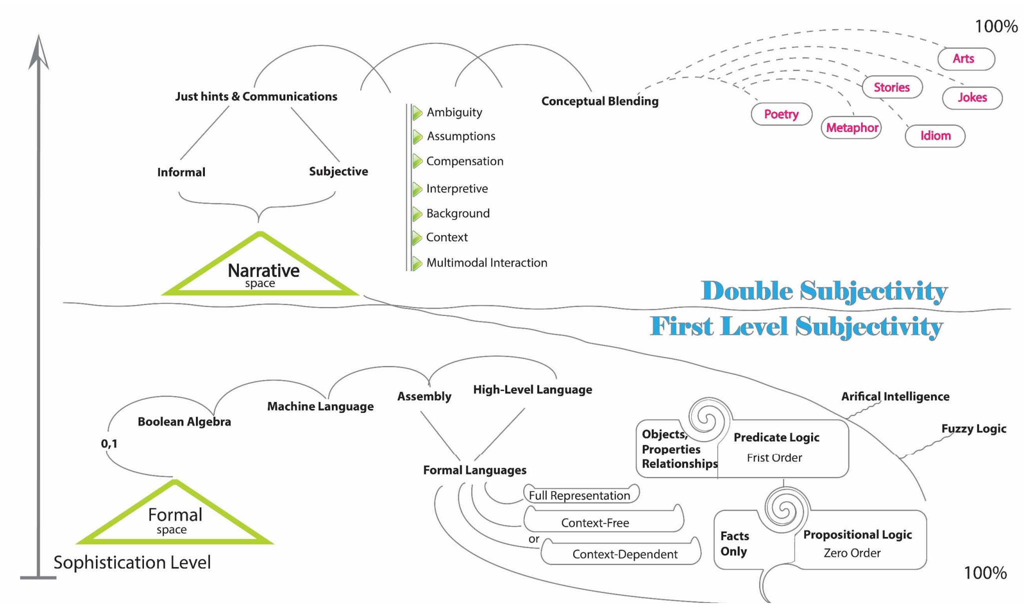

Figure 1. A depiction of the evolution of First Level Subjectivity and Double Subjecting potentials.

- Humans have different expressive abilities and styles in reporting events using natural languages, leading to increase in subjectivity.
- The ambiguity of natural languages allows for different interpretations of the spoken or written sentences.
- Therefore, even for different people directly witnessing the same situation, significant subjectivity can be present in reporting or discussing it amongst themselves. This is seen in long debates and discussions or exchanges in correspondences.

This first level subjectivity has been identified and we often try to reduce it. For example, TV reporters try to abide by certain rules to maximize neutrality and leave it up to the viewers to build their own opinions freely. The scientific community tries to prove the stated facts through rigorous research and scientific experiments in an attempt to eliminate or reduce subjectivity. However, this is easier said than done. We've seen reporters who are known for their "opinions." Even some scientific research based on rigorous experimental results was later refuted after being followed by more experiments.

An analogy drawn from photocopying an original document illustrates "Double Subjectivity" in that, ideally, a photocopy should be a replica of the original with no distortion (with distortion being metaphorically equivalent to bias). Yet, in actuality, the resulting photocopy entails some distortion. Two photocopies of the original will not be the same as a photocopy of a photocopy. The latter case will have a cumulative distortion effect similar to that in Double Subjectivity.

Online news is exemplary of "Double Subjectivity," in that its assemblage consists of subjective properties; such as, context, interpretation, multi-modal interaction, background, compensation, and assumptions. These properties are critical building blocks that characterizes narrative text for its a) signification, i.e., the semantic content signifying an association or concept given a word, sentence, or phrase and b) significance, i.e., the relevance, rank, importance, or capacity to make a difference. Unfortunately, there remain significant gaps for advancing what we know about the production of language in narrative text where “Double Subjectivity” exists. Therefore, exploring deeper levels of expressivity with an emphasis on “Double Subjectivity” may inform a new direction in research.

Learning “Double Subjectivity,” in narrative text poses a difficult problem, as it seeks to give insight and meaning to open-ended statements and indirect data. More complex structures within the narrative text provide an expansive landscape for advancing effective tools and crafting applications to automate many language-related tasks. For example, document summarization, automated text generation, and many others. To show the gains in logic and reasoning, Fig. 1 depicts advances made using current tools and new frontiers that Double Subjectivity may shed light on, such as, poems, metaphors, idioms, jokes, art, and storytelling. This is to be achieved in exploring levels of narrative text.

