Computational Linguistics and Chinese Language Processing
Vol. 19, No. 2, June 2014, pp. 1-16

Social Metaphor Detection via Topical Analysis

Ting-Hao (Kenneth) Huang*

Abstract

With the massive amount of social media data becoming available, there is a rising interest in automatic metaphor detection and interpretation from open social text. One of the most well-known approaches to this subject is identifying the violation of selectional preference. The basic concept of selectional preference is that verbs tend to have semantic preferences of their arguments and that violations of these preferences are strong indicators of metaphorical language use. Nevertheless, previously, few works have focused on metaphor detection of social media data. In response to this problem, we propose a three-step framework that is based on the technology of selection preference modeling to detect metaphors in social media data. We conduct a pilot study of this framework on the data of a real-world online support group. Furthermore, to improve our approach, we also leverage topical analysis techniques in our framework. As a result, we address the challenges of the task of metaphor detection in social media data, provide qualitative analysis for our experiments, and illustrate our insight based on the results.

Keywords: Metaphor, Cluster, Selectional Preference, Social Media Data.

1. Introduction

With massive social media data, *e.g.*, comments, blog articles, or tweets, becoming available, there is a rising interest in automatic metaphor detection from open social text. One of the most well-known approaches to this subject is detecting the violation of selectional preference. The idea of selectional preference is that the predicates (*i.e.*, mostly verbs) tend to have semantic preferences of their arguments. For instance, the verb "flex" has a strong preference of "muscle" and "bone" as its object. If we find that, in some text, the object of "flex" is not of the semantic class of "muscle" and "bone," it is very likely to be a metaphorical use.

Previously, researchers have studied metaphor identification by modeling selectional preference (Loenneker-Rodman & Narayanan, 2010; Shutova *et al*., 2010; Shutova, 2010;

* Language Technologies Institute, Carnegie Mellon University, 5000 Forbes Ave., Pittsburgh, PA 15213, USA
E-mail: windx@cmu.edu

Resnik, 1997; Shutova & Teufel, 2010; Calzolari *et al*., 2010; Preiss *et al*., 2007), while few papers have focused on social media data. In our work, we call the metaphors occurring in social media “social metaphor” to emphasize their different properties and difficulty.

Furthermore, to improve the technology of metaphor detection, we also leverage topic analysis techniques in our approach. The intuition behind combining metaphor identification and topic analysis is that both verbs and arguments exhibit strong tendencies towards a few specific topics, and this topical information provides additional evidence to facilitate identification of selectional preference among text. For instance, in the topic of sports, the subjects of “flex” are mostly humans; but in the topic of finance or politics, the subjects of “flex” are mostly organizations or countries, *e.g*., “*China to flex its financial muscles at US meeting*.” In this paper, we study how the metaphor detection technique can be influenced by topical analysis techniques.

The problem of automatic social metaphor detection poses two main challenges. First, as social media data is usually noisy, how to effectively preprocess the input texts before an actual detection component is employed should be studied carefully. We should estimate empirically the performance of existing NLP tools, especially lemmatizers and POS taggers. Second, how to apply and evaluate the proposed approach on a real world data set is not straight-forward. As there is neither an existing data set nor benchmark to evaluate metaphor detection, we need to create a benchmark that can show the performance difference effectively.

Furthermore, incorporating topical analysis into metaphor detection has another layer of challenges: how to automatically discover the topical distribution for each term (including verbs and nouns) within open text, which is not a trivial problem. Moreover, we need to study how to leverage the topical distribution of each verb and noun to metaphor detection.

In this paper, we will define the problem before proposing our 3-step approach for meta-phor detection. Specifically, we first preprocess the input text by extracting tokens and further clustering nouns then detect selectional association outliers. Finally, we apply a selectional preference strength filter to extract metaphor-embedded text snippets.

We then conduct experiments on a real-world social media data set. The LDA model is applied to partition the input corpus based on topics, and we adopt the 3-step approach both on the whole corpus and on every single topic data partition. Finally, we compare the metaphor detection results between those with and without the influence of topics, and we observe which one performs better.

