2016 IEEE Tenth International Conference on Semantic Computing

“Climate Change” Frames Detection and Categorization Based on Generalized Concepts

Saud Alashri
School of Computing, Informatics, and Decision Systems Engineering
Arizona State University
Tempe, Arizona
Email: salashri@asu.edu

Jiun-Yi Tsai
Hugh Downs School of Human Communication
Arizona State University
Tempe, Arizona
Email: jtsai8@asu.edu

Sultan Alzahrani
School of Computing, Informatics, and Decision Systems Engineering
Arizona State University
Tempe, Arizona
Email: ssalzahr@asu.edu

Steven R. Corman
Hugh Downs School of Human Communication
Arizona State University
Tempe, Arizona
Email: steve.corman@asu.edu

Hasan Davulcu
School of Computing, Informatics, and Decision Systems Engineering
Arizona State University
Tempe, Arizona
Email: hdavulcu@asu.edu

***Abstract*— The subliminal impact of framing of social, political and environmental issues such as climate change has been studied for long time in political science and communications research. Media framing offers “interpretative package” for average citizens on how to make sense of climate change and its consequences to their livelihoods, how to deal with its negative impacts, and which mitigation or adaptation policies to support. A line of related work has used bag of words and word-level features to detect frames automatically in text. Such works face limitations since standard keyword based features may not generalize well to accommodate surface variations in text when different keywords are used for similar concepts. In this paper, we develop a new type of textual features that generalize (subject,verb,object) triplets extracted from text, by clustering them into high-level concepts. We utilize these concepts as features to detect frames in text. Our corpus comprises more than 45,000 climate change related sentences. Expert coders annotated those sentences as frame/non-frame and framed sentences were mapped into one of four general frame categories: solution, problem threat, cause, and motivation. Compared to unigram and bigram based models, classification using our generalized concepts yielded better discriminating features and a higher accuracy classifier with a 12% boost (i.e. from 74% to 83% in f-measure) for frame/no frame detection.**

***Keywords*—Text mining, Frames Detection, Concepts, Big Data, Climate Change, Natural Language Processing**

I. Introduction

Climate change has provoked heated debates on the global political and media arenas. Media framing offers “interpretative package” for average citizens on how to make sense of climate change and its consequences to their livelihoods, how to deal with its negative impacts, and which mitigation or adaptation policies to support [1], [2], [3]. News frames encourage salient interpretation of debated issues through the usage of rhetorical devices (e.g. words, repetitive phrases, and metaphors). Increasingly, governments and international communities are concerned about the security implications of climate change as empirical research has documented that climate change is linked to increased risk of violent conflict [4]. For example, in May 2015, U.S. President Barack Obama suggested that extreme weather is a threat to national security and elevates the risk of global instability and conflict. Some popular press adopted security threat frame to gain public attention. Therefore, systematic detection of news frames related to climate change offers better understanding of stakeholders and their competing perspectives.

Politicians have used framing on hotly debated issues to shift public opinion, gain support and pursue their agenda. A **frame** is the bundling of a component of oratory to urge certain perceptions and to dishearten others [5]. Framing is accomplished when a choice of words, expressions, subjects and other logical gadgets support one understanding of an arrangement of realities, and debilitate other interpretations. One of those framed issues is climate change. Internet created a public space for politicians and stakeholders to frame climate change and related issues to push for their agenda. Online tools such as blogsphere, microblogging and social media streams have increased the availability of data on climate change related debate and made it feasible for researchers to analyze them.

Framing research requires qualitative analysis of a number of texts by subject matter experts to identify and code a set of frames. This is a time consuming process that does not scale well. In order to address the scalability problem, machine learning techniques can be utilized to detect and classify frames. In this paper we propose a system for automatic detection of frames in sentences in a climate change related corpus, and map them to one of four expert identified frame

978-1-5090-0662-5/16 $31.00 © 2016 IEEE
DOI 10.1109/ICSC.2016.14

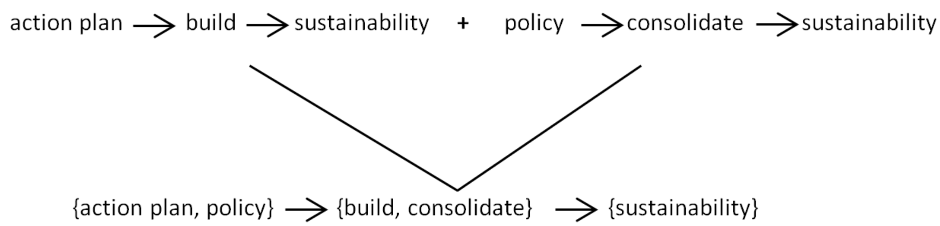

Fig. 1. Example of merging two related concepts

categories: solution, problem threat, cause, and motivation. Our problem here can be described as a multi-level multi-class classification problem where we first classify each sentence as Frame or Non-Frame. Then, the Frame sentences are further classified into one of four predefined frame categories. In particular, we show that generalized concepts and relations [6] as features outperform classical textual features (e.g. uni-grams and bi-grams) while detecting and categorizing *Frame/Non-Frame* sentences. We experimented with SVM [7], Random Forests [8] and sparse logistic regression [9] classifiers, and identified sparse logistic regression as the best performing classifier for these tasks.