The Double Subjectivity framework may be useful for learning some embedded internal structures of unstructured narrative text that contain valuable embedded information cues and biases. The human ability is unprecedented, but unscalable in text comprehension. Currently, much of the online news and other forms of narrative text (e.g., studies, earning call transcripts, annual reports, letter to shareholders, and disclosures statements) remains untapped due in large part to human limitations of readability. Moreover, the Double Subjectivity framework visualization will expose the dynamics of migration paths of narrative text, the bias intensity, and the intersection of forces where shifts in beliefs are realized.

We seek to give insight into the following questions in this research inquiry: What is the migration path of online news frames evolving over time? What is the effect of “Double Subjectivity?” What are the conditions observed for determining shift in beliefs?

In this manuscript we cover an examination of online news framing of articles (i.e., a form of narrative text) for exposing a deeper level of expressivity we coin Double Subjectivity. It is our aim to demonstrate news framing as a step toward a deeper learning about the production of language that may inform scholars about machine intelligent behavior.

To the best of our knowledge, examining Double Subjectivity using online news and framing effects has not previously been addressed for this class of learning problem in computer science scholarly research. The following sections will describe the methods, observations, and a discussion.

II. Related works

A. *Text Data Mining, Text Summarization, Information Retrieval*

Scholarly research in computer science on latent meanings in association of terms and documents to reveal relationships is found in literature related to text summarization, information retrieval, and text data mining. The earliest paper on text summarization is that of [15] that describes work done at IBM in the 1950s. In his work, [15] proposed that the frequency of a particular word provides a useful measure of its significance and identified the concept Term Frequency (TF), which states it is possible to identify significant terms solely based on the term calculated frequency of the term within that document. It relates to average information, or entropy, of a term or group of terms ranking in relationship to each other.

Other outstanding contributions to text data mining in the 1990s and beyond with the advent of machine learning include text representation and models construction [16]; [17]; [18]; [19]; [20]; data dimensions reduction research in feature extraction [21]; [22]; and deep semantic mining [23].

Our research of Online News Frames is motivated by framing theory which focuses on understanding the latent meanings of observable messages in their contexts [24], and can provide important insight into how the presentation or “framing” of an issue affects the choices people make. Other disciplines have focused on framing: in linguistics research, similar approaches are also described as “latent semantic analysis” (LSA) [25]. Social Network Analysis (SNA) focuses on the importance of relationships among interacting units [26].

B. *Opinions and Social Influence*

While text data mining techniques and methods have proven to give insight to first-order subjectivity, we look to dynamical systems modelling for insights into Double Subjectivity that exist in narrative text. Models of social influence, cultural dynamics, and information diffusion (i.e., topic modeling, sentiment analysis, and opinion mining), are active areas of research [7], [1], [27]; [28]).

The opinions models, such as the voter model give insight into the spread and distribution of opinions [1]. According to [6] model, social impact theory, it posits the impact of a social group on an agent as being dependent on the prominence of the social sources, their proximity, and on source group mass (i.e., the number in the group). Political opinions and Axelrod’s cultural dynamics model behave similar to these models that capture the interplay between selection and influence [2], [3], [4], [5].

III. Backgroung On The Issue

Water is a fundamental resource affecting all aspects of life on earth. Water is used for human consumption, industrial processes, production of food, sanitation, as well as other usages. The way water policy and water decisions are framed affects water rights allocations, policy decisions, human consumptions, emerging technologies, farming techniques, and agricultural outcomes [29]. The issue of water insecurity in the Southwest Region is particularly important in that the seven states that make up the Southwest Region (Arizona, California, Colorado, Nevada, New Mexico, Utah, and Wyoming) rely primarily on fresh groundwater flows originating from the Colorado River. The Colorado River was an efficient source of water for decades, however, due to a decade of drought, this once plentiful source of water cannot meet the demand to sustain life as it exists in this region. This contributes to water insecurity throughout the Southwest Region of the United States.

When online news sources use strategic devices for presenting prominent aspects and perspectives about an issue using certain keywords as well as stereotyped images and sentences for the purpose of conveying latent meanings about an issue, it is called framing [30]. Water insecurity in the Southwest Region of the United States exemplifies an issue that is appropriate for study through the lens of Double Subjectivity that exists in framing by online news sources because this issue is context specific, complex and characterized by uncertainty. Newspapers in the Southwest regularly produce in-depth articles about drought, water and climate change that are published online. However, the general public outside of the Southwest has very little experience with water insecurity. Biased framing or terminology are more likely to influence uninformed respondents [31] or respondents with reduced exposure to or interest in an issue [23]. Therefore, citizens' attitudes and beliefs about water insecurity are likely influenced by the way reporters frame this issue. This source of unstructured text offers researchers a body of content in which to explore the production of language in the context of a socio-environmental issue. The issue of water insecurity is a case study needed to explore the patterns of interaction between online news sources and the effects of their frame choices. This research can generally apply to any large, unstructured dataset which is affected (or biased) by framing. It can therefore be extended to other applications such as economics or political discourse.

