Extracting Signed Social Networks From Text

Ahmed Hassan
Microsoft Research
Redmond, WA, USA
hassanam@microsoft.com

Amjad Abu-Jbara
EECS Department
University of Michigan
Ann Arbor, MI, USA
amjbara@umich.edu

Dragomir Radev
EECS Department
University of Michigan
Ann Arbor, MI, USA
radev@umich.edu

Abstract

Most of the research on social networks has almost exclusively focused on positive links between entities. There are much more insights that we may gain by generalizing social networks to the signed case where both positive and negative edges are considered. One of the reasons why signed social networks have received less attention that networks based on positive links only is the lack of an explicit notion of negative relations in most social network applications. However, most such applications have text embedded in the social network. Applying linguistic analysis techniques to this text enables us to identify both positive and negative interactions. In this work, we propose a new method to automatically construct a signed social network from text. The resulting networks have a polarity associated with every edge. Edge polarity is a means for indicating a positive or negative affinity between two individuals. We apply the proposed method to a larger amount of online discussion posts. Experiments show that the proposed method is capable of constructing networks from text with high accuracy. We also connect out analysis to social psychology theories of signed network, namely the structural balance theory.

1 Introduction

A great body of research work has focused on social network analysis. Social network analysis plays a huge role in understanding and improving social computing applications. Most of this research has almost exclusively focused on positive links between individuals (e.g. friends, fans, followers, etc.). However, if we carefully examine the relationships between individuals in online communities, we will find out that limiting links to positive interactions is a very simplistic assumption. It is true that people show positive attitude by labeling others as friends, and showing agreement, but they also show disagreement, and antagonism toward other members of the online community. Discussion forums are one example that makes it clear that considering both positive and negative interactions is essential for understanding the rich relationships that develop between individuals in online communities.

If considering both negative and positive interactions will provide much more insight toward understanding the social network, why did most of previous work only focus on positive interactions? We think that one of the main reasons behind this is the lack of a notion for explicitly labeling negative relations. For example, most social web applications allow people to mark others as friends, like them, follow them, etc. However, they do not allow people to explicitly label negative relations with others.

Previous work has built networks from discussions by linking people who reply to one another. Even though, the mere fact that X replied to Y's post does show an interaction, it does not tell us anything about the type of that interaction. In this case, the type of interaction is not readily available; however it may be mined from the text that underlies the social network. Hence, if we examine the text exchanged between individuals, we may be able to come up with conclusions about, not only the existence of an interaction, but also its type.

In this work, we apply Natural Language Processing techniques to text correspondences exchanged between individuals to identify the under-

Proceedings of the TextGraphs-7 Workshop at ACL, pages 6–14,
Jeju, Republic of Korea, 12 July 2012.

lying signed social structure in online communities. We present and compare several algorithms for identifying user attitude and for automatically constructing a signed social network representation. We apply the proposed methods to a large set of discussion posts. We evaluate the performance using a manually labeled dataset.

The input to our algorithm is a set of text correspondences exchanged between users (e.g. posts or comments). The output is a *signed network* where edges signify the existence of an interaction between two users. The resulting network has polarity associated with every edge. Edge polarity is a means for indicating a positive or negative affinity between two individuals.

The proposed method was applied to a very large dataset of online discussions. To evaluate our automated procedure, we asked human annotators to examine text correspondences exchanged between individuals and judge whether their interaction is positive or negative. We compared the edge signs that had been automatically identified to edges manually created by human annotators.

We also connected our analysis to social psychology theories, namely the Structural Balance Theory (Heider, 1946). The balance theory has been shown to hold both theoretically (Heider, 1946) and empirically (Leskovec et al., 2010b) for a variety of social community settings. Showing that it also holds for our *automatically* constructed network further validates our results.

The rest of the paper is structured as follows. In section 2, we review some of the related prior work on mining sentiment from text, mining online discussions, extracting social networks from text, and analyzing signed social networks. We define our problem and explain our approach in Section 3. Section 4 describes our dataset. Results and discussion are presented in Section 5. We present a possible application for the proposed approach in Section 6. We conclude in Section 7.

2 Related Work

In this section, we survey several lines of research that are related to our work.

2.1 Mining Sentiment from Text

Our general goal of mining attitude from one individual toward another makes our work related to a huge body of work on sentiment analysis. One such line of research is the well-studied problem of identifying the of individual words. In previous work, Hatzivassiloglou and McKeown (1997) proposed a method to identify the polarity of adjectives based on conjunctions linking them in a large corpus. Turney and Littman (2003) used statistical measures to find the association between a given word and a set of positive/negative seed words. Takamura et al. (2005) used the spin model to extract word semantic orientation. Finally, Hassan and Radev (2010) use a random walk model defined over a word relatedness graph to classify words as either positive or negative.

