Social Metaphor Detection via Topical Analysis

Ting-Hao (Kenneth) Huang
Language Technologies Institute, Carnegie Mellon University
5000 Forbes Ave., Pittsburgh, PA 15213, USA
windx@cmu.edu

Abstract

With massive social media data, e.g., comments, blog articles, or tweets, become available, there is a rising interest towards automatic metaphor detection from open social text. One of the most well-known approaches is detecting the violation of selectional preference. The idea of selectional preference is that verbs tend to have semantic preferences of their arguments. If we find that in some text, any arguments of these predicates are not of their preferred semantic classes, and it's very likely to be a metaphor. However, previously only few papers have focuses on leveraging topical analysis techniques in metaphor detection. Intuitively, both predicates 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. In this paper, we study how the metaphor detection technique can be influenced by topical analysis techniques based on our proposed three-step approach. We formally define the problem, and propose our approach for metaphor detection, and then we conduct experiments on a real-world data set. Though our experimental result shows that topics do not have strong impacts on the metaphor detection techniques, we analyze the result and present some insights based on our study.

1 Introduction

With massive social media data, e.g., comments, blog articles, or tweets, become available, there is a rising interest towards automatic metaphor detection from open social text. One of the most well-known approaches is detecting the violation of selectional preference. The idea of selectional preference is that the predicates (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's very likely to be a metaphorical use.

Previously, researchers have studies metaphor identification by modeling selectional preference (Loenneker-Rodman and Narayanan, 2010; Shutova *et al.*, 2010; Shutova, 2010; Resnik, 1997; Shutova and Teufel, 2010; Calzolari *et al.*, 2010; Preiss *et al.*, 2007), while only few papers have focused on leveraging topical analysis techniques in metaphor detection. 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 sport, 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 aim to study how the metaphor detection technique can be influenced based on topical analysis techniques.

The problem of automatic social metaphor detection via topical analysis poses several 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 carefully studied. We should empirically estimate the performance of existing NLP tools, especially lemmatizers and POS taggers.

Second, how we can automatically discover the topical distribution for each term (including verbs and nouns) within open text is not a trivial problem. Moreover, we also need to study how to leverage topical distribution of each verb and noun to metaphor detection.

Finally, how to apply and evaluate the proposed approach on a real world data set is not straight-forward, as there is hardly existing data set nor benchmark to evaluate metaphor detection, we need to create a benchmark that can effectively show that the performance difference.

In this paper, we formally define the problem, and propose our 3-step approach for metaphor

IJCNLP 2013 Workshop on Natural Language Processing for Social Media (SocialNLP), pages 14–22,
Nagoya, Japan, 14 October 2013.

detection, specifically, we first preprocess the input text by extracting tokens and further clustering nouns, and then we detect selectional association outlier, 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 every single topic data partitions, respectively. Finally, we compare the metaphor detection results between that with and without influences of topics, and to observe which one performs better.

The rest of the paper is organized as follows: In Section 2, we briefly summarize related work for metaphor detection based on selectional preference detection. In Section 3, we formally define the problem of automatic social metaphor detection. Then, in Section 4, we first conduct a preliminary test to compare two technologies for metaphor detection, and choose one to establish the 3-step framework 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 whole work in Section 9.

2 Related Work

In this section, we briefly survey papers that investigate approaches to detect metaphor 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 metaphors could be divided into two sub-fields: metaphor detection and metaphor interpretation. In this paper, we focus on metaphor detection and aim to explore some new potential directions of this field.

Speaking of metaphor detection, 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 that we only aim to detect "non-conventionalized metaphor" which usually has low frequency and could reasonably be considered as outliers.

The Met* System (Fass, 1991) can be considered as the first attempt to explore this field, and the following approaches include (Goatly, 1997), (Peters and Peters, 2000), CorMet System (Mason, 2004), and TroFi System (Birke and Sarkar, 2006). Most of them adopt the concept of selectional preference which we mentioned above, along with some hand-coded knowledge base, e.g., VerbNet. VerbNet has the information about the constraint of arguments of verbs. By matching the text with verb and its argument, we're able to detect the violation of arguments. However, in this paper, we apply a different approach that learns the violations purely from statistics based on natural texts. One advantage of this method is that we don't need any hand-coded knowledge, so could be easier to be ported to other languages.