IV. METHODS

Online news framing of critical socio-environmental issues is a complex and dynamic social network, which can be studied as a graph. Graph analysis has become important in understand the dynamical process in the production of news frames. Graphs allows for the visualization of the social network that show interdependency of actors (or nodes) in terms of the social relationships, such as friendship, kinship, or financial exchange [33]; [34]. The presence of the phenomenon of influence appears as a dynamic process for the formation and transmittal of group standards, values, attitudes, and skills as accomplished in online news communications [35]. Influence is the tendency of people to become similar through endogenous interactions [28]. This phenomenon coupled with the vast amounts of online news articles being produced every second poses a potential challenge when seeking to gain knowledge about critical issues. Graphs represent objects where order and disorder co-exist. Graphs prove well for showing social interactions, influence, migration paths, and framing effects. For instance, a graph may expose news sources which hold central positions that function as points of prominence, control, and stability and edges that act as highways for lead relationships and exchanges of news frames. These intertwined dynamics coupled with the vast amounts of online news being produced daily, makes graphs an important tool for visualization and analysis of information flows pertaining to critical issues. Furthermore, the output of our graph will be a human readable network visualization, which show the overall spatial-temporal behavior of the paths. This visualization will enable one to see at a glance the cultural dynamics and social influences and hence make better informed decisions. Having a visual tool that automatically processes vast amounts of online news will expose policy implications, risk, and public perceptions over the short and long term; thereby offering a mechanism for adjusting policies in a positive direction.

A. Dataset Description

The news data for our study was collected from Google News1, which is a news feed aggregator. Our feature selection comprises four different characteristics of a given article. Namely a) the news articles being published by online news sources, b) the news source that generate the articles, c) the frequency of news publications by sources over time, and d) the salient terms contained in the article.

We restrict the time period from which the news articles published on the Internet by online news sources. Which for this research, was January 2006 to March 2015.

Articles were gathered by limiting the search for keywords associated with water or sub-topics of water consequences within Arizona, Colorado, California, and Nevada. 55,000 news articles were collected and the articles were then stored in a database for undergoing data pre-processing that included removing errors and inconsistencies in order to improve the quality of data. After data pre-processing, 30,000 articles were deemed relevant for our news dataset and 25,000 were deemed irrelevant. We plan to publish our dataset in Fall 2017 to the "Data Exchange Network[2]."

B. News Frame Issues Network Construction

We start the construction of our news frame issues network by defining our nodes; which we consider as our edge connections between two nodes, and the boundary of the network in question, i.e. what relations and what nodes should be included in the dataset.

In this research we chose news articles being published by online news sources and the news sources that generate the articles to serve as our nodes. We reason that news sources play a critical role in the propagation of bias about issues and bias in online news drive a wedge between evidence and beliefs [36]. In total, 280 news sources were identified to potentially influence public perceptions of the issue of water insecurity using strategic framing. Coverage of elites in articles is a property of particular interest in this study, as this factor has an effect on the bias intensity that news sources adopt over time. According to Druckman (2001), the way elites frame an issue is a driving force for shaping public opinion. The boundary of our network is water or sub-topics of water consequences within Arizona, Colorado, California, and Nevada, as shown in Fig. 2.

[1] http://news.google.com/

[2] http://www.insna.org/connections.html.

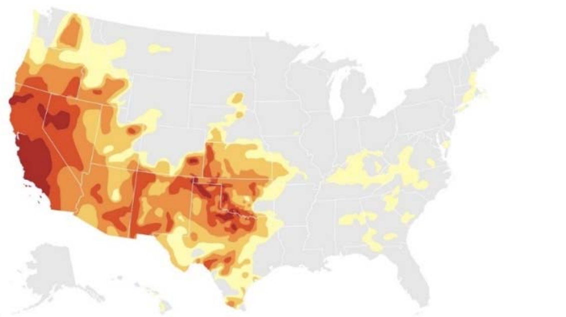

Figure 2. Drought severity impacting the U.S. Southern Region. Source: National Climatic Data Center, National Drought Mitigation Center, U.S. Department of Agriculture, National Oceanic and Atmospheric Administration.

Since water is a fundamental resource affecting all aspects of life, we can expect to see a relationship between the number of articles being produced on the issue as an indicator of its' prominence and the framing choices of the news sources. Which is key for driving policy and decision-making and shaping attitudes, values, and bias. We make connection ties from one news source to another through the capture of framing adoption overtime as it happens during migration. Though we recognize the value of reciprocal relationships (i.e., acknowledgement of both news sources as using the same frame choice), this type of hand shaking is a challenge to capture using the article content itself. This is largely due to news sources not crediting other news sources in their reports. The resulting network aim is to capture structures that were created and maintained through repeated patterns of framing and its migration path overtime.

