

Identifying Stance by Analyzing Political Discourse on Twitter

Kristen Johnson

Purdue University
West Lafayette, IN 47907, USA
john1187@purdue.edu

Dan Goldwasser

Purdue University
West Lafayette, IN 47907, USA
dgoldwas@purdue.edu

Abstract

Politicians often use Twitter to express their beliefs, stances on current political issues, and reactions concerning national and international events. Since politicians are scrutinized for what they choose or neglect to say, they craft their statements carefully. Thus despite the limited length of tweets, their content is highly indicative of a politician’s stances. We present a weakly supervised method for understanding the stances held by politicians, on a wide array of issues, by analyzing how issues are framed in their tweets and their temporal activity patterns. We combine these components into a global model which collectively infers the most likely stance and agreement patterns.

In this work we present a novel approach for modeling the microblogging activity of presidential candidates and other prominent politicians. We look into two aspects of the problem, *stance* prediction over a wide array of issues, as well as *agreement and disagreement* patterns between politicians over these issues. While the two aspects are related, we argue they capture different information, as identifying agreement patterns reveals alliances and rivalries between candidates, across and inside their party. We show that understanding the political discourse on microblogs requires modeling both the content of posted messages as well as the social context in which they are generated, and suggest a joint model capturing both aspects.

1 Introduction

Recently the popularity of traditional media outlets such as television and printed press has decreased, causing politicians to turn their attention to social media outlets, which allow them to directly access the public, express their beliefs, and react to current events. This trend emerged during the 2008 U.S. presidential election campaign and has since moved to the mainstream – in the 2016 campaign, all candidates employ social media platforms. One of the most notable examples of this trend is the microblogging outlet Twitter, which unlike its predecessors, requires candidates to compress their ideas, political stances, and reactions into 140 character long tweets. As a result, candidates have to cleverly choose how to frame controversial issues, as well as react to events and each other (Mejova et al., 2013; Tumasjan et al., 2010).

Converse to other works predicting stance per individual tweet (SemEval, 2016), we use the *overall* Twitter behavior to predict a *politician’s* stance on an issue. We argue that these settings are better suited for the political arena on Twitter. Given the limit of 140 characters, the stance relevance of a tweet is not independent of the social context in which it was generated. In an extreme case, even the lack of Twitter activity on certain topics can be indicative of a stance. Additionally, framing issues in order to create bias towards their stance is a tool often used by politicians to contextualize the discussion (Tsur et al., 2015; Card et al., 2015; Boydston et al., 2014). Previous works exploring framing analyze text in traditional settings, such as congressional speeches or newspaper articles. To apply framing analysis to Twitter data, we allow tweets to hold multiple frames when necessary, as we find that on average many tweets are relevant to two frames per issue. This approach allows our model to make use of changing and similar framing patterns over

- (1) **Hillary Clinton (@HillaryClinton):** We need to keep **guns** out of the hands of **domestic abusers** and **convicted stalkers**.
- (2) **Donald Trump (@realDonaldTrump):** Politicians are trying to chip away at the **2nd Amendment**. I won't let them take away our **guns**!
- (3) **Bernie Sanders (@SenSanders):** We need sensible **gun-control** legislation which prevents **guns** from being used by **people who should not have them**.

Figure 1: Tweets on the issue of gun control, highlighting issue indicators in green and different frame indicators in yellow.

politicians' timelines in order to increase our prediction accuracy.

For example, consider the issue of gun control. Figure 1 shows three issue-related tweets by three politicians. To correctly identify the stance taken by each of the politicians, our model must combine three aspects. First, the relevance of these tweets to the question can be identified using *issue* indicators (marked in green). Second, the similarity between the stances taken by two of the three politicians can be identified by observing how the issue is *framed* (marked in yellow). In this example, tweets (1) and (3) frame the issue of gun control as a matter of safety, while (2) frames it as an issue related to personal freedom, thus revealing the agreement and disagreement patterns between them. Finally, we note the strong negative sentiment of tweet (1). Notice that each aspect individually might not contain sufficient information for correct classification, but combining all three, by propagating the stance bias (derived from analyzing the negative sentiment of (1)) to politicians likely to hold similar or opposing views (derived from frame analysis), leads to a more reliable prediction.