Generalized concepts approach extracts high-level information from text as relationships and concepts forming a semantic network. It first uses shallow semantic parser to generate POS tags to obtain semantic triplets (subject,verb,object) from text. Next, it utilizes a bottom-up agglomerative clustering approach to merge and generalize those triplets into concepts. In NLP, shallow parsing is the task of extracting the *subjects*, predicates or *verb phrases*, and *objects*. Figure 1 shows how two related triplets could be merged into a higher level generalized concept. In this figure, two extracted triplets: $\langle$*action plan*$\rightarrow$*build*$\rightarrow$*sustainability*$\rangle$ and $\langle$*policy* $\rightarrow$*consolidate*$\rightarrow$*sustainability*$\rangle$ are merged to form a high level generalized concept and relationship as: $\langle\{$*action plan, policy*$\}\rightarrow\{$*build, consolidate*$\}\rightarrow\{$*sustainability*$\}\rangle$ by discovering contextual synonyms such as $\{$*action plan, policy*$\}$ and $\{$*build, consolidate*$\}$. Here the definition of contextual synonyms is not based on the one in the traditional dictionary. Rather, they correspond to phrases that may occur in similar semantic roles and associated with similar contexts. In Figure 1 the two triplets share the same object $\{$*sustainability*$\}$ and semantically similar verbs; hence, we can merge their subjects $\{$action plan,policy$\}$ as contextual synonyms.

II. Related Work

A. Media Framing

Mainstream media serve as the main arena where international governments, social and political actors, scientists, social movement organizations interact and make competing claims about climate change issues [10]. Communication surrounding climate change can inhibit or support science and policy interactions, propagate consensus or disagreements [11], and ultimately facilitate social change [12], [13], depending on how messages about climate change have been framed [12].

Media representation of climate change plays a vital role in shaping ongoing policy discourse, public perception and attitudes. [14] suggests that prominent political actors frame climate risk for their own purposes, and align frames with their interests and perspectives through media feedback processes of representing climate change risk. Studies have shown that the lay people learn about climate change mainly through consuming mainstream media news [15]. Consequently, [2] argued news media framing can catalyze public engagement and help trigger collective concern of climate change. Put differently, media framing is a powerful tool to highlight different aspects of the policy options, and promote specific interpretations or evaluations that influence decision making [16].

Existing typologies of climate change framing, focusing on dichotomous categories, are limited by their inability to link framing processes with movement interaction. We argue that, in order to understand how the media reflect different organizations interests in addressing climate change as a social problem, it is necessary to supplement the social movement focus on resource mobilization to framing processes of collective action problems. To do that, this study develops a nuanced typology for studying climate change framing and its adequacy for supporting a social movements that would be necessary to overcome the collective action problem. Our typology provides a holistic map to evaluate how climate change media framing can enable appropriate social and policy actions that ultimately can mitigate risks of social unrest. We apply this framework to examine framing of climate change in media and social media texts collected from the Niger Basin region over seven months from August 2014 to February 2015, using a novel coding technique to assess diagnostic, prognostic, and motivational framing described by [17] as the keys to effective social movements.

B. Framing Research in Computer Science

[18] examined twitter stream on extracted frames and pointed out a strong ties between frames collected from news with the public opinions expressed in tweeter feeds. [19] went further to distill agenda from news and link them to action. Content analysis of frames in news is performed either by (1) manual frame coding, that is done by trained coders, which is costly as well as not scalable, or by (2) frame identification by using machine learning techniques that overcome human