C. *Problem Formalization of News Frame Issue Network*

The overarching research problem is expressed through the construction of an issues network defined as follows:

We model a news frame issues network as a graph $G = ({S_g}^t, {A_h}^t, {E_r}^t)$ consisting of a) S_g that denote a set of source nodes containing $g = \{s_1, s_2, \cdots, s_g\}$ elements and b) A_h is a set of article nodes containing $h = \{a_1, a_2, \cdots, a_h\}$ elements that operate within the time window t. In general, t indicates one discrete time step of a reporting period by S_g. Let $S_g = \{C_i, F_i\}$ be the set of source properties, where a) C_i is the set of terms (i.e., sub-issue keywords or content-descriptors) and b) F_i is dominate frame choice (identity) the news source uses. Let R be the number of relations. Each relation has a corresponding set of edge connections, ${E_r}^t$, directed/undirected as edge elements. The issues network, subscript $r = 3$ correspond to the total number of relations. The edge set represents the communication channels between node pairs. A_h are children of S_g, therefore, they may inherit the properties of the parent news source, as shown in Fig. 3.

The relations (or rules) for E_r edge connections are as follows:

<u>$\boldsymbol{R_1}$. A & A</u> (articles to articles). A non-directional edge connection is constructed for E_1 when $A_h(a) = \{a: (a_1, a_2, \cdots, a_h) \in E_1\}$ is the set of neighbors of $\boldsymbol{a_h}$ such that $\boldsymbol{a_h}$ are articles with the similar sub-issue.

<u>$\boldsymbol{R_2}$. A & S</u> (articles produced by the same source). A directional edge connection is constructed for E_2 when $S_g \in E_2 \rightarrow A_h(a) = \{a: (a_1, a_2, \cdots, a_h) \in E_2\}$ is the source (or producer) of the set of neighbors of $\boldsymbol{a_h}$ articles, where $\boldsymbol{a_h}$ is similar article sub-issues.

<u>$\boldsymbol{R_3}$. S & F</u> (sources with sub-issue to sub-issue with the same frame choice). A non-directional edge connection is constructed for E_3 when $S_g(s) = \{s: (s_1, s_2, \cdots, s_g) \in E_3\}$ is the set of neighbors of $\boldsymbol{s_g}$ such that $\boldsymbol{s_g}$ share the same dominant frame choice.

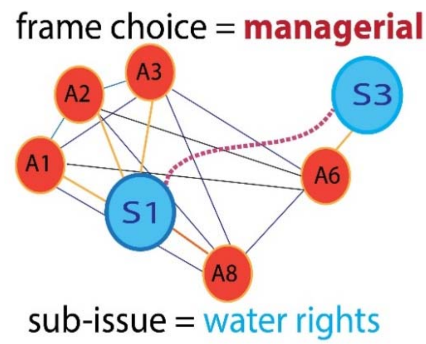

Figure 3. A subset of the News Frame Issues Network that comprises two news source nodes with connecting articles using the same sub-issue and using the same frame choice. Our visual design language is expressed as follows: a) blue circles around the notation S_g denotes online news sources node that produced the article and propagates a central organizing idea (i.e., framing), b) orange circles around the notation A_h denotes an online news articles node, c) thin black lines correspond to edge connections among neighboring articles covering the similar issue, d) thick golden lines correspond to edge connections of news sources who produce articles covering similar issue and e) thick red lines correspond to managerial news frame.

Problem: Given $G = ({S_g}^t, {A_h}^t, {E_r}^t)$ about I issue, construct a graph for determining how frames are produced in online news media over time.

The aim of constructing a news frame issues network using water insecurity as a contextual case is a) for deeper understanding of how framing advances novel knowledge discovery with sensitivity to "Double Subjectivity" as a factor in the production of language in text narratives and b) to explicitly examine the temporal dynamics of structural changes that may exist within the network.

The definition of our overarching research problem was defined: "How is the issue of water insecurity in the Southwest Region of the United States being framed in online news media over time?"