The rest of the paper is organized as follows: In Section 2, we summarize related work for metaphor detection based on selectional preference detection. In Section 3, we formally de-fine the problem of automatic social metaphor detection. Then, in Section 4, we conduct a preliminary test to compare two technologies for metaphor detection and choose one to establish the 3-step framework we will describe in Section 5. In Section 6, we further discuss the details of topic analysis. Finally, we demonstrate the experiment in Section 7, discuss the results in Section 8, and conclude the work in Section 9.

2. Related Work

In this section, we briefly survey papers that investigate approaches to detect metaphors in text.

2.1 Automatic Metaphor Detection

There have been many computational approaches in the field of natural language processing toward modeling metaphors. Based on Shutova *et al*. (2010), the research of modeling meta-phors could be divided into two sub-fields: metaphor detection and metaphor interpretation. In this paper, we focus on metaphor detection. In this field, the first challenge is how to define a metaphor. As mentioned in Loenneker-Rodman and Narayanan (2010), "*there is rich continuing theoretical debate on the definition and use of metaphor*." In our work, we limited the scope of our research in that we only aim to detect a "non-conventionalized metaphor," which usually has low frequency and could reasonably be considered as an outlier of selectional preferences. For instance, conventional metaphors like "Life is a journey" or "Time is running out," which would not strongly violate the selectional preference, are considered to be out of scope of this work.

In the field of metaphor detection, the Met* System (Fass, 1991) can be considered the first attempt to explore this field, and the following approaches include (Goatly, 1997), (Peters & Peters, 2000), CorMet System (Mason, 2004), and TroFi System (Birke & Sarkar, 2006). Most of them adopt the concept of selectional preference that we mentioned above, along with some hand-coded knowledge base, *e.g*., VerbNet. VerbNet contains information about the constraint of arguments of verbs. By matching the text with the verb and its argument, we are able to detect the violation of arguments. Nevertheless, in this paper, we apply a different approach that learns the violations directly from statistics based on natural texts. One advantage of this approach is that we do not need any hand-coded knowledge base, so it could be ported to other languages more easily.

2.2 Topical Analysis

Many topical analysis techniques have been developed, *e.g*., latent semantic analysis, proba-bilistic LSA, NMF, and LDA. Latent Dirichlet Allocation (LDA) (Blei, Ng, & Jordan, 2003) models documents using a latent topic layer. In LDA, for each document d, a multinomial distribution θ_d over topics first is sampled from a Dirichlet distribution with parameter α. Second, for each word w_{di}, a topic z_{di} is chosen from this topic distribution. Finally, the word w_{di} is generated from a topic-specific multinomial distribution φ_{zdi}. Accordingly, the generating probability of word w from document d is:

$$P(w \mid d,\theta,\phi) = \sum_{z \in T} P(w \mid z,\phi_z)P(z \mid d,\theta_d)$$

Basically, we will use this approach as our topical analysis component to discover the un-derlying topic distribution for nouns, verbs, and adjectives.

3. Problem Definition

In this section, we formally define the problem of the social metaphor detection via topic diversity identification.

Social Metaphor detection: We aim to recognize non-conventionalized metaphors in so-cial media text by a fully automatic approach, where the input would be real text from social media. Based on the word distribution among the input data, we aim to detect metaphors without using any external knowledge resources.

There are many sub-categories of metaphors. In this work, we only focus on “non-conventionalized metaphors,” which reasonably could be considered as an outlier of language behavior. One advantage of non-conventionalized metaphors is that the approach can be language-independent and there is no need of external knowledge resources. This type of framework reasonably could be ported to other languages.

We will present how to tackle the problem by our proposed 3-step framework and discuss how to take the advantage of topical analysis for metaphor detection. We will also show how to quantitatively calculate these values in the next section.

4. Preliminary Test

As mentioned above, one of the most important approaches of metaphor detection is to identify the violation of selectional preference. Nevertheless, none of the other approaches are proposed as a baseline model to compare with the proposed model. In this section, to investigate the reliability of selectional preference modeling, we adopted another possible approach for metaphor detection, *i.e.*, the semantic outlier word detection, and run a preliminary test to compare their effectiveness.