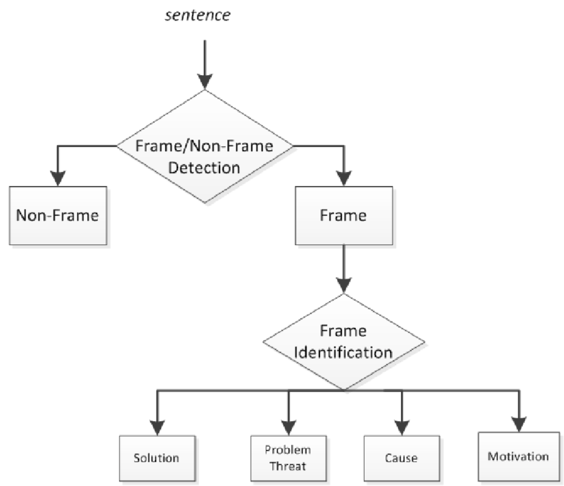

Fig. 2. Multi-level multi-class classification

limitations by automatically detecting frames after training a classifier [20]. Many studies have addressed media framing as a document classification problem by building a learning model to classify documents or paragraphs by utilizing different features. Aside from document level, [21], [22] examined the classification task at the sentence level and even at the phrase level. Previous work on sentence level classification has focused on experimenting with different classifiers and different features. [23] examined: bag of words, n-grams, and topic models to classify news articles and map them to a set of frames. Others, [24] employed POS-tags [25] and named entities [26] as features to detect and classify frames. Ceran et al [27] experimented with {subject,verb,object} based features and benchmarked " paragraph level" classifier for story detection against standard keyword based features, which showed significant improvement in classification accuracy. More advanced conceptual features engineering was developed in [6] as they showed how generalized concepts performed better in detecting stories in paragraphs. We utilize their generalized concepts as features to detect and categorize frames. Our paper works on sentence level classification compared to their paragraph level. Also, our task is a multi-level multi-class classification task where we first examine if a sentence contains a frame, and then we identify which one of four frame categories it belongs to. Moreover, we developed triple-extraction techniques where we can extract more features and incorporate a larger percentage of sentences into the classification model (i.e. 80% of sentences compared to 40%).

III. Problem Definition

Given a set of documents $\{D_1,...,D_M\}$ where each document contains one or more paragraphs. First, we split documents into sentences $\{S_1,...,S_N\}$. Next, using sentences as data points, we aim to resolve whether a sentence S_i contains a frame or not. And, if the sentence contains a frame, then we aim to identify its category, as one of: {*Solution, Problem Threat, Cause, Motivation*}. Figure 2 shows our multi-level multi-class classification problem for a given sentence.

IV. Methodology

A. Overall System Model

The overall system consists of documents collected from nearly 100 RSS feeds that are related to climate change in the Niger Delta region. We also perform sentence splitting of documents, identification of key frames and their categories by the coders, feature extraction (uni-grams, bi-grams, and generalized concepts), identification of discriminative features, and a predictive model to detect and identify the frame categories for sentences containing frame references.

B. Climate Change Corpus

Our climate change corpus is comprised of nearly $45,054$ sentences extracted from news and social media websites, that are related to climate change topics in Niger Basin region over a seven months period from August 2014 to February 2015. There are $16,050$ sentences coded as frame sentences and $29,004$ as non-frame sentences by domain experts. Frame sentences are further categorized into one of four categories: Solution, Problem Threat, Cause, and Motivation.

C. N-gram Features

We experimented with both uni-gram and bi-gram features. We run a simple term frequency - inverse document frequency (TF-IDF) [28] based technique on the entire corpus to generate a large ranked list of, stopword eliminated, uni-grams and bi-grams, and we experimented with them separately as features in our classification models.

D. Generalized Concepts Features

In [6], they extracted concepts from paragraphs where only 40% of the paragraphs generated concepts. In this paper, since we are working on sentence level, we improved the concept extraction approach, by extracting more triplets by utilizing a larger number of triplet extractors and pre-processing their output to include about 80% of the sentences in our experimental evaluations.

1) Triplets Extraction: In order to extract *Subject,Verb,Object* triplets, first we run a pronoun resolver [29], [30], [31], [32]. Since triplets extraction is an ongoing research topic in NLP, we proceeded to use four state-of-the-art triplets extraction tools: ClearNLP [33], Reverb [34], Everest [35], AlchemyAPI [36] as complementary systems. Additionally, any triplet slots with phrases were segmented into keywords, stemmed, stop-word removed and their cartesian product were produced as additional triplets.

2) Concepts Generation: Triplets extraction algorithms typically produce noisy and sparse triplets. Therefore, we apply a hierarchical bottom-up clustering algorithm that generalizes triplets into more meaningful relationships. To do so, we employ both syntactic and semantic criteria that are based on the corpus to generalize triplets into high level concepts without *drift*. In *syntactic criteria*, a pair of subjects-verbs-objects are merged only if they share common context related to their different arguments (i.e. a pair of different subjects are merged only if they co-occur with an identical verb-object context).

Additionally, we capture contextual synonyms for subjects, verbs and objects by defining a *semantic criterion* which is based on our corpus as well as WordNet [37]. Corpus-based contextual synonyms for subjects, verbs and objects is based on their common verb-object, subject-object and subject-verb contexts respectively. Also, we capture contextual synonyms that are not derivable from our corpus by applying WordNet synonyms and hyponyms on the memebers of the concepts to further expand and generalize them.

In order for the information to propagate between clusters of subjects/objects and clusters of relations, we apply a hierarchical bottom-up clustering algorithm [38]. High level concepts and relations are merged to form clusters. Each cluster is represented by graph of nodes and edges where nodes represent concepts and edges represent relations between concepts. The details of the above criteria and the generalization algorithm are available in [6].