Given the dynamic nature of this domain, we design our approach to use minimal supervision and naturally adapt to new issues. Our model builds on several weakly supervised local learners that use a small seed set of issue and frame indicators to characterize the stance of tweets (based on lexical heuristics (O'Connor et al., 2010) and framing dimensions (Card et al., 2015)) and activity statistics which capture temporally similar patterns between politicians' Twitter activity. Our final model represents agreement and stance bias by combining these weak models into a weakly supervised joint model through Probabilistic Soft Logic (PSL), a recently

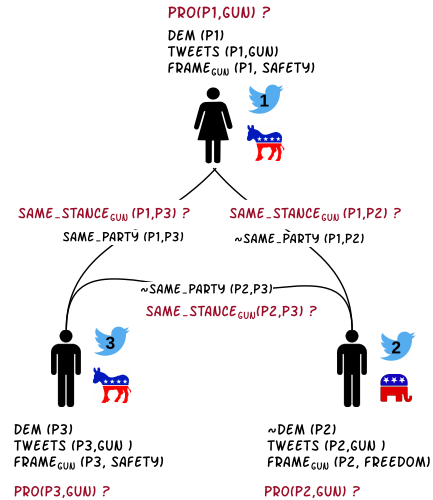


Figure 2: Relational Representation of Politicians' Twitter Activity. P1, P2, and P3 represent 3 different politicians. GUN refers to the issue of gun control; SAFETY and FREEDOM refer to different frames. Prediction target predicates are marked in red.

introduced probabilistic modeling framework (Bach et al., 2013). PSL combines these aspects declaratively by specifying high level rules over a relational representation of the politicians' activities (as shown in Figure 2), which is further compiled into a graphical model called a hinge-loss Markov random field (Bach et al., 2013), and used to make predictions about stance and agreement between politicians.

We analyze the Twitter activity of 32 prominent U.S. politicians, some of which were candidates for the U.S. 2016 presidential election. We collected their recent tweets and stances on 16 different issues, which were used for evaluation purposes. Our experiments demonstrate the effectiveness of our global modeling approach which outperforms the weak learners that provide the initial supervision.

2 Related Work

To the best of our knowledge this is the first work to use Twitter data, *specifically content, frames, and temporal activity*, to predict *politicians' stances*. Previous works (Sridhar et al., 2015; Hasan and Ng, 2014; Abu-Jbara et al., 2013; Walker et al., 2012; Abbott et al., 2011; Somasundaran and Wiebe, 2010; Somasundaran and Wiebe, 2009) have studied mining opinions and predicting stances in online debate forum data, exploiting argument and threaded conversation structures, or analyzed social interaction

and group structure (Sridhar et al., 2015; Abu-Jbara et al., 2013; West et al., 2014). In our Twitter dataset, there were few “@” mention or retweet examples forming a conversation concerning the investigated issues, thus we did not have access to argument or conversation structures for analysis. Works which focus on inferring signed social networks (West et al., 2014), stance classification (Sridhar et al., 2015), social group modeling (Huang et al., 2012), and PSL collective classification (Bach et al., 2015) are closest to our work, but these typically operate in supervised settings. In this work, we use PSL *without direct supervision*, to assign *soft* values (in the range of 0 to 1) to output variables ¹.

Using Twitter to analyze political discourse and influence has gained in popularity over recent years. Predicting characteristics of Twitter users, including political party affiliation has been explored (Volkova et al., 2015; Volkova et al., 2014; Conover et al., 2011). Previous works have also focused on sentiment analysis (Pla and Hurtado, 2014; Bakliwal et al., 2013), predicting ideology (Djemili et al., 2014), analyzing types of tweets and Twitter network effects around political events (Maireder and Ausserhofer, 2013), automatic polls based on Twitter sentiment and political forecasting using Twitter (Birmingham and Smeaton, 2011; O’Connor et al., 2010; Tumasjan et al., 2010), as well as uses of distant supervision (Marchetti-Bowick and Chambers, 2012).

Analyzing political tweets specifically has also attracted considerable interest. Recently, SemEval Task 6 (SemEval, 2016) aimed to detect the stance of *individual tweets*. Unlike this task and most related work on stance prediction (e.g., those mentioned above), we *do not assume that each tweet expresses a stance*. Instead, we combine tweet content and temporal indicators into a representation of a politician’s overall Twitter behavior, to determine if these features are indicative of a *politician’s stance*. This approach allows us to capture when politicians fail to tweet about a topic, which indicates a lack of stance as well.

To the best of our knowledge, this work is also the *first attempt to analyze issue framing in Twitter data*. To do so we used the frame guidelines developed by Boydston et al. (2014). Issue framing is related to

¹Conversely, Markov Logic Networks assign *hard* (0 or 1) values to model variables.

both analyzing biased language (Greene and Resnik, 2009; Recasens et al., 2013) and subjectivity (Wiebe et al., 2004). Several previous works have explored topic framing of public statements, congressional speeches, and news articles (Tsur et al., 2015; Card et al., 2015; Baumer et al., 2015). Other works focus on identifying and measuring political ideologies (Iyyer et al., 2014; Bamman and Smith, 2015; Sim et al., 2013; Lewenberg et al., 2016) and policies (Gerrish and Blei, 2012; Nguyen et al., 2015; Grimmer, 2010).

Finally, unsupervised and weakly supervised models of Twitter data for several various tasks have been suggested, such as user profile extraction (Li et al., 2014b), life event extraction (Li et al., 2014a), and conversation modeling (Ritter et al., 2010). Further, Eisenstein (2013) discusses methods for dealing with the unique language used in micro-blogging platforms.

