

# Identifying Ideological Perspectives of Web Videos using Patterns Emerging from Folksonomies

Wei-Hao Lin and Alexander Hauptmann\*

Language Technologies Institute  
School of Computer Science  
Carnegie Mellon University  
Pittsburgh, PA 15213 USA  
{whlin,alex}@cs.cmu.edu

**Abstract.** We are developing a classifier that can automatically identify a web video’s ideological perspective on a political or social issue (e.g., pro-life or pro-choice on the abortion issue). The problem has received little attention, possibly due to inherent difficulties in content-based approaches. We propose to develop such a classifier based on the pattern of tags emerging from folksonomies. The experimental results are positive and encouraging.

## 1 Introduction

Video sharing websites such as YouTube, Metacafe, and Imeem have been extremely popular among Internet users. More than three quarters of Internet users in the United States have watched video online. In a single month in 2008, 78.5 million Internet users watch 3.25 billion videos on YouTube. On average, YouTube viewers spend more than one hundred minutes a month watching videos on YouTube [1].

Video sharing websites have also become an important platform for expressing and communicating different views on various social and political issues. In 2008, CNN and YouTube held United States presidential debates in which presidential candidates answered questions that were asked and uploaded by YouTube users. In March 2008 YouTube launched YouChoose’08<sup>1</sup> in which each presidential candidate has their own channel. The accumulative viewership for one presidential candidate as of June 2008 has exceeded 50 millions [2]. In addition to politics, many users have authored and uploaded videos expressing their views on social issues. For example, Figure 1 is an example of a “pro-life” web video on the abortion issue<sup>2</sup>, while Figure 2 is an example of “pro-choice” web video<sup>3</sup>.

---

\* We would like to thank the anonymous reviewers for their valuable comments and suggestions. This work was supported in part by the National Science Foundation (NSF) under Grants No. IIS-0535056 and CNS-0751185.

<sup>1</sup> <http://youtube.com/youchoose>

<sup>2</sup> <http://www.youtube.com/watch?v=TddCILTWnr8>

<sup>3</sup> <http://www.youtube.com/watch?v=oWeX0jsv58c>



**Fig. 1.** The key frames of a web video expressing a “pro-life” view on the abortion issue, which is tagged with *prayer*, *pro-life*, and *God*.



**Fig. 2.** The key frames of a web video expressing a “pro-choice” view on the abortion issue, which is tagged with *pro*, *choice*, *feminism*, *abortion*, *women*, *rights*, *truth*, and *Bush*.

We are developing a computer system that can automatically identify highly biased broadcast television news and web videos. Such a system may increase an audience’s awareness of individual news broadcasters’ or video authors’ biases, and can encourage viewers to seek videos expressing contrasting viewpoints. Classifiers that can automatically identify a web video’s ideological perspective will enable video sharing sites to organize videos on various social and political views according to their ideological perspectives, and allow users to subscribe to videos based on their personal views. Automatic perspective classifiers will also enable content control or web filtering software to filter out videos expressing extreme political, social or religious views that may not be suitable for children.

Although researchers have made great advances in automatically detecting “visual concepts” (e.g., car, outdoor, and people walking) [3], developing classifiers that can automatically identify whether a video is about *Catholic* or *abortion* is still a very long-term research goal. The difficulties inherent in content-based approaches may explain why the problem of automatically identifying a video’s ideological perspective on an issue has received little attention.

- In this paper we propose to identify a web video’s ideological perspective on political and social issues using associated tags. Videos on video sharing sites such as YouTube allow users to attach tags to categorize and organize videos. The practice of collaboratively organizing content by tags is called folksonomy, or collaborative tagging. In Section 3.3 we show that a unique pattern of tags emerges from videos expressing opinions on political and social issues.
- In Section 2 we apply a statistical model to capture the pattern of tags from a collection of web videos and associated tags. The statistical model

simultaneously captures two factors that account for the frequency of a tag associated with a web video: what is the subject matter of a web video? and what ideological perspective does the video’s author take on an issue?