E. Frame Classification

To classify each sentence as Frame/Non-Frame and identify its relevant frame category we utilize sparse learning framework [9], with the underlined motivation to select a subset of discriminating concepts that can (1) identify sentences containing frame references and (b) classify a sentence into a frame category. The following steps describe our algorithm:

1) Generate features from the entire corpus
2) Filter the features $\times$ sentences matrix to include only resultant generalized concepts/features
3) Formulate the problem in a general sparse learning framework [9]. In particular, the logistical regression formulation presented below fits this application, since it is a dichotomous frame classification problem (e.g. each sentence classified as Frame/Non-Frame), and multi-class classification problem (e.g. each Frame sentence is further classified as one of four frame $\{$Solution, Problem Threat, Cause, and Motivation$\}$):

$$\min_x \sum_{i=1}^{m} w_i \log(1 + \exp(-y_i(x^t a_i + c))) + \lambda|x| \quad (1)$$

In formula (1), a_i is the vector representation of the i^{th} sentence, w_i is the weight assigned to the i^{th} sentence ($w_i = 1/m$ by default), and $A = [a_1, a_2, \ldots, a_m]$ is the features $\times$ sentences matrix, y_i is the label of each sentence, and the x_j , the j^{th} element of x, is the unknown weight for each feature, ($\lambda > 0$) is a regularization parameter that controls the sparsity of the solution, $|x|_1 = \sum |x_i|$ is 1-norm of the x vector. We used the SLEP [39] sparse learning package that utilizes gradient descent approach to solve the above convex and non-smooth optimization problem. The features with non-zero values on the sparse x vector yield the discriminant factors for classifying a sentence.

V. Experimental Evaluation

A. Sentence Annotation

Our experts developed four categories of climate change related frames as follows:

- **Solution framing (prognostic)**: Covering the prognostic function of defining what should be done about problems, solution framing refers to actions taken to prevent further impact of climate change effects or further impact of the causes of climate change such as greenhouse gas emissions. Solutions can also emphasize ongoing measures to deal with existing effects of climate change. Six frames capture an array of mitigation and adaptation efforts conservation, education, investment, infrastructure and development, creation or implementation of policy and programs, and goal.
- **Problem Threat framing (diagnostic)**: This diagnostic framing class stresses on how climate change or outcomes of climate change impact various actors, industries, human health, and the environment, Eight codes capture negative consequences and threats brought by climate change, including environmental systems and ecosystem, public health, economic development, food security, water scarcity, national security, social unrest, and general or multiple impacts. Both cause framing and problem/threat framing comprise the diagnostic function in defining social problems.
- **Cause framing**: This group of diagnostic frames focus on attributing the blame for causing climate change to either human activity, natural variation or other reasons. Six subcategories captured different explanations for causal attribution of climate change: (a) human activity, (b) natural variation, (c) scientific uncertainty, (d) policy causes, (e) insufficient actions, and (f) human disruption to mitigate climate change impact.
- **Motivation framing (motivational)**: Motivational framing refers to statements that explicit call for definitive course(s) of action and explain why the audience should make an effort to enact solutions [17]. In other words, motivational frames elaborate on the rationale for action that goes beyond diagnosis and prognosis, and include vocabularies of severity, urgency, efficacy, and propriety [40]. We added a general category to analyze statements that call for actions without providing readers with above-mentioned reasons.

We assigned sentence annotation to three different expert coders where we break ties by using the majority vote.

B. Quantitaive Evaluation

Once sentences are labeled as Frame/Non-Frame and categorized with their corresponding frame category, we utilize uni-gram keywords, bi-gram terms, and generalized concepts separately as features and the sparse logistical regression

TABLE I
FRAME/NON-FRAME CLASSIFICATION

Method	Class Label	Precision	Recall	F-measure
Concepts	Frame	0.80	0.88	0.84
	Non Frame	0.87	0.77	0.82
	Average	**0.83**	**0.83**	**0.83**
Bi-grams	Frame	0.75	0.42	0.54
	Non Frame	0.74	0.92	0.82
	Average	**0.74**	**0.67**	**0.68**
Uni-grams	Frame	0. 75	0.48	0.59
	Non Frame	0.76	0.91	0.89
	Average	**0.75**	**0.70**	**0.74**

classifier SLEP [39] to identify weighted discriminative features and classify sentences. We experimented with three different classifiers (SVM [7], Random Forests [8]) and found that SLEP outperformed both these other classifiers. Using different types of features generated from the entire corpus, we perform ten-fold cross-validation for measuring the classifier's predictive accuracy to detect Frame/Non-Frame sentences. Next, using features generated from frame sentences only, we train a multi-class model to classify sentences into their corresponding frame category. We report precision, recall, and F-measure as quantitative evaluation metrics. Qualitative analysis of the identified discriminating concepts is also presented in the next section.

Table 1 presents the accuracies for detecting Frame/Non-Frame sentences using different features. Using generalized concepts approach as features, the resultant average accuracy (F-measure of 83%) outperforms both accuracies with uni-grams (74%) and bi-grams (68%) features by 12% and 22% respectively.