2.2 Topical analysis

Many topical analysis techniques have been developed, e.g., latent semantic analysis, probabilistic LSA, NMF, LDA, etc.

Latent Dirichlet Allocation (LDA) (Blei, Ng, and Jordan, 2003) models documents using a latent topic layer. In LDA, for each document d, a multinomial distribution θ_d over topics is first 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 $\varphi_{z_{di}}$. Accordingly, the generating probability of word w from document d is:

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

Basically, we will use this approach as our topical analysis component to discover underlying 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 the non-conventionalized metaphors in social 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 could be reasonably considered as an outlier of language behavior. One advantage of non-conventionalized metaphors is that the approach can be language independent and need no external knowledge resource. This kind of framework could be simply ported to various 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 next section.

4 Preliminary Test

As mentioned above, one of the most important approaches of metaphor detection is to detect the violation of selectional preference. However, none of other approaches are proposed as a baseline model to compare with it. 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 ran a preliminary test to compare their performance.

4.1 Semantic Outlier Word Detection

Intuitively, for a certain topic, people tend to use the words that are "semantically more related" to the topic. Therefore, we can estimate that the set of words which are usually used 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, business, are semantically more similar (or related) to each other than to the words used to describe "entertainment", e.g., movie, music, star, etc. 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 and Désilets, 2005) to detect the semantic outlier words. For one input sentence, we first use the DISCO[1] package to calculate the pair-wise semantic similarities between any two words within the sentence, and then calculate the average of 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.

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

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 relation. 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 of metaphorical languages.

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

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

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

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

Finally, similar as the Section 4.1, the selectional preference outliers tend to have obviously lower SA value than others, so we just set an empirical threshold to capture those outliers. Note 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 are effort consuming, we conduct experiment on a benchmark corpus, which contains 122 sentences extracted from the Web, where 61 (50%) of them contain metaphor, and 61 of them don't contain metaphor.

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 huge amount of false positive semantic outliers. Mainly due to two reasons: First, the semantic coherence can be easily affected by very general words, which usually have very high similarities and also occur very often. If one sentence has more than one very general context words, e.g., "take", "put", or "get", the semantic coherences of all other words could be systematically increased, and thus fail to represent the outlier words. We believe that's the main reason why this method can not detect the semantic outlier we expected. Second, the measure of semantic similarity between word pairs is not very reliable for low frequency words. The similarities calculations which are based on the text of big corpus usually have this problem: It's reliable on high frequency words, but not on low frequency words, which exactly are what we aim to capture.

To conclude, the selectional association outlier detection method obviously outperform 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, (3) selectional preference strength filter. The first step deals with noisy input social media data, and produce relatively clean output with richer NLP information labeled on the text, and in the second step, we use 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 filtered out false positives based on empirical threshold.

5.1 Step 1: Word extraction and noun clustering

Different from well phased corpora, e.g., Wall Street Journal, or Wikipedia pages, that are used by other metaphor detection methods, social metaphors tend to be embedded in noisy social media text, 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. However, since the POS taggers unlikely 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 the word frequency greater than 50, and over 80% of all occurrences should be adjectives or verbs.

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 spectral clustering algorithm. Specifically,

1. For each noun, we use the DICSO toolkit to extract their top 100 semantically similar nouns (from Wikipedia). For the first similar words, the similarity weight is set to 1/2; the second is 1/3, the third is 1/4, and so on.
2. We use this information as feature, and run the spectral clustering algorithm among all nouns we extracted.

Note that though the DISCO toolkit calculates word similarity based on Wikipedia, which is a reliable corpus, we only focus on the nouns actually occur 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.

5.2 Step 2: Selectional association outlier detection

Based on the formula mentioned in the Section 4.2 and the semantic noun clusters built in Step 1, we measure the selctional 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 relation. When speaking of the selectional preference, most previous work focus only on verbal predicates. However, in the grammatical relation of adjective modifier, the modifier can also be considered as a predicate, and the modifyees are mostly also nouns. Therefore, we also aim to apply our approach on the *amod* relation, and see if the method also effectively captures adjective metaphors.