3 Data and Problem Setting

REPUBLICAN POLITICIANS

Jeb Bush, Ben Carson, Chris Christie, Ted Cruz, Carly Fiorina, Lindsey Graham, Mike Huckabee, Bobby Jindal, John Kasich, George Pataki, Rand Paul, Rick Perry, Marco Rubio, Rick Santorum, Donald Trump, Scott Walker

DEMOCRATIC POLITICIANS

Joe Biden, Lincoln Chafee, Hillary Clinton, Kirsten Gillibrand, John Kerry, Ben Lujan, Ed Markey, Martin O’Malley, Nancy Pelosi, Harry Reid, Bernie Sanders, Chuck Schumer, Jon Tester, Mark Warner, Elizabeth Warren, Jim Webb

Table 1: Politicians tracked in this study.

Collection and Pre-Processing of Tweets: We collected tweets for the 32 politicians listed in Table 1, initially beginning with those politicians participating in the 2016 U.S. presidential election (16 Republicans and 5 Democrats). To increase representation of Democrats, we collected tweets of Democrats who hold leadership roles within their party. For all 32 politicians we have a total of 99,161 tweets, with an average of 3,000 per person. There are 39,353 Democrat and 59,808 Republican tweets.

Using tweets from both parties, we compiled a set of frequently appearing keywords for each issue, with an average of seven keywords per issue. A Python script was then used on these preselected keywords to filter all tweets, keeping only those that represent our 16 political issues of interest (shown in

ISSUE	QUESTION
ABORTION	<i>Do you support abortion?</i>
ACA	<i>Do you support the Patient Protection and Affordable Care Act (Obamacare)?</i>
CONFEDERATE	<i>Should the federal government allow states to fly the confederate flag?</i>
DRUGS	<i>Do you support the legalization of Marijuana?</i>
ENVIRONMENT	<i>Should the federal government continue to give tax credits and subsidies to the wind power industry?</i>
GUNS	<i>Do you support increased gun control?</i>
IMMIGRATION	<i>Do you support stronger measures to increase our border security?</i>
IRAN	<i>Should the U.S. conduct targeted airstrikes on Irans nuclear weapons facilities?</i>
ISIS	<i>Should the U.S. formally declare war on ISIS?</i>
MARRIAGE	<i>Do you support the legalization of same sex marriage?</i>
NSA	<i>Do you support the Patriot Act?</i>
PAY	<i>Should employers be required to pay men and women, who perform the same work, the same salary?</i>
RELIGION	<i>Should a business, based on religious beliefs, be able to deny service to a customer?</i>
SOCIAL SECURITY	<i>Should the government raise the retirement age for Social Security?</i>
STUDENT	<i>Would you support increasing taxes on the rich in order to reduce interest rates for student loans?</i>
TPP	<i>Do you support the Trans-Pacific Partnership?</i>

Table 2: Issues taken from ISideWith.com and their corresponding Yes/No questions. Each issue serves as a prediction target for each politician. For example, for each politician we predict if they are for (PRO) or against (\neg PRO) increased gun control (GUNS), as well as if every pair of politicians shares the same stance for that issue (SAMESTANCE_I).

Table 2), and automatically eliminating all irrelevant tweets (e.g., those about personal issues).

Annotating Stances and Agreement: We used ISideWith.com, a popular website that matches users to politicians based on their answers to a series of 58 questions, to choose 16 of these issues (shown in Table 2) for our prediction goals. ISideWith.com uses a range of yes/no answers in their questions and provides proof of the politician’s stance on that issue, *if available*, through public information such as quotes. Since we use the stances as the ground truth for evaluating our prediction, all politicians with unavailable answers or those not listed on the site were manually annotated via online searches of popular newspapers, political channels, and voting records. Since ISideWith.com does not contain answers to all questions for all politicians, especially those that are less popular, we design our approach to be generalizable to such situations by requiring minimal supervision.

Predicting Stance and Agreement: Based on the collected stances, which represent our ground truth of whether a politician is for or against an issue, we define two target predicates using PSL notation (see Section 4.1) to capture the desired output as soft truth assignments to these predicates. The first predicate, PRO(P1, ISSUE) captures the idea that politician P1 is in support of an ISSUE. Consequently,

an opposing stance would be captured by the negation: \neg PRO(P1, ISSUE). In this work, we do not make use of stance correlations among party members (Lewenberg et al., 2016; Maddox and Lilie, 1984). For example, in U.S. politics Republicans are known to be against gun control and abortion, while Democrats support both issues. Since we are interested in determining the effectiveness of our local models (described in Section 4.2) to capture the stance of each politician, we do not encode such cross-issue information into the models. Additionally, in a weakly supervised setting, we assume we do not have access to such information.

The second target predicate, SAMESTANCE_I(P1, P2) classifies if two politicians share a stance for a given issue, i.e., if both are for or against an issue, where *I* represents 1 of the 16 issues being investigated. Although the two predicates are clearly inter-dependent, we model them as separate predicates since they can depend on different Twitter behavioral and content cues and we can often identify indicators of shared stance, without mention of the actual stance.