D. *Model Setup*

To model our news frame issues network, we propose formulating a new version of information feedback on a social system in the framework of [1] model for cultural dynamics with an emphasis on bias that emerges as "Double

Subjectivity" in narrative text. Axelrod's [7] model is built on two simple assumptions: a) people are more likely to interact with those who are more similar to them, and/or to be more receptive to influence from those who are similar—a mechanism of Selection and b) these interactions tend to cause similarity among interacting actors—Influence. Though our model behaves similar to Axelrod's [7] when deciding how to interact in the network, it differs in that we give insight into the endogenous and exogenous forces that may be at play in dynamic systems. Therefore, our Double Subjectivity social influence model, integrates cultural dynamics extensions [28] with cusp catastrophe model [38]. Cusp catastrophe model is useful for describing nonlinear relationships such as found in narrative text. We particularly focus on the different forces at work in the bias slant employed in news frames as it allows for the degrees of freedom whereby alternative pathways may be realized. The cusp catastrophe model allows for understanding the different forces at work when quantifying bias and the migration path of bias formation.

The integration of cultural dynamics and cusp catastrophe model for defining our new Double Subjectivity social influence model may shed insights on deeper levels of expressivity in narrative text where bias offers alternative pathways and possibilities. This remains an emerging research topic for understanding the production of language that may help in synthesizing vast amounts of unstructured text and learning online social behaviors. Here we explore the process news sources use when producing online news articles. The main factors we consider in this production are the facts, bias intensity, and signals that derive from neighboring news sources. The signals that derives from news sources is liken to the reporter-relative or a floating baseline as mentioned earlier. In the process of time, the news sources Absolute Baseline will undergo adjustments as the bias intensity change.

E. Double Subjectivity Social Influence Model

Online news sources tend to produce articles about an issue I of interest to society, such as water insecurity, the state of the economy, or affordability of health care. Additionally, news sources may receive signals β from other news sources on the importance of the issue, based on observed increase to the number of articles other neighboring news sources produce. They receive facts f about an issue. The news sources then, selects a Bias strategy α using as the basis of its calculation of the bias intensity. More specifically, each news source S_g with interest in I issue at time t generate news articles

$$\mathrm{I(g)} = \alpha_g + \beta_g(f) \qquad (1)$$

where α_g and β_g are the coefficients of the news source while f are the facts received. In the first interaction, set $\alpha_g = 0$, denoting bias intensity; this value will be calculated upon interacting with neighboring nodes and should change for giving shape to the news source Absolute Baseline. Time evolves in discrete steps t = 0, 1, 2, . . ., and $S_g(t)$ denotes the mass count on S_g at time t. The maximization of $\{S_g\}$ gives the sub-issue most important, M_i; a count of the largest cluster—the mass—of nodes reporting on a water sub-issue. This initialization of α_g and β_g provides access to new nodes to enter the overall conversation about water consequences by linking to the most prominent mass. Thereby, leveraging facts and neighboring nodes relative baseline. This is similar to operation of observed in Hidden Markov Model (HMM), whereby the news sources absolute baseline is hidden with adjustments made over time. In contrast, the relative baseline is likened to the observations one can make in the HMM for getting a sense of the state [13]. Furthermore, we imagine that α_g correspond to the endogenous propensity of the news source nodes to express its own bias (or Double Subjectivity), while β_g represent the exogenous force in the network.

At the start of the process, each news source $g \in S$ has a non-negative node mass associated with it, corresponding to the fraction of the news sources that initially report (i.e., through signaling) on similar sub-issues about water insecurity. The dynamical system allows for each news source to switch frame choice, $F(e)$, and sub-issue interest, thereby enabling random selection of news source interaction. As well, news sources are susceptible to be influenced when they interact with neighboring news sources who share similar bias intensity. The full state space may be calculated by counting the news sources as expressed as the mass vector $S_g(t)$.

F. Defining Bias Intensity for Expressing Double Subjectivity

We situate the quantification of Bias Intensity in terms of a mathematical catastrophe theory [38]. Catastrophe theory is a branch of non-linear dynamic systems theory that originates with the work of the mathematician Rene´ Thom [38] to help explain biological morphogenesis as one of the great mysteries confronting mathematical biology. A key property in catastrophe theory is that the system under study is driven toward an equilibrium through its use of gradient descent or potential function for seemingly automatic guidance (i.e., through the law of attraction) occurring in the system, which is important in research of social influence, particularly when considering the property of convergence and stable states. The cusp model is the most well-known and simple model of catastrophe theory positing that nonlinear transition from one state to another of a system are guided by two controlling variables, the asymmetry, and bifurcation factor.

Consider, for example, the social and economic inequality movement, Occupy Wall Street Movement. What began as small grassroots pressure points to address pay inequality by fast food workers and Walmart employees has transition to giving voice to 99% and moving the political conversation in the U.S. election. This transition, from a grassroots protest in Manhattan's Zuccotti Park to making inequality and the wealth gap the core of the political race, is a catastrophe.