4.1 Semantic Outlier Word Detection

Intuitively, for a certain topic, people tend to use the words that are “semantically related” to the topic. Therefore, we can assume that the set of words that are used frequently to describe a certain topic are more strongly related to each other than to the words used to describe other topics. For instance, the words used to describe “finance,” *e.g.*, bank, money, and business, are semantically more similar (or related) to each other than to the words used to describe “entertainment,” *e.g.*, movie, music, and star. Based on this idea, we can detect the “semantic outlier” in a chunk of text, which can indicate the words that are borrowed from other topics to establish metaphors.

In this paper, we basically followed the method proposed by (Inkpen & Désilets, 2005) to detect the semantic outlier words. For a chunk of an input sentence, we first use the DISCO[1] package to calculate the pair-wise semantic similarities between any two words within the text, before calculating the average of the three greatest similarities of each word as its “semantic coherence (SC).” Finally, the semantic outliers tend to have obviously lower semantic coherence than other words, so we just set an empirical threshold to capture those outliers.

4.2 Selectional Association Outlier Detection

Selectional preference (also referred to as selectional association or selectional restriction) describes the semantic preference of predicates to noun classes in a given grammatical rela-tion. For instance, the predicate “eat” prefers the noun class of “food” as its *direct object* more than the noun class “building” and also prefers the noun class of “human” and “animal” as its *subject* more than the noun class “vehicle”. Modeling selectional preference could help us to find the anomaly grammatical argument, which is an important clue to metaphorical language.

In this paper, for a given predicate *p* and a semantic noun class *c*, we adopt the measure of selectional association (SA), which was proposed by Resnik (1997), to present the selectional preference value between them. The selectional association equation can be calculated similar to point-wise mutual information, as follows:

$$A_R(p,c) = \frac{1}{S_R(p)} \Pr(c \mid p) \log \frac{\Pr(c \mid p)}{\Pr(c)}$$

A_R is the selectional association value between a given predicate *p* and a semantic noun class *c*. SR is the selectional preference strength of *p*, which can be formally defined similar to the K-L divergence between prior and posterior, as follows:

$$\begin{aligned} S_R(p,c) &= D(\Pr(c \mid p) \,\|\, \Pr(c)) \\ &= \sum_c \Pr(c \mid p) \log \frac{\Pr(c \mid p)}{\Pr(c)} \end{aligned}$$

Finally, similar to Section 4.1, the selectional preference outliers tend to have obviously lower SA value than others, so we set an empirical threshold to capture those outliers. Note

[1] http://www.linguatools.de/

that, for this preliminary test, we only focus on the direct-object (*dobj*) and subject (*subj*) grammatical relations.

4.3 Experiment and Discussion

Since labeling metaphor embedded sentences is laborious, we conduct experiments on a relatively small benchmark corpus, which contains 122 sentences extracted from the Web, where 61 (50%) of them contain metaphors and 61 of them do not contain metaphors.

We apply both approaches on this data set. For the selectional association outlier detection, the best resulting F-1 score is 0.58, with precision of 0.60 and recall of 0.56. On the other hand, for the semantic outlier word detection, regardless of which value of threshold we set, the performance remains very low. This method returns a huge number of false positive semantic outliers, which is mainly caused by two reasons.

First, the semantic coherence can be affected easily by very general words, which usually have very high similarities and occur very often. If one sentence has more than one very general context word, *e.g.*, "take," "put," or "get," the semantic coherences of all other words could be systematically increased, and thereby fail to present the outlier words. We believe this is the main reason this method cannot detect the semantic outliers we expected.

Second, the measure of semantic similarity between word pairs is not very reliable for in-frequent words. The similarities calculations that are based on the text of a large corpus usually have this problem – being reliable on high frequency words, but not on low frequency words, which are exactly what we aim to capture.

To conclude, the selectional association outlier detection method outperformed the semantic outlier word detection in the preliminary test. Therefore, in this paper, we only focus on selectional association to develop our technology.