Table 2 shows the accuracies for identifying the corresponding frame category. Using generalized concepts, these accuracies vary between 73% and 83% (F-measure) for different categories. In this table, utilizing generalized concepts yields slightly better performance compared to both uni-grams and bi-grams with an overall average accuracy (F-measure) of 79%.

C. *Qualitative Analysis of Resultant Concepts*

Table III shows top five discriminative concepts for each frame category. Our team of experts explored the highly significant generalized concepts germane to four-class framing in media discourse surrounding climate change across West African RSS feeds and provided qualitative evaluations as follows:

1) Cause Framing: Causal responsibility of climate change and its effects was often attributed to anthropogenic activities, particularly man-made greenhouse gas emissions, human-induced pollution, and fossil fuel use. Carbon dioxide and greenhouse gas emission emerged as highly significant concepts, as indicated by high weigh value. Media texts often

TABLE II
FRAME CLASSIFICATION INTO FOUR CATEGORIES

Method	Frame Category	Precision	Recall	F-measure
Concepts	Solution	0.75	0.93	0.83
	Problem Threat	0.77	0.84	0.79
	Cause	0.85	0.77	0.80
	Motivation	0.89	0.62	0.73
	Average	**0.82**	**0.79**	**0.79**
Bi-grams	Solution	0.87	0.77	0.81
	Problem Threat	0.84	0.77	0.80
	Cause	0.86	0.73	0.76
	Motivation	0.90	0.58	0.71
	Average	**0.87**	**0.71**	**0.77**
Uni-grams	Solution	0.78	0.87	0.82
	Problem Threat	0.81	0.81	0.81
	Cause	0.83	0.62	0.82
	Motivation	0.85	0.57	0.64
	Average	**0.82**	**0.72**	**0.77**

associated global warming with carbon dioxide emissions using the following triplets to construct a cohesive story:

- Scientific research indicate that atmospheric carbon dioxide increase at a large level.
- Cars and trucks were major sources of air pollution and carbon dioxide emissions, which directly increased local temperature.

2) Problem Threat Framing: Next, we turned our attention to identify the dominant concepts representing the problem and threat framing of climate change. Media texts tended to highlight devastating environmental impacts caused by climate change, such as floods, prolonged drought, loss of landmass and soil, desertification, sea-level rise, storm surge, heat waves, and more. Flooding, in particular, is a severe concern as nine out of sixteen triplets of high weigh values explicitly mentioned the negative impacts of heavy rainfall or torrential rain. Consequently, economic condition and food insecurity were influenced, infrastructure was damaged, and health diseases were exacerbated with the increased intensity and frequency of floods.

3) Solution Framing: The most representative discourse of solution framing is discussed next in Section D.

4) Motivation Framing: When discussing motivation for why policy actors and citizens should act upon, the most salient concepts emphasized that international communities (e.g. U.S., EU, and China) should negotiate a legal agreement to reduce greenhouse gas emissions at the end of 2015. There is little attention to stating specific reasons for offering localized adaptation strategies that people can undertake. Although the awareness of climate change impacts among African government officials was generally high, the prevailing generalized concept of calling for international actions on mitigation from mainstream media discourse reflected a lack

TABLE III
TOP FIVE GENERATED CONCEPTS FOR EACH FRAME CATEGORY

Cause	Problem Threat	Solution	Motivation
{Greenhouse,Emissions,Gases} ↓ {Cause,Attribute to} ↓ {Global warming}	{Flood} ↓ {Associate,Create} ↓ {Poverty,Disease}	{Action plan,Policy} ↓ {Build,Consolidate} ↓ {Sustainability,Resilience future}	{International,Community} ↓ {Urge,Warn} ↓ {Threat}
{Industry,Anthropogenic} ↓ {Raise} ↓ {Earth temperature,CO2,CO5}	{Heavy rainfall, Torrential rain} ↓ {Create,Bring,Increase} ↓ {Flooding,Disaster,Landslide}	{Development, Sustainability,National program} ↓ {Enhance} ↓ {Community}	{Agreement,Leaders,World} ↓ {Help} ↓ {Future,Hope}
{Fossil fuel} ↓ {Impact,Harm} ↓ {Planet,Environment,Weather}	{Drought} ↓ {Cause,Impact,Reduce} ↓ {Food-shortage,Food-production,Crop}	{Brown} ↓ {Sign} ↓ {Local legislation, CA groundwater,Management framework}	{USA,EU,China} ↓ {Recognize,Reduce} ↓ {Emissions}
{Coal combustion,Diesel,Man-Made} ↓ {Create} ↓ {Extreme weather,Temperature-up}	{Sea-level rise} ↓ {Result in,Cause} ↓ {Tsunami,Damage,Flood}	{Sustainability,Energy} ↓ {Can help,Improve} ↓ {Food security,Households}	{Africa} ↓ {Need,Implement} ↓ {Policy,Awareness,Partnership}
{Truck,Car} ↓ {Rise} ↓ {Carbon pollution,Pollute}	{Extreme Weather, Hailstorm} ↓ {Cause,Affect} ↓ {Mudslide,Floods,Farming}	{Smart agriculture,Africa countries} ↓ {Meet,Breathe} ↓ {Life}	{Nigerian} ↓ {Apply,Take} ↓ {Measures,Renewable Energy,Policy}

of effective national and local polices.