4 From Local to Global Models of Twitter Activity

Our approach uses a collection of weakly supervised local models to capture the similarities between stance bias, tweet content, and temporal ac-

tivity patterns of users’ timelines. These local models are used to provide the initial bias when learning the parameters of the global PSL model, which uses PSL to combine all of the local models together into a joint global model. In addition to the PSL local model predicates (described below), we also use directly observed information: party affiliation, denoted $\text{DEM}(P1)$ for Democrat and $\neg\text{DEM}(P1)$ for Republican, and $\text{SAMEPARTY}(P1, P2)$ to denote if two politicians belong to the same party. As shown by the baseline measurements in Section 5, local information alone is not strong enough to capture stance or agreement for politicians. However, by using PSL, we are able to build connections between each local model in order to increase the overall accuracy of each global model’s prediction.

4.1 Global Modeling using PSL

PSL is a recent declarative language for specifying weighted first-order logic rules. A PSL model is specified using a set of weighted logical formulas, which are compiled into a special class of graphical model, called a hinge-loss MRF, defining a probability distribution over the possible continuous value assignments to the model’s random variables and allowing the model to scale easily (Bach et al., 2015). The defined probability density function has the form:

$$P(\mathbf{Y} | \mathbf{X}) = \frac{1}{Z} \exp \left(- \sum_{r=1}^M \lambda_r \phi_r(\mathbf{Y}, \mathbf{X}) \right)$$

where λ is the weight vector, Z is a normalization constant, and

$$\phi_r(\mathbf{Y}, \mathbf{X}) = (\max\{l_r(\mathbf{Y}, \mathbf{X}), 0\})^{\rho_r}$$

is the hinge-loss potential corresponding to the instantiation of a rule, specified by a linear function l_r , and an optional exponent $\rho_r \in 1, 2$. The weights of the rules are learned using maximum-likelihood estimation, which in our weakly supervised setting was estimated using the Expectation-Maximization algorithm. For more details we refer the reader to Bach et al. (2015).

Specified PSL rules have the form:

$$\begin{aligned} \lambda_1 &: P_1(x) \wedge P_2(x, y) \rightarrow P_3(y), \\ \lambda_2 &: P_1(x) \wedge P_4(x, y) \rightarrow \neg P_3(y) \end{aligned}$$

where P_1, P_2, P_3, P_4 are predicates, and x, y are variables. Each rule is associated with a weight λ , which indicates its importance in the model. Given concrete constants a, b respectively instantiating the

variables x, y , the mapping of the model’s atoms to soft $[0,1]$ assignments will be determined by the weights assigned to each one of the rules. For example, if $\lambda_1 > \lambda_2$, the model will prefer $P_3(b)$ to its negation. This contrasts with “classical” or other probabilistic logical models in which rules are strictly true or false. In our work, the constant symbols correspond to politicians and predicates represent party affiliation, Twitter activities, and similarities between politicians based on Twitter behaviors.

4.2 Local Models of Basic Twitter Activity

Issue: We use a keyword based heuristic, similar to the approach described in O’Connor et al. (2010), to capture which issues politicians are tweeting about. Each issue is associated with a small set of keywords, which may be mutually exclusive, such as those concerning Iran or Environment. However, some may fall under multiple issues at once (e.g., *religion* may indicate the tweet refers to ISIS, Religion, or Marriage). The majority of matching keywords determines the issue of the tweet, with rare cases of ties manually resolved. The output of this classifier is all of the issue-related tweets of a politician, which are used as input for the PSL predicate $\text{TWEETS}(P1, \text{ISSUE})$. This binary predicate indicates if politician P1 has tweeted about the issue or not.

Sentiment Analysis: Based on the idea that the sentiment of a tweet can help expose a politician’s stance on a certain issue, we use OpinionFinder 2.0 (Wilson et al., 2005) to label each politician’s issue-related tweets as positive, negative, or neutral. We observed, however, that for all politicians, a majority of tweets will be labeled as neutral. This may be caused by the difficulty of labeling sentiment for Twitter data. If a politician has no positive or negative tweets, they are assigned their party’s majority sentiment assignment for that issue. This output is used as input to the PSL predicates $\text{TWEETPOS}(P1, \text{ISSUE})$ and $\text{TWEETNEG}(P1, \text{ISSUE})$.