- We evaluate the idea of using associated tags to classify a web video’s ideological perspective on an issue in Section 3. The experimental results in Section 3.2 are very encouraging, suggesting that Internet users holding similar ideological beliefs upload, share, and tag web videos similarly.

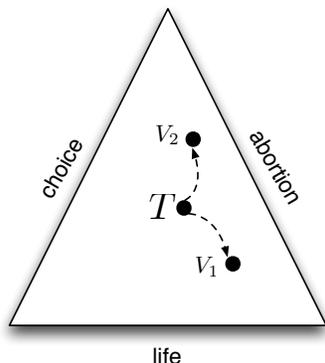
## 2 Joint Topic and Perspective Model

We apply a statistical model to capture how web videos expressing strongly a particular ideological perspective are tagged. The statistical model, called the Joint Topic and Perspective Model [4], is designed to capture an *emphatic* pattern empirically observed in many ideological texts (editorials, debate transcripts) and videos (broadcast news videos). We hypothesize that the tags associated with web videos on various political and social issues also follow the same emphatic pattern.

The emphatic pattern consists of two factors that govern the content of ideological discourse: *topical* and *ideological*. For example, in the videos on the abortion issue, tags such as **abortion** and **pregnancy** are expected to occur frequently no matter what ideological perspective a web video’s author takes on the abortion issue. These tags are called *topical*, capturing what an issue is about. In contrast, the occurrences of tags such as **pro-life** and **pro-choice** vary much depend on a video author’s view on the abortion issue. These tags are emphasized (i.e., tagged more frequently) on one side and de-emphasized (i.e., tagged less frequently) on the other side. These tags are called *ideological*.

The Joint Topic and Perspective Model assigns *topical* and *ideological* weights to each tag. The topical weight of a tag captures how frequently the tag is chosen because of an issue. The *ideological* weight of a tag represents to what degree the tag is emphasized by a video author’s ideology on an issue. The Joint Topic and Perspective Model assumes that the observed frequency of a tag is governed by these two sets of weights combined.

We illustrate the main idea of the Joint Topic and Perspective Model in a three tag world in Figure 3. Any point in the three tag simplex represents the proportion of three tags (e.g., **abortion**, **life**, and **choice**) chosen in web videos about the abortion issue (also known as a multinomial distribution’s parameter).  $T$  represents how likely we would be to see **abortion**, **life**, and **choice** in web videos about the abortion issue. Suppose a group of web video authors holding the “pro-life” perspective choose to produce and tag more **life** and fewer **choice**. The *ideological* weights associated with this “pro-life” group in effect move the proportion from  $T$  to  $V_1$ . When we sample tags from a multinomial distribution of a parameter at  $V_1$ , we would see more **life** and fewer **choice** tags. In contrast, suppose a group of web video authors holding the “pro-choice” perspective choose to make and tag more **choice** and fewer **life**. The *ideological* weights associated with this “pro-choice” group in effect move the proportion



**Fig. 3.** A three tag simplex illustrates the main idea behind the Joint Topic and Perspective Model.  $T$  denotes the proportion of the three tags (i.e., topical weights) that are chosen for a particular issue (e.g., abortion).  $V_1$  denotes the proportion of the three tags after the topical weights are modulated by video authors holding the “pro-life” view;  $V_2$  denotes the proportion of the three tags modulated by video authors holding the contrasting “pro-choice” view.

from  $T$  to  $V_2$ . When we sample tags from a multinomial distribution of a parameter at  $V_2$ , we would see more **life** and fewer **choice** tags. The topical weights determine the position of  $T$  in a simplex, and each ideological perspective moves  $T$  to a biased position according to its ideological weights.