D. Visualizing Concepts

To visualize the generalized concept and relation clusters, we utilize a semantic network [41] of nodes (V) and edges (E) to describe the semantic space of the underlying texts. Circle nodes represent *subjects/objects* and square nodes represent *verbs*. Edges represent relations between concepts. In such a network, distinct combinations of actors (subjects) perform or recommend various sets of actions (verbs) on distinct combinations of targets (objects). The sample semantic network in Figure 3 (next page) illustrates how *sustainability* emerges as a concept that is central to addressing climate change impacts. The semantic network represents the contextual relationships between generalized triplets relating to strategies for sustainable adaptation. In the media discourse, sustainable adaptation is predominantly framed as an effective solution to reduce impacts of climate change and contribute to social, economic, and environmental development. As shown in Figure 3, developing sustainable national programs (or actions) can enhance local community resilience. According to the IPCC (Intergovernmental Panel on Climate Change) report, majority of rural communities rely on rain-fed agriculture to sustain their livelihoods in West Africa, the region worst affected by climate change. With changing rainfall patterns, prolonged droughts and flooding, sustainable system of developing agriculture-smart technologies can help improve food security at the household level. Interestingly, the African media discussed that California Governor Jerry Brown has signed the most significant framework for regulating underground water resources to achieve sustainable development in September, 2014.

VI. CONCLUSION AND FUTURE WORK

Climate change framing has pervasive influence, and this paper presents a new computational approach based on generalized concepts to identify popular media frames and map them to different categories: solution, problem threat, cause, and motivation. A line of related work has used bag of words and word-level features to detect frames automatically in text. Such