Agreement and Disagreement: To determine how well tweet content similarity can capture stance agreement, we computed the pair-wise cosine similarity between all of the politicians. Due to the usage of similar words per issue, most politicians are grouped together, even across different parties. To overcome this noise, we compute the *frequency* of similar words within tweets about each issue. For

PSL MODEL	EXAMPLE OF PSL RULE
LOCAL BASELINE (LB)	$\text{LOCALSAMESTANCE}_I(P1, P2) \rightarrow \text{SAMESTANCE}_I(P1, P2)$ $\text{TWEETS}(P1, \text{ISSUE}) \wedge \text{TWEETPOS}(P1, \text{ISSUE}) \rightarrow \text{PRO}(P1, \text{ISSUE})$
MODEL 1 (M1)	$\text{SAMEPARTY}(P1, P2) \wedge \text{DEM}(P1) \rightarrow \text{PRO}(P2, \text{ISSUE})$ $\text{SAMEPARTY}(P1, P2) \rightarrow \text{SAMESTANCE}_I(P1, P2)$
MODEL 2 (M2)	$\text{TWEETS}(P1, \text{ISSUE}) \wedge \text{DEM}(P1) \rightarrow \text{PRO}(P1, \text{ISSUE})$ $\text{TWEETPOS}(P1, \text{ISSUE}) \wedge \text{TWEETPOS}(P2, \text{ISSUE}) \rightarrow \text{SAMESTANCE}_I(P1, P2)$
MODEL 3 (M3)	$\text{LOCALSAMESTANCE}_I(P1, P2) \wedge \text{PRO}(P1, \text{ISSUE}) \rightarrow \text{PRO}(P2, \text{ISSUE})$ $\text{SAMETEMPORALACTIVITY}_I(P1, P2) \wedge \text{SAMEPARTY}(P1, P2) \rightarrow \text{SAMESTANCE}_I(P1, P2)$ $\text{FRAME}(P1, \text{ISSUE}) \wedge \text{FRAME}(P2, \text{ISSUE}) \rightarrow \text{SAMESTANCE}_I(P1, P2)$

Table 3: Subset of examples of the rules that are used by each PSL model. Each model also contains negated versions of rules, as well as similar rules where DEM has been replaced by \neg DEM to represent Republicans.

each issue, all of a politician’s tweets are aggregated and the frequency of each word is compared to all other politicians’ word frequencies. Politicians, P1 and P2, are considered to have a similar $\text{LOCALSAMESTANCE}_I(P1, P2)$ if their frequency counts per word for an issue are within the same range.

4.3 Baseline PSL Model: Using Local Models Directly

Previous stance classification works typically predict stance based on a single piece of text (e.g., forum posts or tweets) in a supervised setting, making it difficult to directly compare to our approach. To provide some comparison, we implement a baseline model which, as expected, has a weaker performance than our models. The baseline model does not take advantage of the global modeling framework, but instead learns weights over the rules listed in the first two lines of Table 3. These rules directly map the output of the local noisy models to PSL target predicates.

4.4 PSL Model 1: Party Based Agreement

The tendency of politicians to vote with their political party on most issues is encoded via the Model 1 PSL rules listed in Table 3, which aim to capture party based agreement. For some issues we initially assume Democrats (DEM) are for an issue, while Republicans (\neg DEM) are against that issue, or vice versa. In the latter case, the rules of the model would change, e.g. the second rule would become: \neg DEM(P1) \rightarrow PRO(P1, ISSUE), and likewise for all other rules. Similarly, if two politicians are in the same party, we expect them to have the same stance, or agree, on an issue. For all PSL rules, the reverse also holds, e.g., if two politicians are not in the same

party, we expect them to have different stances.

4.5 PSL Model 2: Basic Twitter Activity

Model 2 builds upon the initial party line bias of Model 1. In addition to political party based information, we also include representations of the politician’s Twitter activity, as shown in Table 3. This includes whether or not a politician tweets about an issue (TWEETS) and what sentiment is expressed in those tweets. The predicate TWEETPOS models if a politician tweets positively on the issue, whereas TWEETNEG models negative sentiment. Two different predicates are used instead of the negation of TWEETPOS, which would cause all politicians for which there are no tweets (or sentiment) on that issue to also be considered.

4.6 Local Models of High Level Twitter Activity

Temporal Activity Patterns: We observed from reading Twitter feeds that most politicians will comment on an event the day it happens. For general issues, politicians comment as often as desired to express their support or lack thereof for a particular issue. To capture patterns between politicians, we align their timelines based on days where they have tweeted about an issue. When two or more politicians tweet about the same issue on the same day, they are considered to have similar temporal activity, which may indicate stance agreement. This information is used as input for our PSL predicate $\text{SAMETEMPORALACTIVITY}_I(P1, P2)$.

Political Framing: The way politicians choose to contextualize their tweets on an issue is strongly indicative of their stance on that issue. To investigate this, we compiled a list of unique keywords for each

political framing dimension as described in Boyd-stun et al. (2014) and Card et al. (2015). We use the keyword matching approach described in Section 4.2 to classify all tweets into a political frame with some tweets belonging to multiple frames. We sum over the total number of each frame type and use the frame with the maximum and second largest count as that politician’s frames for that issue. In the event of a tie we assign the frame that appears most frequently within that politician’s party. These frames are used as input to the PSL predicate `FRAME(P1, ISSUE)`.