Let $M = \{Managerial, Economic, Conflict\}$ and $S = \{Human\ Interest, Science\}$ be the aggregate count of the sets of frame choices obtained from the edge connection. S_g is assigned the Frame Identity corresponding to their F_i property at time t.

Definition: $\boldsymbol{z}$ is *Bias*, a projection on the behavior surface for predicting the migration path that show the gradual shift in bias intensity of the online news source when shaping the frame narrative about the issue of water insecurity. This factor is a measure of the strength of bias, that is, “weak bias (safe)” or “strong bias (insecure or harmful)”—respectively *S(ecure)* and *I(nsecure)*.

Definition: $\boldsymbol{x}$ is the asymmetry control (or normal) factor x. This factor receives the label *Justice*, denoted on a scale that range from $J(ust)\ to\ U(njust)$. Justice is a measure of the news source perception of fairness of water consequences.

Definition: $\boldsymbol{y}$ is the splitting factor or the bifurcation factor. This factor receives the label *Frame Identity*, denoted as $S(trong)\ and\ W(eak)$.

We argue that, *Frame Identity* and the perception of *Justice* are key forces for predicting the Bias Intensity of news sources when producing articles. Naturally, Frame Identity is one form of bias, because of the strategic devices for presenting prominent aspects and perspectives about an issue using a strong slant for the purpose of conveying latent meanings about an issue [30]. We view Justice as being of equal importance as it captures the perceived sense of fairness about an issue. Fig. 4 provides a depiction of the News Source Bias Intensity

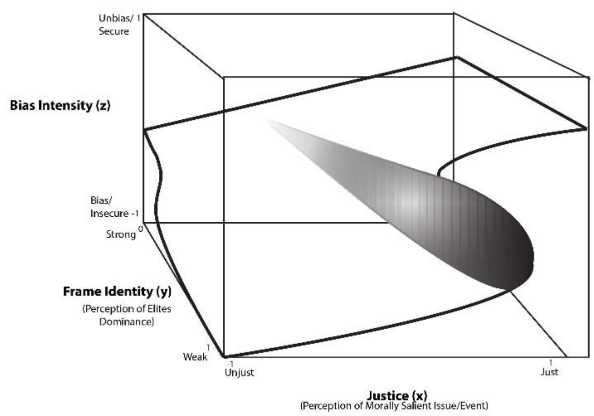

Figure 4. We depict the Cusp Catastrophe Model for measuring the shift in News Sources Bias Intensity.

following cusp catastrophe model for predicting gradual shifts of views and showing the possibility space. A discretized approach to opinion research has been conducted in scholarly research [39] and applications of cusp catastrophe [40].

V. POSSIBILITY SPACE TRANSITIONS

Nonlinear relationships represent a variety of possibilities and pathways of which the linear case is but a limiting one. Here we propose expressing our possibility space using three variables; a) *Frame Identity*=[Index-1 ((S, S(trong), 0), (W, W(eak), 1))], b) *Justice*=[Index-2 ((J, J(ust), 1), (U, U(njust), -1), (0, N(eutral), 0))], and *Bias Intensity*=[Index-3 ((S, S(ecure), 1), (I, I(nsecure), -1), (0, N(eutral), 0))]. Index-1 represents the control variable Frame Identity, which is the perceived dominance of elites in the reporting the news. Index-2 represents the sense of Justice in the handling of the issue. Index-3 represents the dependent variable Bias Intensity, which has the effect of signaling secure (or safe) or insecure (or harmful) issue.

We combine Index-1, Index-2, and Index-3 to denote state transitions for updating shrinkages and growth. For instance, SJS denotes strong frame identity (S), a sense of justness (J), and a sense of security or being safe (S).

VI. OBSERVATIONS OF THE MODELS

A. Influence Dynamics and Migration Path

Traditional cultural dynamics models show the network sizes are power law distributed. We integrate models of cultural dynamics with the cusp catastrophe model to explore Double Subjectivity through quantifying the bias intensity. Our analysis show this measure has potential to change the network structure as the heavy reliance on prominent mass count is not the only law of attraction at work.

Though our model is posed to scale to full dataset capacity, we focus on a small subset of the network for this preliminary evaluation. Our model allows for news sources generating articles about water consequences to interact freely with other news sources. However, because news sources receive signals that provide a relative-baseline—leveraging other news sources opinions, interpretations, and perceptions—it is most likely that a news source will choose interactions with those that are similar. There is a chance for a news source to interact with another news source based on an increase in their bias intensity, even though there exist no connection. And there is a higher chance that news sources will only interact with those whom they’re connected with, meaning breaking out of the cluster is caused by catastrophic shifts in bias.