5. 3-Step Framework of Metaphor Detection

In this section, we introduce our approach to the problem of social metaphor detection.

In particular, our approach consists of three steps: (1) word extraction and building noun clustering, (2) selectional association outlier detection, and (3) selectional preference strength filtering. The first step deals with noisy input social media data, and it produces relatively clean output with richer NLP information labeled on the text. In the second step, we use a statistical method to calculate the selectional association scores of particular types of token pairs, based on the tokens and noun clusters extracted from the first step. Finally, as a post-process step, the output generated from the first step will be further analyzed and false positives will be filtered out via an empirical threshold.

5.1 Step 1: Word Extraction and Noun Clustering

Different from well-phased corpora, *e.g.*, Wall Street Journal or Wikipedia pages, which are used by other metaphor detection works, social metaphors tend to be embedded in noisy social media texts, *e.g.*, blog and forum texts. The goal of word extraction is to filter out the noise from grammatically structured phrases and tokens.

We first use a POS tagger to label the tokens with part-of-speech tags. Nevertheless, since the POS taggers are unlikely to produce high quality results on noisy data, we only select nouns with word frequency greater than 5 and greater than 70% of the overall occurrences as a noun. For adjectives and verbs, more strictly, we require a word frequency greater than 50 and over 80% of all occurrences should be adjectives or verbs. All of these parameterized thresholds are set experimentally.

Then, based on the nouns we extracted, we build a set of semantic noun clusters, which is the foundation for modeling the selectional preference. In this work, we apply the spectral clustering algorithm as follows.

1. For each noun W_N, we use the DICSO toolkit, which uses Wikipedia as the knowledge source, to generate its top 100 semantically similar nouns. For the first similar word W_{S1}, the similarity weight $Sim(W_N, W_{S1})$ is set to 1/2; for the second word, $Sim(W_N, W_{S2})$ is 1/3; for the third word, $Sim(W_N, W_{S3})$ is 1/4, and so on.
2. For all nouns, the first step will generate an asymmetric graph of word similarity. Based on the graph, we run the spectral clustering algorithm on it and get the noun cluster.

Note that, although the DISCO toolkit calculates word similarity based on Wikipedia, which is a reliable corpus, we only focus on the nouns actually occurring in the input data set, *i.e.*, the social media data. Namely, if a certain noun appears in the extracted “top 100 semantically similar nouns” but never occurs in the input data, we just ignore it. Moreover, we ignore the similarity score produced by the toolkit and calculate the similarity based on the similarity ranking. This is because, for the top 100 similar words, we tend to trust the ranking more than the scores, which is a common engineering trick for a clustering problem.

5.2 Step 2: Selectional Association Outlier Detection

Based on the formula mentioned in Section 4.2 and the semantic noun clusters built in Step 1, we measure the selectional associations for the most frequent verbs we extracted, particularly on the three kinds of grammatical relations, namely, adjective modifier (*amod*), direct object (*dobj*), and subject (*subj*).

In this work, we intentionally include the adjective modifier (*amod*) relation. When speak-ing of the selectional preference, most previous works have focused only on verbal predicates. Nevertheless, in the grammatical relation of adjective modifier, the modifier can also be considered as a predicate and the words being modified are mostly also nouns. Therefore, we also apply our approach on the *amod* relation and see if the method effectively captures adjective metaphors as well.

We considered the relations with negative SA values as “SA outliers,” and we labeled the sentences containing “SA outliers” as metaphors.

5.3 Step 3: Selectional Preference Strength Filter

As mentioned in Section 4.3, the selectional preference strength of a predicate is defined as the K-L divergence between the prior and the posterior of noun clusters. For the predicates with strong preference, *e.g*., “filmmake,” it significantly affects the posterior probability distribution of noun clusters. In the case of the direct object of “film-make,” the probability of the “movie/film” noun class is increased considerably. On the other hand, some “light verbs,” *e.g*., “get,” “put,” or “take,” have quite weak preferences toward their direct object or subject.