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

5.3 Step 3: Selectional preference strength filter

As mentioned in the Section 4.3, 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 affect the posterior probability distribution of noun clusters. In the case of the direct object of “filmmake”, the probability of “movie/film” noun class is hugely increased. On the other hand, some light verb, e.g., “get”, “put”, “take”, have quite weak preferences toward their direct object or subject.

The idea of selectional preference filter is first proposed by (Shutova, *et al.*, 2010), which suggests that the predicates with less strong selectional preference would barely “violates” their own weak preference. Therefore, if we filter out the predicates with weak selectional preference, the false positives of metaphor detection will be reduced, and the precision will significantly increase. In our framework, we apply this filter as the final step. Note that due to the lack of training and developing data, we just set the same threshold, which is 1.32, as suggested as that 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 the Section 5 on the whole data set, and also on the *k* different document collections, respectively.
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, and 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 the fully-parsed data set, so we decide to choose a reasonable size of social media data. We collected the whole text of posts from a large online breast cancer support community which is also used in (Wen, *et al.*, 2013), and then parse it by the Stanford Parser toolkit[2]. In our word extraction step, we extract 55,511 distinct nouns, 3,242 distinct adjectives, and 1,827 distinct verbs.

Note that in the noun clustering step, we manually removed the following 3 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 adopt 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 the Table 1. We also demonstrate the result of one out of twenty topic document collections in the Table 2 for comparison. Note that example usages are lightly disguised based on the techniques suggested by (Bruckman, 2006).

[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.

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 1: Examples of Selectional Association Violation Identified without Topical Analysis

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

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

We found that the strength of selectional preference of each predicates are actually increased in split topics. However, the increase has no clear benefits to metaphor detection in our result. It successfully detects "outliers", but those outliers are not necessarily to be metaphors.

Take the result of direct object for example. Without topic analysis, the top outliers we detected are (accomplish, Bianca), (defy, breast), (occupy, breast), (sprinkle, germ). Most of them are just rarely used verb-object combinations, but not metaphors. With topic analysis, we picked one topic out of twenty as example, the top outliers we detected are (celebrate, cancer), (join, skin), (draw, brow), (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. However, those pairs are also only rare, but not metaphors.

8 Discussion

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

Firstly, parsing error is lethal for our approach. It would hurt our performance in at least two aspects. Parsing errors would put incorrect nouns in the noun cluster, which is the foundation of the whole method. Furthermore, it would also create significant amount of noise in the data, and thus affect the statistical modeling phase. Therefore, the pre-processing is critical. After we added the strict word extraction strategy into our system, the quality of outputs is significantly improved.

Secondly, 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 collect documents by word co-occurrences. Therefore, predicates are more concentrated on their preferred grammatical arguments. However, the enhancement of selectional preference strength turned out not strong enough to improve metaphor detection. For some certain topic, the top SA outliers are even worse than that 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 also 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 languages, or some "weird" usage, e.g., (hug, multiply) in "t*he hugs we are storing will multiply*" of the Table 1, or the (play, head) in "*It keeps playing through my head now*." of the Table 2. We might need to reconsider about the hypothesis we adopt in the future.

9 Conclusion and Future Work

In this paper, we try 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 preference among text. Though our experimental result shows that topics do not have strong impacts on the metaphor detection techniques, we analyze the result and present some insights based on our study.

As our next step, for reconsidering our hypothesis, we need to quantitatively compare our results on the gold-standard benchmark. Another interesting experiment might be to cluster the predicates, similar to nouns, as in our experiments, the predicates still suffer from 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 representing 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 Nguyen for the work of data collection.

References

Birke, J., and Sarkar, A. 2006. A clustering approach for the nearly unsupervised recognition of nonliteral language. In Proceedings of EACL, volume 6, 329–336.

Blei, D.; Ng, A.; and Jordan, M. 2003. Latent dirichlet allocation. the Journal of machine Learning research 3:993–1022.

Bruckman, Amy. (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.; and Tapias, D., eds. 2010. 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., and Désilets, A. 2005. Semantic similarity for detecting recognition errors in automatic speech transcripts.

Loenneker-Rodman, B., and 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 2:849–856.

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

Preiss, J.; Briscoe, T.; and 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, volume 45, 912.

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

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

Shutova, E.; Sun, L.; and 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., and Rose, C. (2013). Extracting Events with Informal Temporal References in Personal Histories in Online Communities. ACL'13.