4.7 PSL Model 3: Agreement Patterns

The last three lines of Table 3 present a subset of the rules used in Model 3 to incorporate higher level Twitter information into the model. Our intuition is that politicians who tweet in a similar manner would also have similar stances on issues, which we represent with the predicate `LOCALSAMESTANCEI`. `SAMETEMPORALACTIVITY` represents the idea that if politicians tweet about an issue around the same times then they also share a stance for that issue. Finally, `FRAME` indicates the frame used by that politician for different issues. The use of these rules allows Model 3 to overcome Model 2 inconsistencies between stance and sentiment (e.g., if someone attacks their opposition).

5 Experiments

Experimental Settings: As described in Section 4, the data generated from the local models is used as input to the PSL models. Stances collected in Section 3 are used as the ground truth for evaluation of the results of the PSL models. We initialize Model 1, as described in Section 4.4, using political party affiliation knowledge. Model 2 builds upon Model 1 by adding the results of the issue and sentiment analysis local models. Model 3 combines all previous models with higher level Twitter activities: tweet agreement, temporal activity, and frames. We implement our PSL models to have an initial bias that candidates do not share a stance and are against an issue.

Experimental Results By Issue: Table 4 presents the results of using our three proposed PSL models. Local Baseline (LB) refers to using only the weak local models for prediction with no additional information about party affiliation. We observe that for prediction of stance (PRO) LB performs better

than random chance in 11 of 16 issues; for prediction of agreement (`SAMESTANCEI`), LB performs much lower overall, with only 5 of 16 issues predicted above chance.

Using Model 1 (M1), we improve stance prediction accuracy for 11 of the issues and agreement accuracy for all issues. Model 2 (M2) further improves the stance and agreement predictions for an additional 8 and 10 issues, respectively. Model 3 (M3) increases the stance prediction accuracy of M2 for 4 issues and the agreement accuracy for 9 issues. The final agreement predictions of M3 are significantly improved over the initial LB for all issues.

The final stance predictions of M3 are improved over all issues except Guns, Iran, and TPP. For Guns, the stance prediction remains the same throughout all models, meaning additional party information does not boost the initial predictions determined from Twitter behaviors. For Iran, the addition of M1 and M2 lower the accuracy, but the behavioral features from M3 are able to restore it to the original prediction. For TPP, this trend is likely due to the fact that all models incorporate party information and the issue of TPP is the most heavily divided within and across parties, with 8 Republicans and 4 Democrats in support of TPP and 8 Republicans and 12 Democrats opposed. Even in cases where M1 and/or M2 lowered the initial baseline result (e.g. stance for Religion or agreement for Environment), the final prediction by M3 is still higher than that of the baseline.

Framing and Temporal Information: As shown in Table 4, performance for *some* issues does not improve in Model 3. Upon investigation, we found that for all issues, except Abortion which improves in agreement, one or both of the top frames for the party are the same across party lines. For example, for ACA both Republicans and Democrats have the *Economic* and *Health and Safety* frames as their top two frames. For TPP, both parties share the *Economic* frame.

In addition to similar framing overlap, the Twitter timeline for ACA also exhibits overlap, as shown in Figure 3(a). This figure highlights one week before and after the Supreme Court ruling (seen as the peak of activity, 6/25/2015) to uphold the ACA. Conversely, Abortion, which shares no frames between parties (Democrats frame Abortion with *Constitutionality* and *Health and Safety* frames; Repub-

Issue	STANCE				AGREEMENT			
	LB	M 1	M 2	M 3	LB	M 1	M 2	M 3
ABORTION	81.25	96.88	96.88	96.88	49.31	93.75	93.75	95.36
ACA	96.88	100	100	100	51.61	100	100	100
CONFEDERATE	34.38	78.12	87.5	84.38	51.31	69.6	77.7	80.18
DRUGS	87.5	78.12	96.88	88.88	50.42	63.6	84.07	84.07
ENVIRONMENT	53.12	78.12	78.13	81.08	45.16	68.75	65.59	69.28
GUNS	93.75	93.75	93.75	93.75	48.59	68.54	99.59	99.59
IMMIGRATION	37.5	81.25	81.25	86.36	53.62	68.55	69.06	69.56
IRAN	84.38	65.62	65.63	84.38	35.57	79.73	100	100
ISIS	40.32	76.28	93.75	93.75	59.68	76.28	76.28	90.04
MARRIAGE	62.5	90.62	90.62	90.62	50.57	87.12	87.43	87.43
NSA	37.5	53.12	53.12	61.54	34.15	49.2	56.66	59.65
PAY	84.38	84.38	90.62	89.47	64.30	72.92	80.31	74.31
RELIGION	75	68.75	81.25	81.25	47.62	86.24	76.46	79.44
SOCIAL SECURITY	28.12	78.12	78.13	78.13	53.76	73.25	90.03	90.88
STUDENT	93.75	96.88	96.88	96.88	51.61	100	100	100
TPP	62.5	62.5	62.5	62.5	45.43	48.39	54.64	65.32

Table 4: Stance and Agreement Accuracy by Issue. LB uses weak local models, M1 represents party line agreement, M2 adds Twitter activity, and M3 adds higher level Twitter behaviors.