Further analysis of the Bias Intensity function, a variant of cusp catastrophe model, shows interesting behavior for making decisions about trust and distrust. We note that any space on the surface response represent the state of the news source under observation. For instance, as the news source perceive the issue will not lead to fair water consequences, as they move toward the cusp, this will trigger them to start attributing blame toward the population who is causing the harm. A contextualization of this case with water insecurity would show, as a news source who has ties to nodes employing scientific frame choice perceive unfairness (unjust), they will notice the population more aligned with elites are becoming increasingly bullish in their handling of the water and are attracting other news sources (i.e., homophily), because their narrative is spreading. At this point when feelings of unfairness intersect with the dominance of the elites, they will start scanning for those to attribute blame; thereby entering a space of distrust and strong bias. This kind of model exhibit hysteresis, thereby making it hard to shift between surface response planes, as the migration path will be different. We will give treatment to other hypotheses associated with bias that lead to trust, as a result of perceived positive water consequences in future work.

VII. CONCLUSION AND OPEN QUESTIONS

We have presented a new concept that allows for an adjustment of the absolute baseline of a news source (or agent) that takes into account ones' bias for expressing Double Subjectivity through news frames.

We present in this article a) the first formal computer science news frame issues network, b) a model for learning the migration paths and patterns about issues, and c) the first formulation of a bias intensity function using the Cusp Catastrophe Model for showing gradual shifts surface response space. Preliminary experiments suggest the integration of cultural dynamic models with Cusp Catastrophe Model are promising for exploring Double Subjectivity for revealing latent relationships found in online news articles. Future work involves testing other hypothesis and unfolding the potential function to explore conditions of convergence.

REFERENCES

[1] P. Cliord and A. Sudbury. A model for spatial conict. Biometrika, 60(3):581{588, 1973.

[2] Eli Ben-Naim, Paul Krapivsky, and Sidney Redner. Bifurcations and patterns in compromise processes. Physica D, 183, 2003.

[3] Claudio Castellano, Santo Fortunato, and Vittorio Loreto. Statistical physics of social dynamics. Reviews of Modern Physics, 81:591–646, 2009.

[4] Guillaume Deffuant, David Neau, Frederic Amblard, and G´erard Weisbuch. Mixing beliefs among interacting agents. Advances in Complex Systems, 3:87–98, 2000.

[5] Rainer Hegselmann and Ulrich Krause. Opinion dynamics and bounded confidence: Models, analysis and simulation. Journal of Artificial Societies and Social Simulation, 5(3), 2002.

[6] B. Latane′. The psychology of social impact. American psychologist, 36(4):343, 1981.

[7] R Axelrod. The dissemination of culture - A model with local convergence and global polarization. Journal of Confict Resolution, 41(2):203{226, APR 1997. ISSN 0022-0027. doi: f10.1177/002200279704100200 1g.

[8] K. C. Chen, M. Chiang, H. V. Poor, "From technological networks to social networks," accepted and to be published in IEEE J. Sel. Areas Commun.., 2013

[9] Y. F. Chou, R. G. Cheng, and H. H. Huang, "Performance evaluation of referral selection algorithms in human centric communication networks,"in Proc. 2012 International Symposium on Wireless Personal Multimedia Communications, pp. 638–642.

[10] Miller McPherson, Lynn Smith-Lovin, and James M. Cook. Birds of a feather: Homophily in social networks. Annual Review of Sociology, 27:415–444, 2001.

[11] Mullainathan, Sendhil, and Andrei Shleifer. 2005. "The Market for News." American Economic Review 95 (1): 1031-1053.

[12] Thomas A. DiPrete and Gregory M. Eirich, Cumulative Advantage as a Mechanism for Inequality: A Review of Theoretical and Empirical Developments, Annual Review of Sociology, Vol. 32: 271-297 (Volume publication date August 2006)

[13] L. R. Rabinerand B. H. Juang,"An introduction to hidden Markovmodels," IEEE ASSP Mag., vol. 3, no. l, pp. 4-16, 1986.

[14] P. Raghavan, S. Amer-Yahia and L. Gravano eds., "Structure in Text: Extraction and Exploitation", In. Proceeding of the 7th international Workshop on the Web and Databases(WebDB), ACM SIGMOD/PODS 2004, ACM Press, Vol 67, 2004.

[15] H. P. Luhn. "The Automatic Creation of Literature Abstracts," IBM Journal of Research and Development, 2, No. 2, 159 (April 1958).

[16] Salton, G., A. Wong, and C. S. Yang. "A vector space model for automatic indexing." In Communications of the ACM 18 (11), 1975: 613-620.

[17] Steinheiser, R., and C. Clifton. "Data Mining on Text." 2012 IEEE 36th Annual Computer Software and Applications Conference IEEE Computer Society, 1998:630.