We can fit the Joint Topic and Perspective Model on data to simultaneously uncover topical and ideological weights. These weights succinctly summarize the emphatic patterns of tags associated with web videos about an issue. Moreover, we can apply the weights learned from training videos, and predict the ideological perspective of a new web video based on associated tags.

## 2.1 Model Specification and Predicting Ideological Perspectives

Formally, the Joint Topic and Perspective Model assumes the following generative process for the tags associated with web videos:

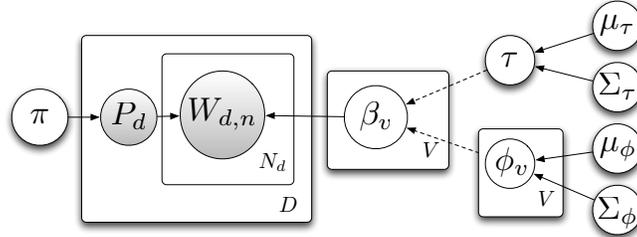
$$\begin{aligned}
 P_d &\sim \text{Bernoulli}(\pi), d = 1, \dots, D \\
 W_{d,n} | P_d = v &\sim \text{Multinomial}(\beta_v), n = 1, \dots, N_d \\
 \beta_v^w &= \frac{\exp(\tau^w \times \phi_v^w)}{\sum_{w'} \exp(\tau^{w'} \times \phi_v^{w'})}, v = 1, \dots, V \\
 \tau &\sim \text{N}(\mu_\tau, \Sigma_\tau) \\
 \phi_v &\sim \text{N}(\mu_\phi, \Sigma_\phi).
 \end{aligned}$$

The ideological perspective  $P_d$  from which the  $d$ -th web video in a collection was produced (i.e., its author or uploader’s ideological perspective) is assumed

to be a Bernoulli variable with a parameter  $\pi$ . In this paper, we focus on bipolar ideological perspectives, that is, those political and social issues with only two perspectives of interest ( $V = 2$ ). There are a total of  $D$  web videos in the collection. The  $n$ -th tag in the  $d$ -th web video  $W_{d,n}$  is dependent on its author's ideological perspective  $P_d$  and assumed to be sampled from the multinomial distribution of a parameter  $\beta$ . There are a total of  $N_d$  tags associated with the  $d$ -th web video.

The tag multinomial's parameter,  $\beta_v^w$ , subscripted by an ideological perspective  $v$  and superscripted by the  $w$ -th tag in the vocabulary, consists of two parts: a topical weight  $\tau^w$  and ideological weights  $\{\phi_v^w\}$ . Every tag is associated with one topical weight  $\tau^w$  and two ideological weights  $\phi_1^w$  and  $\phi_2^w$ .  $\beta$  is an auxiliary variable, and is deterministically determined by (unobserved) topical and ideological weights.

$\tau$  represents the *topical* weights and is assumed to be sampled from a multivariate normal distribution of a mean vector  $\mu_\tau$  and a variance matrix  $\Sigma_\tau$ .  $\phi_v$  represents the *ideological* weights and is assumed to be sampled from a multivariate normal distribution of a mean vector  $\mu_\phi$  and a variance matrix  $\Sigma_\phi$ . Every tag is associated with one topical weight  $\tau^w$  and two ideological weights  $\phi_1^w$  and  $\phi_2^w$ . Topical weights are modulated by ideological weights through a multiplicative relationship, and all the weights are normalized through a logistic transformation. The graphical representation of the Joint Topic and Perspective Model is shown in Figure 4.



**Fig. 4.** A Joint Topic and Perspective model in a graphical model representation. A dashed line denotes a deterministic relation between parents and children nodes.