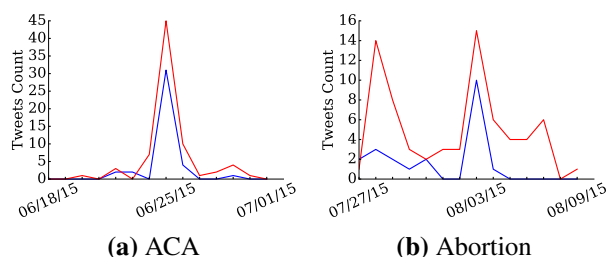


Figure 3: Temporal Twitter Activity by Party. Republican (red) and Democrat (blue) event based temporal overlaps.

licans use *Economic* and *Capacity and Resources* frames), exhibits a timeline with greater fluctuation. The peak of Figure 3(b) is 8/3/2015, which is the day that the budget was passed to include funding for Planned Parenthood. Overall both parties have different patterns over this time range, allowing M3 to increase agreement prediction accuracy by 1.61%.

6 Conclusion

In this paper we take a first step towards understanding the dynamic microblogging behavior of politicians. Though we concentrate on a small set of politicians and issues in this work, this framework can be modified to handle additional politicians or issues, as well as those in other countries, by incorporating appropriate domain knowledge (e.g., using new keywords for different issues in other countries), which we leave as future work. Unlike previous works, which tend to focus on one

aspect of this complex microblogging behavior, we build a holistic model connecting temporal behaviors, party-line bias, and issue frames into a single predictive model used to identify fine-grained policy stances and agreement. Despite having no explicit supervision, and using only intuitive “rules-of-thumb” to bootstrap our global model, our approach results in a strong prediction model which helps shed light on political discourse framing inside and across party lines.

Acknowledgements

We thank the anonymous reviewers for their thoughtful comments and suggestions.

References

- Rob Abbott, Marilyn Walker, Pranav Anand, Jean E. Fox Tree, Robeson Bowmani, and Joseph King. 2011. How can you say such things?!?: Recognizing disagreement in informal political argument. In *Proc. of the Workshop on Language in Social Media*.
- Amjad Abu-Jbara, Ben King, Mona Diab, and Dragomir Radev. 2013. Identifying opinion subgroups in arabic online discussions. In *Proc. of ACL*.
- Stephen H. Bach, Bert Huang, Ben London, and Lise Getoor. 2013. Hinge-loss Markov random fields: Convex inference for structured prediction. In *Proc. of UAI*.
- Stephen H Bach, Matthias Broecheler, Bert Huang, and Lise Getoor. 2015. Hinge-loss markov random