[18] A. Tan, N. Zhong and L. Zhou, "Text Mining: The state of the art and the challenges." Proceedings of the PAKDD Workshop on Knowledge Discovery and Data Mining, 65—70, Beijing, China 2000.

[19] S. Tan, X. Cheng, B. Wang, H. Xu, M. M. Ghanem, Y. Guo, "Using dragpushing to refine centroid text classifiers", In Proceedings of the 28th annual international ACM SIGIR conference on Research and development in information retrieval (SIGIR '05). ACM, New York, NY, USA, 653-654. 2005.

[20] M. Tsytsarau & T. Palpanas, Survey on mining subjective data on the web, Data Mining and Knowledge Discovery, 24(3), 478-514, 2012.

[21] H. Karanikas, C. Tjortjis, and B. Theodoulidis. "An Approach to Text Mining using Information Extraction", Proc. Workshop Knowledge Management Theory Applications (KMTA 00), 2000.

[22] Q. Hu, D. Yu, Y. Duan, & W. Bao, A novel weighting formula and feature selection for text classification based on rough set theory, in Proceedings of the 2003 International Conference on Natural Language Processing and Knowledge En C. Jiang, F. Coenen, R. Sanderson, M. Zito, "Text Classification Using Graph Mining-Based Feature Extraction", Knowledge-Based Systems, vol. 23, no. 4, pp. 302-308, 2010.gineering (pp. 638-645), 2003. IEEE.

[23] B. Liu, Sentiment analysis and opinion mining, Synthesis lectures on human language technologies, 5(1), 1-167, 2012.

[24] D. A. Scheufele, "Framing as a theory of media effects", The Journal of Communication, 49(1), 103-122, 1999.

[25] S. Dumais, G.W. Furnas, T.K. Landauer, S. Deerwester, Using latent semantic analysis to improve information retrieval. In Proceedings of CHI'88: Conference on Human Factors in Computing, New York: ACM, 281-285, 1988.

[26] C. Fellbaum, "A Semantic Network of English: The Mother of all WordNets," in: Computers and the Humanities 32: 209-220. 1998.

[27] D. Bindela, J. Kleinberg, & S. Orenb, How bad is forming your own opinion?, Games and Economic Behavior, 92, 248–265, 2015.

[28] D. Kempe, J. Kleinberg, and ´ E. Tardos, "Maximizing the spread of influence through a social network," in Proceedings of the ninth ACM SIGKDD international conference on Knowledge discovery and data mining. ACM, pp. 137–146, 2003.

[29] M. Lombard, J. Snyder-Duch, C. C. Bracken, Content analysis in mass communication: Assessment and reporting of intercoder reliability. Human Communication Research, 28, 587–604, 2002.

[30] R. M. Entman, "Framing: Towards clarification of a fractured paradigm", Journal of Communication, 43(4), 51-58, 1993.

[31] G.F. Bishop, R.W. Oldendick, A.J.Tuchfarber,. Effects of presenting one versus two sides of an issue in survey questions. The Public Opinion Quarterly 46, 69–85, 1982.

[32] J.A., Krosnick, Survey research. Annu. Rev. Psychol. 50, 537–567, 1999

[33] K. C. Chen, M. Chiang, H. V. Poor, "From technological networks to social networks," accepted and to be published in IEEE J. Sel. Areas Commun.., 2013

[34] Y. F. Chou, R. G. Cheng, and H. H. Huang, "Performance evaluation of referral selection algorithms in human centric communication networks,"in Proc. 2012 International Symposium on Wireless Personal Multimedia Communications, pp. 638–642.

[35] D. P. Cartwright, "Analysis of qualitative material", In L. Festinger & D. Katz (Eds.), Research Methods in the Behavioral Sciences (pp. 421-470). New York: Dryden, 1953.

[36] R. K.Garrett, B. E. Weeks, R. L. Neo, Driving a Wedge Between Evidence and Beliefs: How Online Ideological News Exposure Promotes Political Misperceptions, Journal of Computer-Mediated Communication, 4 August 2016

[37] Robert Axelrod. The dissemination of culture. Journal of Conflict Resolution, 41(2):203–226, April 1997.

[38] Thom, Rene´. Structural Stability and Morphogenesis, 1972. D. H. Fowler (trans.). Reading, MA: W.A. Benjamin.

[39] Sobkowicz, P. (2012). Discrete model of opinion changes using knowledge and emotions as control variables. PloS one, 7(9), e44489.

[40] Homer-Dixon, T. F. (2006). The upside of down: Catastrophe, creativity, and the renewal of civilization. Washington: Island Press.