Given a set of  $D$  documents on a particular topic from differing ideological perspectives  $\{P_d\}$ , the joint posterior probability distribution of the topical and

ideological weights under the Joint Topic and Perspective model is

$$\begin{aligned}
& P(\tau, \{\phi_v\} | \{W_{d,n}\}, \{P_d\}; \Theta) \\
& \propto P(\tau | \mu_\tau, \Sigma_\tau) \prod_v P(\phi_v | \mu_\phi, \Sigma_\phi) \prod_{d=1}^D P(P_d | \pi) \prod_{n=1}^{N_d} P(W_{d,n} | P_d, \tau, \{\phi_v\}) \\
& = N(\tau | \mu_\tau, \Sigma_\tau) \prod_v N(\phi_v | \mu_\phi, \Sigma_\phi) \prod_d \text{Bernoulli}(P_d | \pi) \prod_n \text{Multinomial}(W_{d,n} | P_d, \beta),
\end{aligned}$$

where  $N(\cdot)$ ,  $\text{Bernoulli}(\cdot)$  and  $\text{Multinomial}(\cdot)$  are the probability density functions of multivariate normal, Bernoulli, and multinomial distributions, respectively.

The joint posterior probability distribution of  $\tau$  and  $\{\phi_v\}$ , however, are computationally intractable because of the non-conjugacy of the logistic-normal prior. We have developed an approximate inference algorithm [4]. The approximate inference algorithm is based on variational methods, and parameters are estimated using variational Expectation Maximization [5].

To predict a web video’s ideological perspective is to calculate the following conditional probability,

$$\begin{aligned}
& P(\tilde{P}_d | \{P_d\}, \{W_{d,n}\}, \{\tilde{W}_n\}; \Theta) \\
& = \int \int P(\{\phi_v\}, \tau | \{P_d\}, \{W_{d,n}\}, \{\tilde{W}_n\}; \Theta) \\
& \quad P(\tilde{P}_d | \{\tilde{W}_n\}, \tau, \{\phi_v\}; \Theta) d\tau d\phi_v
\end{aligned} \tag{1}$$

The predictive probability distribution in 1 is not computationally tractable, and we approximate it by plugging in the expected values of  $\tau$  and  $\{P_d\}$  obtained in variational inference.

### 3 Experiments

#### 3.1 Data

We collected web videos expressing opinions on various political and social issues from YouTube<sup>4</sup>. To identify web videos expressing a particular ideological perspective on an issue, we selected “code words” for each ideological perspective, and submitted the code words as query to YouTube. All of the returned web videos are labeled as expressing the particular ideological perspective. For example, the query words for the “pro-life” perspective on the abortion issue are “pro-life” and “abortion.”

We downloaded web videos and associated tags for 16 ideological views in May 2008 (two main ideological perspectives for eight issues), as listed in Table 1. Tags are keywords voluntarily added by authors or uploaders<sup>5</sup>. The total number of downloaded videos and associated tags are shown in Table 2. Note that the

<sup>4</sup> <http://www.youtube.com/>.

<sup>5</sup> <http://www.google.com/support/youtube/bin/answer.py?hl=en&answer=55769>

Issue	View 1	View 2
1 Abortion	pro-life	pro-choice
2 Democratic party primary election in 2008	pro-Hillary	pro-Obama
3 Gay rights	pro-gay	anti-gay
4 Global warming	supporter	skeptic
5 Illegal immigrants to the United States	Legalization	Deportation
6 Iraq War	pro-war	anti-war
7 Israeli-Palestinian conflict	pro-Israeli	pro-Palestinian
8 United States politics	pro-Democratic	pro-Republican

**Table 1.** Eight political and social issues and their two main ideological perspectives

	total videos	total tags	vocabulary
1	2850	30525	4982
2	1063	13215	2315
3	1729	18301	4620
4	2408	27999	4949
5	2445	25820	4693
6	2145	25766	4634
7	1975	22794	4435
8	2849	34222	6999

**Table 2.** The total number of downloaded web videos, the total number of tags, and the vocabulary size (the number of unique tags) for each issue

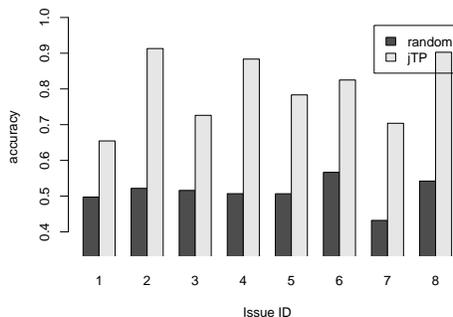
number of downloaded videos is equal to less than the total number of videos returned by YouTube due of the limit on the maximum number of search results in YouTube APIs.