- fields and probabilistic soft logic. *arXiv preprint arXiv:1505.04406*.
- Akshat Bakliwal, Jennifer Foster, Jennifer van der Puil, Ron O'Brien, Lamia Tounsi, and Mark Hughes. 2013. Sentiment analysis of political tweets: Towards an accurate classifier. In *Proc. of ACL*.
- David Bamman and Noah A Smith. 2015. Open extraction of fine-grained political statements. In *Proc. of EMNLP*.
- Eric Baumer, Elisha Elovic, Ying Qin, Francesca Polletta, and Geri Gay. 2015. Testing and comparing computational approaches for identifying the language of framing in political news. In *Proc. of ACL*.
- Adam Bermingham and Alan F Smeaton. 2011. On using twitter to monitor political sentiment and predict election results.
- Amber Boydston, Dallas Card, Justin H. Gross, Philip Resnik, and Noah A. Smith. 2014. Tracking the development of media frames within and across policy issues.
- Dallas Card, Amber E. Boydston, Justin H. Gross, Philip Resnik, and Noah A. Smith. 2015. The media frames corpus: Annotations of frames across issues. In *Proc. of ACL*.
- Michael D Conover, Bruno Gonçalves, Jacob Ratkiewicz, Alessandro Flammini, and Filippo Menczer. 2011. Predicting the political alignment of twitter users. In *Proc. of Privacy, Security, Risk and Trust (PASSAT) and SocialCom*.
- Sarah Djemili, Julien Longhi, Claudia Marinica, Dimitris Kotzinos, and Georges-Elia Sarfati. 2014. What does twitter have to say about ideology? In *NLP 4 CMC: Natural Language Processing for Computer-Mediated Communication*.
- Jacob Eisenstein. 2013. What to do about bad language on the internet. In *Proc. of NAACL*.
- Sean Gerrish and David M Blei. 2012. How they vote: Issue-adjusted models of legislative behavior. In *Advances in Neural Information Processing Systems*, pages 2753–2761.
- Stephan Greene and Philip Resnik. 2009. More than words: Syntactic packaging and implicit sentiment. In *Proc. of NAACL*.
- Justin Grimmer. 2010. A bayesian hierarchical topic model for political texts: Measuring expressed agendas in senate press releases. In *Political Analysis*.
- Kazi Saidul Hasan and Vincent Ng. 2014. Why are you taking this stance? identifying and classifying reasons in ideological debates. In *Proc. of EMNLP*.
- Bert Huang, Stephen H. Bach, Eric Norris, Jay Pujara, and Lise Getoor. 2012. Social group modeling with probabilistic soft logic. In *NIPS Workshops*.
- Mohit Iyyer, Peter Enns, Jordan L Boyd-Graber, and Philip Resnik. 2014. Political ideology detection using recursive neural networks. In *Proc. of ACL*.
- Yoad Lewenberg, Yoram Bachrach, Lucas Bordeaux, and Pushmeet Kohli. 2016. Political dimensionality estimation using a probabilistic graphical model. In *Proc. of UAI*.
- Jiwei Li, Alan Ritter, Claire Cardie, and Eduard H Hovy. 2014a. Major life event extraction from twitter based on congratulations/condolences speech acts. In *Proc. of EMNLP*.
- Jiwei Li, Alan Ritter, and Eduard H Hovy. 2014b. Weakly supervised user profile extraction from twitter. In *Proc. of ACL*.
- William Maddox and Stuart Lillie. 1984. Beyond liberal and conservative: Reassessing the political spectrum.
- Axel Maireder and Julian Ausserhofer. 2013. National politics on twitter: Structures and topics of a networked public sphere. In *Information, Communication, and Society*.
- Micol Marchetti-Bowick and Nathanael Chambers. 2012. Learning for microblogs with distant supervision: Political forecasting with twitter. In *Proc. of EACL*.
- Yelena Mejova, Padmini Srinivasan, and Bob Boynton. 2013. Gop primary season on twitter: popular political sentiment in social media. In *WSDM*.
- Viet-An Nguyen, Jordan Boyd-Graber, Philip Resnik, and Kristina Miler. 2015. Tea party in the house: A hierarchical ideal point topic model and its application to republican legislators in the 112th congress. In *Proc. of ACL*.
- Brendan O'Connor, Ramnath Balasubramanian, Bryan R Routledge, and Noah A Smith. 2010. From tweets to polls: Linking text sentiment to public opinion time series. In *Proc. of ICWSM*.
- Ferran Pla and Lluís F Hurtado. 2014. Political tendency identification in twitter using sentiment analysis techniques. In *Proc. of COLING*.
- Marta Recasens, Cristian Danescu-Niculescu-Mizil, and Dan Jurafsky. 2013. Linguistic models for analyzing and detecting biased language. In *Proc. of ACL*.
- Alan Ritter, Colin Cherry, and Bill Dolan. 2010. Unsupervised modeling of twitter conversations. In *Proc. of NAACL*.
- SemEval. 2016. Task 6. <http://alt.qcri.org/semeval2016/task6/>.
- Yanchuan Sim, Brice DL Acree, Justin H Gross, and Noah A Smith. 2013. Measuring ideological proportions in political speeches. In *Proc. of EMNLP*.
- Swapna Somasundaran and Janyce Wiebe. 2009. Recognizing stances in online debates. In *Proc. of ACL*.

- Swapna Somasundaran and Janyce Wiebe. 2010. Recognizing stances in ideological on-line debates. In *Proc. of NAACL Workshops*.
- Dhanya Sridhar, James Foulds, Bert Huang, Lise Getoor, and Marilyn Walker. 2015. Joint models of disagreement and stance in online debate. In *Proc. of ACL*.
- Oren Tsur, Dan Calacci, and David Lazer. 2015. A frame of mind: Using statistical models for detection of framing and agenda setting campaigns. In *Proc. of ACL*.
- Andranik Tumasjan, Timm Oliver Sprenger, Philipp G Sandner, and Isabell M Welp. 2010. Predicting elections with twitter: What 140 characters reveal about political sentiment.
- Svitlana Volkova, Glen Coppersmith, and Benjamin Van Durme. 2014. Inferring user political preferences from streaming communications. In *Proc. of ACL*.
- Svitlana Volkova, Yoram Bachrach, Michael Armstrong, and Vijay Sharma. 2015. Inferring latent user properties from texts published in social media. In *Proc. of AAAI*.
- Marilyn A. Walker, Pranav Anand, Robert Abbott, and Ricky Grant. 2012. Stance classification using dialogic properties of persuasion. In *Proc. of NAACL*.
- Robert West, Hristo S Paskov, Jure Leskovec, and Christopher Potts. 2014. Exploiting social network structure for person-to-person sentiment analysis. *TACL*.
- Janyce Wiebe, Theresa Wilson, Rebecca Bruce, Matthew Bell, and Melanie Martin. 2004. Learning subjective language. *Computational linguistics*.
- Theresa Wilson, Paul Hoffmann, Swapna Somasundaran, Jason Kessler, Janyce Wiebe, Yejin Choi, Claire Cardie, Ellen Riloff, and Siddharth Patwardhan. 2005. Opinionfinder: A system for subjectivity analysis. In *Proc. of EMNLP*.