The idea of selectional preference strength filtering was first proposed by Shutova *et al*. (2010) and suggests that the predicates with less strong selectional preference would rarely “violate” their own weak preference. Therefore, if we filter out the predicates with weak se-lectional preference, the false positives of metaphor detection will be reduced, and the preci-sion will increase significantly. In our framework, we apply this filtering method as the final step. Note that, due to the lack of a training and development data set, we just set the same threshold, which is 1.32, as suggested in Shutova *et al*. (2010).

6. Topic Model Analysis

We use LDA to model the topical distribution of words and documents of corpora, and we want to observe the changes of selectional preferences among various topics. The steps are as follows.

1. We train an LDA topic model with k various topics based on the whole input data set, *i.e*., social media corpus.
2. For each document d in the input data set, we assign d to its favorite topic. Namely, we partition the corpus into k document collections, based on topics.
3. Run the 3-step process mentioned in Section 5 on the whole data set and on the k dif-ferent document collections.
4. Compare the SA outlier detection results among the data with and without topic modeling.

The underlying hypothesis in this comparison is that the selectional preference would increase for certain predicates in certain topics; thus, the outlier of SA values would be further emphasized. In that case, the metaphor detection technique could be improved.

7. Experiment

7.1 Data and Setting

Our method requires a fully-parsed data set, so we decided to choose a relatively small size of social media data. We collected the whole text of posts from a large online breast cancer support community, Breastcancer.org, which also is used in Wen *et al*. (2013). We have collected all of the public posts, users, and their profiles on the discussion board platform from October 2001 to January 2011. During this period, there were a total of 90,242 unique users who posted 1,562,459 messages. We then parsed it by the Stanford Parser toolkit[2]. In our word extraction step, we extracted 55,511 distinct nouns, 3,242 distinct adjectives, and 1,827 distinct verbs.

In the noun clustering step, we experimentally set the number of clusters (k) as 2,000. Note that we also manually removed the following three clusters to avoid some systematic parsing errors of the Stanford parser:

- hours, minutes, times, days, weeks, months, seconds, …
- yourselves, oneself, somebody, everybody, someone, anything, everything, anyone, …
- boy, girl, child, woman, children, guy, kid, person, …

In the topic model analysis phase, we adopted the JGibbLDA[3] toolkit to build the model and set the number of topics (k) as 20.

7.2 Results and Case Study

For the whole data set, the top 10 sample detected selectional association outliers[4] (of the three grammatical relationships) are listed in Table 1. We also demonstrate the result of one

[2] http://nlp.stanford.edu/software/lex-parser.shtml

[3] A Java Implementation of Latent Dirichlet Allocation (LDA) using Gibbs Sampling for Parameter Estimation and Inference: http://jgibblda.sourceforge.net/

[4] For each pair of predicate and noun cluster, we try to select the most “metaphor-like” usage if multiple outliers are detected. To protect the privacy of forum users, we also skip all the examples which contain name entities.

out of 20 topic document collections in Table 2 for comparison. Note that example usages are lightly disguised based on the techniques suggested by Bruckman (2006).

We found out that the strength of selectional preference of each predicate was actually in-creased in split topics. Nevertheless, the increase had no clear benefits to metaphor detection in our results. It successfully detected “outliers,” but those outliers were not necessarily metaphors.

Take the results of direct object for example. Without topic analysis, the top outliers we detected were *(accomplish, Bianca)*, *(defy, breast)*, *(occupy, breast)*, and *(sprinkle, germ)*. Most of them are just rarely used verb-object combinations, but not metaphors. With topic analysis, we picked one topic out of 20 as an example, and the top outliers we detected were *(celebrate, cancer)*, *(join, skin)*, *(draw, brow)*, and *(play, head)*. We can observe that the verbs and nouns are actually more concentrated. In this case, the topic seems like celebration/play/event/play. Nevertheless, those pairs are rare, but not metaphors.

8. Discussion

Though the final result is not very promising, we gain some valuable experience in this work.