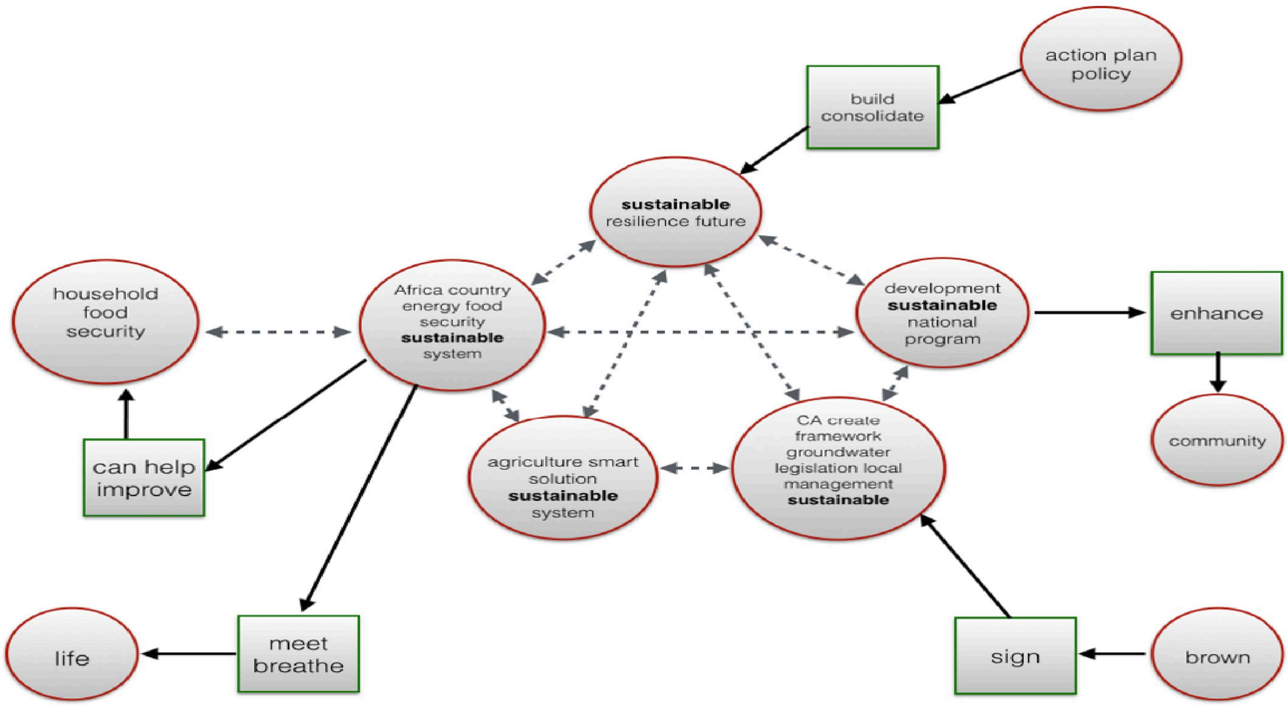

Fig. 3. A sample semantic network of frame concepts

work face limitations since standard keyword based features may not generalize well to accommodate surface variations in text when different keywords are used for similar concepts. In this paper, we developed a new type of textual features that generalize (subject,verb,object) triplets extracted from text, by clustering them into high-level concepts. Compared to unigram and bigram based models, frame classification using our generalized concepts yielded better discriminating features with a 12% boost in accuracy (i.e. from 74% to 83% in f-measure) for frame/no frame detection. In our future work, we plan to utilize discriminating generalized concepts indicating actor-action-target sequences to infer causal chains of events, frames, and actions that might lead to better indicators of climate-change related social unrest.

Acknowledgment

Some of the material presented here was sponsored by Department of Defense and it is approved for public release, case number:15-467.

References

[1] D. Chong and J. N. Druckman, "A theory of framing and opinion formation in competitive elite environments," *Journal of Communication*, vol. 57, no. 1, pp. 99–118, 2007. [Online]. Available: http://dx.doi.org/10.1111/j.1460-2466.2006.00331.x

[2] M. C. Nisbet, "Communicating climate change: Why frames matter for public engagement," *Environment: Science and Policy for Sustainable Development*, vol. 51, no. 2, pp. 12–23, 2009. [Online]. Available: http://dx.doi.org/10.3200/ENVT.51.2.12-23

[3] A. Shehata and D. N. Hopmann, "Framing climate change," *Journalism Studies*, vol. 13, no. 2, pp. 175–192, 2012. [Online]. Available: http://dx.doi.org/10.1080/1461670X.2011.646396

[4] J. Barnett and W. N. Adger, "Climate change, human security and violent conflict," *Political Geography*, vol. 26, no. 6, pp. 639 – 655, 2007, climate Change and Conflict. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S096262980700039X

[5] S. Alashri, S. Alzahrani, L. Bustikova, D. Siroky, and H. Davulcu, "What animates political debates? analyzing ideological perspectives in online debates between opposing parties," in *Proceedings of the ASE/IEEE International Conference on Social Computing (SocialCom-15)*, 2015.

[6] B. Ceran, N. Kedia, S. Corman, and H. Davulcu, "Story detection using generalized concepts and relations," in *Proceedings of International Symposium on Foundation of Open Source Intelligence and Security Informatics (FOSINT-SI), in conj. with IEEE ASONAM*, 2015.

[7] C. Cortes and V. Vapnik, "Support-vector networks," *Machine learning*, vol. 20, no. 3, pp. 273–297, 1995.

[8] L. Breiman, "Random forests," *Mach. Learn.*, vol. 45, no. 1, pp. 5–32, Oct. 2001. [Online]. Available: http://dx.doi.org/10.1023/A:1010933404324

[9] J. Liu, J. Chen, and J. Ye, "Large-scale sparse logistic regression," in *Proceedings of the 15th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, ser. KDD '09. NY, USA: ACM, 2009, pp. 547–556. [Online]. Available: http://doi.acm.org/10.1145/1557019.1557082

[10] S. Hilgartner and C. L. Bosk, "The rise and fall of social problems: A public arenas model," *American journal of Sociology*, pp. 53–78, 1988.

[11] M. Hulme, *Why we disagree about climate change: Understanding controversy, inaction and opportunity*. Cambridge University Press, 2009.

[12] M. T. Boykoff, *Who speaks for the climate?: Making sense of media reporting on climate change*. Cambridge University Press, 2011.

[13] S. C. Moser and L. Dilling, *Creating a climate for change*. Cambridge University Press, 2006.

[14] A. Carvalho, "Ideological cultures and media discourses on scientific knowledge: re-reading news on climate change," *Public understanding of science*, vol. 16, no. 2, pp. 223–243, 2007.

[15] R. J. Brulle, J. Carmichael, and J. C. Jenkins, "Shifting public opinion on climate change: an empirical assessment of factors influencing concern over climate change in the us, 2002–2010," *Climatic change*, vol. 114, no. 2, pp. 169–188, 2012.