We assume that web videos containing the “code words” of an ideological perspective in tags or descriptions convey the particular view, but this assumption may not be true. YouTube and many web video search engines are so far not designed to retrieve videos expressing opinions on an issue, let along to retrieve videos expressing a particular ideological view using keywords. Moreover, a web video may mention the code words of an ideological perspective in titles, descriptions, or tags but without expressing any opinions on an issue. For example, a news clip tagged with “pro-choice” may simply report a group of pro-choice activists in a protest and do not express strongly a so-called pro-choice point of view on the abortion issue.

### 3.2 Identifying Videos’ Ideological Perspectives

We evaluated how well a web video’s ideological perspective can be identified based on associated tags in a classification task. For each issue, we trained a binary classifier based on the Joint Topic and Perspective model in Section 2,

and applied the classifier on a held-out set. We reported the average accuracy of the 10-fold cross-validation. We compared the classification accuracy using the Joint Topic and Perspective Model with a baseline that randomly guesses one of two ideological perspectives. The accuracy of a random baseline is close but not necessarily equal to 50% because the number of videos in each ideological perspective on an issue are not necessarily equivalent.



**Fig. 5.** The accuracies of classifying a web video’s ideological perspective on eight issues

The experimental results in Figure 5 are very encouraging. The classifiers based on the Joint Topic and Perspective Model (labeled as jTP in Figure 5) outperform the random baselines for all eight political and social issues. The positive results suggest that the ideological perspectives of web videos can be identified using associated tags. Note that because the labels of our data are noisy, the results should be considered as a lower bound. The actual performance may be further improved if less noisy labels are available.

The positive classification results also suggest that Internet users sharing similar ideological beliefs on an issue appear to author, upload, and share similar videos, or at least, to tag similarly. Given that these web videos are uploaded and tagged at different times without coordination, it is surprising to see any pattern of tags emerging from folksonomies of web videos on political and social issues. Although the theory of ideology has argued that people sharing similar ideological beliefs use similar rhetorical devices for expressing their opinions in the mass media [6], we are the first to observe this pattern of tags in user-generated videos.

The non-trivial classification accuracy achieved by the Joint Topic and Perspectives Model suggests that the statistical model seem to closely match the real data. Although the Joint Topic and Perspective Model makes several modeling assumptions, including a strong assumption on the independence between tags (through a multinomial distribution), the high classification accuracy supports that these assumptions are not violated by the real data too much.



pollution energy green environment oil eco gas renewable nature  
 conservation coal ecology health sustainable air globalwarming water recycle  
 environmental emissions planet alternative solar comedy bbc politics 2008  
 democrats sea polar save power earth day the sustainability war ice mccain  
 clinton greenhouse clean tv fuel edwards election social house melting on carbon  
 david live music change car climate michael richard peace news obama  
 global warming sun to greenpeace hot commercial video bush un hillary  
 funny of gotcha documentary political president co2 al gore science an  
 effect inconvenient grassroots john government dioxide commentary in george  
 analysis outreach truth nonprofit canada weather public jones media alex  
 kyoto new tax beck robert debate skeptic crisis swindle hoax  
 scam nwo paul world fraud order god great false abc is exposed  
 invalid lies bosneanu sorin

**Fig. 7.** The color text cloud summarizes the topical and ideological weights learned in the web videos expressing contrasting ideological perspectives on the global warming issue. The larger a word's size, the larger its topical weight. The darker a word's color shade, the more extreme its ideological weight. Red represents the ideology of global warming **supporters**, and blue represents the ideology of global warming **skeptics**. The words are ordered by ideological weights, from strongly supporting global warming (red) to strongly skeptical about global warming (blue).