First, a parsing error is lethal for our approach. It would hurt our performance in at least two aspects: putting incorrect nouns in the noun cluster, which is the foundation of the whole method, and creating a significant amount of noise in the data, thereby impacting the statistical modeling phase. Therefore, the pre-processing is critical. After we added the strict word extraction strategy into our system, the quality of output was improved.

Second, from our experiments, we found that the strength of selectional preference is actually increased when clustering the documents by topic modeling. In each topic’s document collection, we collected documents by word co-occurrences. Therefore, predicates are more concentrated on their preferred grammatical arguments. Nevertheless, the enhancement of selectional preference strength turned out not strong enough to improve metaphor detection. For some certain topics, the top SA outliers were even worse than those of the whole set, because selectional association is a linguistic phenomenon with high data sparsity. Partitioning would further reduce the amount of data and affect the reliability of the model.

Finally, we noticed that our fundamental hypothesis might not be accurate. We found that the SA outliers are not necessarily metaphors. Some of the outliers just rarely-used language, or some “weird” usage, *e.g.*, *(hug, multiply)* in “*the hugs we are storing will multiply*” of Ta-ble 1, or the *(play, head)* in “*It keeps playing through my head now*” of Table 2. In the future, we might need to reconsider the hypothesis we adopted.

Table 1. Examples of Selectional Association Violation Identified without Topical Analysis

Relation (arg0, arg1)	**SA(10^{-3})**	**Example Usage**	**Analysis**
amod			
amod(breast, yearly)	-2.7306	“yearly breast MRI”	Parsing Error
amod(skin, circular)	-2.7079	“circular skin patches”	Non-metaphor
amod(skin, greasy)	-2.6896	“greasy skin”	Non-metaphor
amod(head, administrative)	-2.6864	“the administrative head of this institute”	Weak metaphor
amod(hug, weary)	-2.6461	“…get weary. Hugs to you all…”	Sentence Segmentation Error
amod(breast, uncertain)	-2.6138	“The breast dimpling and uncertain mammography…”	Parsing Error
amod(kiss, french)	-2.5970	“…about French kiss…”	Non-metaphor
amod(breast, slim)	-2.5752	“My breasts are not slim but not fat...”	Non-metaphor
amod(tomorrow, crisp)	-2.5636	“…it's expected to be a crisp 72 tomorrow.”	Parsing Error
amod(wing, seasoned)	-2.5510	“seasoned chicken wings”	Non-metaphor
dobj			
dobj(defy, breast)	-2.5893	“gravity defying breasts”	Parsing Error
dobj(occupy, breast)	-2.5749	“…(cancer) occupy the whole breast…”	Non-metaphor
dobj(sprinkle, germ)	-2.5350	“sprinkle wheat germ”	Non-metaphor
dobj(ooze, skin)	-2.5260	"oozing skin"	Non-metaphor
dobj(circulate, breast)	-2.5157	“…let air circulates around patient’s breast.”	Parsing Error
dobj(win, tomorrow)	-2.5095	“If John win tomorrow night, …”	Metonymy
dobj(hire, dvd)	-2.4972	“hire the dvd”	Non-metaphor
dobj(defy, cancer)	-2.4773	“…to defy the cancer and smile…”	Non-metaphor
dobj(float, cancer)	-2.4380	“…cancer cells float around in my blood…”	Non-metaphor
dobj(shut, head)	-2.4141	“…shut my head off…”	Metaphor
nsubj			
nsubj(cleanse, breast)	-2.5783	“breast cleanse”	Parsing Error
nsubj(metabolize, tumor)	-2.5513	“Tumors metabolize …”	Non-metaphor
nsubj(deny, adjuster)	-2.4950	“The claims adjuster denied this claim …”	Non-metaphor
nsubj(occupy, head)	-2.4827	“…keep my head occupied …”	Weak metaphor

nsubj(multiply, hug)	-2.4617	"…the hugs will multiply."	Metaphor
nsubj(constipate, hug)	-2.4286	"… hugs … that percocet is constipating."	Parsing Error
nsubj(overtake, belly)	-2.3276	"… my belly has overtaken the boobs …"	Metaphor
nsubj(multiply, treatment)	-2.2361	"…treatment for.. , multiply that by…"	Weak metaphor
nsubj(pay, patient)	-2.2164	"…patients pay for…"	Non-metaphor
nsubj(manufacture, expander)	-2.2056	"…ask the expander manufactures come up with better tissue expander."	Parsing Error