[16] R. M. Entman, "Framing: Towards clarification of a fractured paradigm," *McQuail's reader in mass communication theory*, pp. 390–397, 1993.
[17] R. D. Benford and D. A. Snow, "Framing processes and social movements: An overview and assessment," *Annual review of sociology*, pp. 611–639, 2000.
[18] S. M. Jang and P. S. Hart, "Polarized frames on climate change and global warming across countries and states: evidence from twitter big data," *Global Environmental Change*, vol. 32, pp. 11–17, 2015.
[19] K. Stalpouskaya and C. Baden, "To do or not to do: the role of agendas for action in analyzing news coverage of violent conflict," *ACL-IJCNLP 2015*, p. 21, 2015.
[20] B. Burscher, D. Odijk, R. Vliegenthart, M. de Rijke, and C. H. de Vreese, "Teaching the computer to code frames in news: Comparing two supervised machine learning approaches to frame analysis," *Communication Methods and Measures*, vol. 8, no. 3, pp. 190–206, 2014.
[21] Y. Kim, "Convolutional neural networks for sentence classification," *arXiv preprint arXiv:1408.5882*, 2014.
[22] T. Wilson, J. Wiebe, and P. Hoffmann, "Recognizing contextual polarity in phrase-level sentiment analysis," in *Proceedings of the conference on human language technology and empirical methods in natural language processing*. Association for Computational Linguistics, 2005, pp. 347–354.
[23] D. Odijk, B. Burscher, R. Vliegenthart, and M. De Rijke, "Automatic thematic content analysis: Finding frames in news," in *Social Informatics*. Springer, 2013, pp. 333–345.
[24] E. Baumer, E. Elovic, Y. Qin, F. Polletta, and G. Gay, "Testing and comparing computational approaches for identifying the language of framing in political news." in *HLT-NAACL*, 2015.
[25] K. Toutanova, D. Klein, C. D. Manning, and Y. Singer, "Feature-rich part-of-speech tagging with a cyclic dependency network," in *Proceedings of the 2003 Conference of the North American Chapter of the Association for Computational Linguistics on Human Language Technology-Volume 1*. Association for Computational Linguistics, 2003, pp. 173–180.
[26] J. Finkel, T. Grenager, and C. Manning, "Incorporating non-local information into information extraction systems by gibbs sampling," in *Proceedings of the 43rd Annual Meeting on Association for Computational Linguistics*. Association for Computational Linguistics, 2005, pp. 363–370.
[27] B. Ceran, R. Karad, A. Mandvekar, S. R. Corman, and H. Davulcu, "A semantic triplet based story classifier," in *Advances in Social Networks Analysis and Mining (ASONAM), 2012 IEEE/ACM International Conference on*. IEEE, 2012, pp. 573–580.
[28] J. A. Hartigan and M. A. Wong, "Algorithm AS 136: A k-means clustering algorithm," *Applied Statistics*, vol. 28, no. 1, pp. 100–108, 1979. [Online]. Available: http://dx.doi.org/10.2307/2346830
[29] K. Raghunathan, H. Lee, S. Rangarajan, N. Chambers, M. Surdeanu, D. Jurafsky, and C. Manning, "A multi-pass sieve for coreference resolution," in *Proceedings of the 2010 Conference on Empirical Methods in Natural Language Processing*. Association for Computational Linguistics, 2010, pp. 492–501.
[30] H. Lee, Y. Peirsman, A. Chang, N. Chambers, M. Surdeanu, and D. Jurafsky, "Stanfords multi-pass sieve coreference resolution system at the conll-2011 shared task," *CoNLL 2011*, p. 28, 2011.
[31] H. Lee, A. Chang, Y. Peirsman, N. Chambers, M. Surdeanu, and D. Jurafsky, "Deterministic coreference resolution based on entity-centric, precision-ranked rules," *Comput. Linguist.*, vol. 39, no. 4, pp. 885–916, Dec. 2013.
[32] M. Recasens, M. C. de Marneffe, and C. Potts, "The life and death of discourse entities: Identifying singleton mentions," in *Proceedings of the 2013 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, 2013, pp. 627–633.
[33] J. D. Choi, "Optimization of natural language processing components for robustness and scalability," Ph.D. dissertation, University of Colorado at Boulder, 2012.
[34] A. Fader, S. Soderland, and O. Etzioni, "Identifying relations for open information extraction," in *Proceedings of the Conference of Empirical Methods in Natural Language Processing (EMNLP '11)*, Edinburgh, Scotland, UK, July 27-31 2011.
[35] (2013) Everest triplet extraction. Next Century Corporation. [Online]. Available: https://github.com/NextCenturyCorporation/EVEREST-TripletExtraction
[36] (2015) Alchemyapi language features. AlchemyAPI, Inc. [Online]. Available: http://www.alchemyapi.com/products/alchemylanguage
[37] (2010) About wordnet. Princeton University. [Online]. Available: http://wordnet.princeton.edu
[38] S. Kok and P. Domingos, "Extracting semantic networks from text via relational clustering," in *Proceedings of the 2008 European Conference on Machine Learning and Knowledge Discovery in Databases - Part I*, 2008, pp. 624–639.
[39] J. Liu, S. Ji, and J. Ye, *SLEP: Sparse Learning with Efficient Projections*, Arizona State University, 2009. [Online]. Available: http://www.public.asu.edu/~jye02/Software/SLEP
[40] R. D. Benford, "" you could be the hundredth monkey": Collective action frames and vocabularies of motive within the nuclear disarmament movement," *Sociological Quarterly*, pp. 195–216, 1993.
[41] M. R. Quillian, "Semantic memory," in *Semantic Information Processing*, 1968, pp. 227–270.