We do not intend to give a full analysis of why each ideology chooses and emphasizes these tags, but to stress that folksonomies of the ideological videos on the Internet are a rich resource to be tapped. Our experimental results in Section 3.2 and the analysis in this section show that by learning patterns of tags associated with web videos, we can identify web videos' ideological perspectives on various political and social issues with high accuracy.

Folksonomies mined from video sharing sites such as YouTube contain up-to-date information that other resources may lack. Due to the data collection time coinciding with the United States presidential election, many videos are related to presidential candidates and their views on various issues. The names of presidential candidates occur often in tags, and their views on various social and political issues become discriminative features (e.g., **Ron Paul's** pro-life position on the abortion issue in Figure 6). Ideological perspective classifiers should build on folksonomies of web videos to take advantage of these discriminative features. Classifiers built on static resources may fail to recognize these current, but very discriminative, tags.

## 4 Related Work

We borrow statistically modeling and inference techniques heavily from research on topic modeling (e.g., [7], [8] and [9]). They focus mostly on modeling text collections that containing many *different* (latent) topics (e.g., academic conference papers, news articles, etc). In contrast, we are interested in modeling

ideology texts that are mostly on the *same* topic but mainly differs in their ideological perspectives. There have been studies going beyond topics (e.g., modeling authors [10]). In this paper we are interested in modeling lexical variation collectively from multiple authors sharing similar beliefs, not lexical variations due to individual authors' writing styles and topic preference.

## 5 Conclusion

We propose to identify the ideological perspective of a web video on an issue using associated tags. We show that the statistical patterns of tags emerging from folksonomies can be successfully learned by a Joint Topic and Perspective Model, and the ideological perspectives of web videos on various political and social issues can be automatically identified with high accuracy. Web search engines and many Web 2.0 applications can incorporate our method to organize and retrieve web videos based on their ideological perspectives on an issue.

## References

1. comScore: YouTube.com accounted for 1 out of every 3 u.s. online videos viewed in january. <http://www.comscore.com/press/release.asp?press=2111>Please (March 2008)
2. techPresident: YouTube stats. <http://www.techpresident.com/youtube> (June 2008)
3. Naphade, M.R., Smith, J.R.: On the detection of semantic concepts at TRECVID. In: Proceedings of the Twelfth ACM International Conference on Multimedia. (2004)
4. Lin, W.H., Xing, E., Hauptmann, A.: A joint topic and perspective model for ideological discourse. In: Proceedings of the 2008 European Conference on Machine Learning and Principles (ECML) and Practice of Knowledge Discovery in Databases (PKDD). (2008)
5. Attias, H.: A variational bayesian framework for graphical models. In: Advances in Neural Information Processing Systems 12. (2000)
6. Van Dijk, T.A.: Ideology: A Multidisciplinary Approach. Sage Publications (1998)
7. Hofmann, T.: Probabilistic latent semantic indexing. In: Proceedings of the 22nd Annual International ACM SIGIR Conference on Research and Development in Information Retrieval. (1999) 50–57
8. Blei, D.M., Ng, A.Y., Jordan, M.: Latent Dirichlet allocation. Journal of Machine Learning Research **3** (January 2003) 993–1022
9. Griffiths, T.L., Steyvers, M.: Finding scientific topics. Proceedings of the National Academy of Sciences **101** (2004) 5228–5235
10. Rosen-Zvi, M., Griffiths, T., Steyvers, M., Smyth, P.: The author-topic model for authors and documents. In: Proceedings of the 20th Conference on Unvertainty in Artificial Intelligence. (2004)