Table 2. Examples of Selectional Association Violation Identified Based on Topical Analysis (for one Particular Topic)

Relation (arg0, arg1)	SA(10^{-3})	Example Usage	Analysis
amod			
amod(head, gray)	-2.5469	"gray head"	Metonymy
amod(belly, former)	-2.5462	"your former belly"	Non-metaphor
amod(carcinoma, vaginal)	-2.5452	"… vaginal squamous cell carcinomas …"	Non-metaphor
amod(cancer, unilateral)	-2.5144	"unilateral breast cancer"	Non-metaphor
amod(breast, unilateral)	-2.4714	"unilateral breast"	Non-metaphor
amod(lesion, bilateral)	-2.3713	"bilateral lesions"	Non-metaphor
amod(treatment, immediate)	-2.3687	"immediate treatment"	Non-metaphor
amod(flyer, weekly)	-2.3064	"weekly flyer"	Non-metaphor
amod(symptom, bilateral)	-2.2976	"bilateral symptoms"	Non-metaphor
amod(tumor, enlarged)	-2.2626	"enlarged malignant tumor"	Non-metaphor
dobj			
dobj(celebrate, cancer)	-2.7801	"…celebrate my 10th cancer free year."	Parsing Error
dobj(weigh, head)	-2.7256	"So many questions … is weighing my head."	Metaphor
dobj(join, skin)	-2.7097	"…join the skin together…"	Non-metaphor
dobj(draw, nose)	-2.4197	"…drew a nose on it."	Non-metaphor
dobj(play, cheek)	-2.3255	"…play up my eyes…"	Non-metaphor
dobj(join, slew)	-2.1792	"Mary joined a slew of women …"	Non-metaphor
dobj(play, tomorrow)	-2.1190	"Playing golf tomorrow…"	Parsing Error

dobj(apply, forehead)	-2.0029	"…apply directly to the forehead."	Non-metaphor
dobj(pay, cancer)	-1.9471	"…price to pay for surviving cancer…"	Non-metaphor
dobj(regain, head)	-1.9457	"…regained a full head of hair…"	Parsing Error
nsubj			
nsubj(specialize, patient)	-2.3001	"…specializes in working with breast cancer patients, …"	Parsing Error
nsubj(pay, treatment)	-2.2237	"…get the treatment and self pay, …"	Parsing Error
nsubj(cover, cheek)	-2.0421	"…my cheeks covered with…"	Non-metaphor
nsubj(pay, head)	-1.8908	"…you're drinking safe and only your head is paying the price."	(Weak) metaphor
nsubj(pay, homeschooling)	-1.7228	"…the homeschooling paid off."	Non-metaphor
nsubj(build, expander)	-1.3925	"... an expander to build ..."	Parsing Error
nsubj(cover, melatonin)	-1.3865	"…melatonin covers the need for…"	Non-metaphor
nsubj(cover, wife)	-1.2500	"…so his wife should be covered…"	Non-metaphor
nsubj(cover, nurse)	-1.1849	"…the nurses talking about the insurance would cover it."	Parsing Error
nsubj(cover, dose)	-1.1708	"…do the single big dose to cover 2 weeks…"	Non-metaphor

9. Conclusion and Future Work

In this paper, we tried to leverage one of the most well-known approaches in detecting the violation of selectional preference with topical analysis techniques. The idea of selectional preference is that verbs tend to have semantic preferences of their arguments, while topical information provides additional evidence to facilitate identification of selectional preferences among text. Although our experimental results show that topics do not have strong impact on the metaphor detection techniques, we analyzed the results and presented some insights from our study.

As our next step, to reconsider our hypothesis, we need to quantitatively compare our re-sults to the gold-standard benchmark. Another interesting experiment might be to cluster the predicates, similar to nouns, as in our experiments, because the predicates still suffer from the sparsity issue.

Acknowledgments

Supported by the Intelligence Advanced Research Projects Activity (IARPA) via Department of Defense US Army Research Laboratory contract number W911NF-12-C-0020. The U.S. Government is authorized to reproduce and distribute reprints for Governmental purposes notwithstanding any copyright annotation thereon. Disclaimer: The views and conclusions contained herein are those of the authors and should not be interpreted as necessarily repre-senting the official policies or endorsements, either expressed or implied, of IARPA, DoD/ARL, or the U.S. Government.

We would also like to thank Zi Yang for his help of the topical analysis experiments, Teruko Mitamura and Eric Nyberg for their instructions, and Yi-Chia Wang and Dong Ngu-yen for the work of data collection.

References

Birke, J., & Sarkar, A. (2006). A clustering approach for the nearly unsupervised recognition of nonliteral language. In *Proceedings of EACL*, *6*, 329-336.

Blei, D., Ng, A., & Jordan, M. (2003). Latent dirichlet allocation. *Journal of machine Learning research*, *3*, 993-1022.

Bruckman, A. (2006). Teaching students to study online communities ethically. *Journal of Information Ethics*, *15*(2), 82-98.

Calzolari, N., Choukri, K., Maegaard, B., Mariani, J., Odijk, J., Piperidis, S., Rosner, M., & Tapias, D. eds. (2010). In *Proceedings of the International Conference on Language Resources and Evaluation, LREC 2010*, 17-23 May 2010, Valletta, Malta. European Language Resources Association.

Fass, D. (1991). met*: A method for discriminating metonymy and metaphor by computer. *Computational Linguistics*, *17*(1), 49-90.

Goatly, A. (1997). *The language of metaphors*. volume 3. Routledge London.

Inkpen, D., & Désilets, A. (2005). Semantic similarity for detecting recognition errors in automatic speech transcripts. In *Proceedings of Conference on Empirical Methods in Natural Language Processing (EMNLP 2005)*. October 6-8, 2005. Vancouver, British Columbia, Canada. NRC 48278.

Loenneker-Rodman, B., & Narayanan, S. (2010). Computational approaches to figurative language. *Cambridge Encyclopedia of Psycholinguistics*.

Mason, Z. (2004). Cormet: a computational, corpus-based conventional metaphor extraction system. *Computational Linguistics, 30*(1), 23-44.

Ng, A., Jordan, M., Weiss, Y., *et al*. (2002). On spectral clustering: Analysis and an algorithm. *Advances in neural information processing systems*, *2*(8), 49-856.

Peters, W., & Peters, I. (2000). Lexicalised systematic polysemy in wordnet. In *Proc. Secondt Intnl Conf on Language Resources and Evaluation*.

Preiss, J., Briscoe, T., & Korhonen, A. (2007). A system for large-scale acquisition of verbal, nominal and adjectival subcategorization frames from corpora. In *ANNUAL MEETING-ASSOCIATION FOR COMPUTATIONAL LINGUISTICS*, 45, 912.

Resnik, P. (1997). Selectional preference and sense disambiguation. In P*roceedings of the ACL SIGLEX Workshop on Tagging Text with Lexical Semantics: Why, What, and How*, 52-57. Washington, DC.

Shutova, E., & Teufel, S. (2010). Metaphor corpus annotated for source-target domain mappings. In *Proceedings of LREC*.

Shutova, E., Sun, L., & Korhonen, A. (2010). Metaphor identification using verb and noun clustering. In *Proceedings of the 23rd International Conference on Computational Linguistics*, 1002-1010. Association for Computational Linguistics.

Shutova, E. (2010). Models of metaphor in nlp. In *Proceedings of the 48th Annual Meeting of the Association for Computational Linguistics*, 688-697. Association for Computational Linguistics.

Wen, M., Zheng, Z., Jang, H., Xiang, G., & Rose, C. (2013). Extracting Events with Informal Temporal References in Personal Histories in Online Communities. In *ACL'13*.