Subjectivity analysis is yet another research line that is closely related to our general goal of mining attitude. The objective of subjectivity analysis is to identify text that presents opinion as opposed to objective text that presents factual information (Wiebe, 2000). Prior work on subjectivity analysis mainly consists of two main categories: subjectivity of a phrase or word is analyzed regardless of the context (Wiebe, 2000; Hatzivassiloglou and Wiebe, 2000; Banea et al., 2008), or within its context (Riloff and Wiebe, 2003; Yu and Hatzivassiloglou, 2003; Nasukawa and Yi, 2003; Popescu and Etzioni, 2005). Hassan et al. (2010) presents a method for identifying sentences that display an attitude from the text writer toward the text recipient. Our work is different from subjectivity analysis because we are not only interested in discriminating between opinions and facts. Rather, we are interested in identifying the polarity of interactions between individuals. Our method is not restricted to phrases or words, rather it generalizes this to identifying the polarity of an interaction between two individuals based on several posts they exchange.

2.2 Mining Online Discussions

Our use of discussion threads as a source of data connects us to some previous work on mining online discussions. Lin et al. (2009) proposed a sparse coding-based model that simultaneously models semantics and structure of threaded discus-

sions. Huang et al. (2007) learn SVM classifiers from data to extract (thread-title, reply) pairs. Their objective was to build a chatbot for a certain domain using knowledge from online discussion forums. Shen et al. (2006) proposed three clustering methods for exploiting the temporal information in discussion streams, as well as an algorithm based on linguistic features to analyze discourse structure information.

2.3 Extracting Social Networks from Text

Little work has been done on the front of extracting social relations between individuals from text. Elson et al. (2010) present a method for extracting social networks from nineteenth-century British novels and serials. They link two characters based on whether they are in conversation or not. McCallum et al. (2007) explored the use of structured data such as email headers for social network construction. Gruzd and Hyrthonthwaite (2008) explored the use of post text in discussions to study interaction patterns in e-learning communities.

Our work is related to this line of research because we employ natural language processing techniques to reveal embedded social structures. Despite similarities, our work is uniquely characterized by the fact that we extract signed social networks from text.

2.4 Signed Social Networks

Most of the work on social networks analysis has only focused on positive interactions. A few recent papers have taken the signs of edges into account.

Brzozowski et al. (2008) study the positive and negative relationships between users of Essembly. Essembly is an ideological social network that distinguishes between ideological allies and nemeses. Kunegis et al. (2009) analyze user relationships in the Slashdot technology news site. Slashdot allows users of the website to tag other users as friends or foes, providing positive and negative endorsements. Leskovec et al. (2010c) study signed social networks generated from Slashdot, Epinions, and Wikipedia. They also connect their analysis to theories of signed networks from social psychology. A similar study used the same datasets for predicting positive and negative links given their context (Leskovec et al., 2010a). Other work addressed the problem of clustering signed networks by taking both positive and negative edges into consideration (Yang et al., 2007; Doreian and Mrvar, 2009).

All this work has been limited to analyzing a handful of datasets for which an explicit notion of both positive and negative relations exists. Our work goes beyond this limitation by leveraging the power of natural language processing to automate the discovery of signed social networks using the text embedded in the network.

3 Approach

The general goal of this work is to mine attitude between individuals engaged in an online discussion. We use that to extract a signed social network representing the interactions between different participants. Our approach consists of several steps. In this section, we will explain how we identify sentiment at the word level (i.e. polarity), at the sentence level (i.e. attitude), and finally generalize over this to find positive/negative interactions between individuals based on their text correspondences.

The first step toward identifying attitude is to identify polarized words. Polarized words are very good indicators of subjective sentences and hence we their existence will be highly correlated with the existence of attitude. The method we use for identifying word polarity is a Random Walk based method over a word relatedness graph (Hassan and Radev, 2010).

The following step is to move to the sentence level by examining different sentences to find out which sentences display an attitude from the text writer to the recipient. We train a classifier based on several sources of information to make this prediction (Hassan et al., 2010). We use lexical items, polarity tags, part-of-speech tags, and dependency parse trees to train a classifier that identifies sentences with attitude.

Finally, we build a network connecting participants based on their interactions. We use the predictions we made both at the word and sentence levels to associate a sign to every edge.

3.1 Identified Positive/Negative Words

The first step toward identifying attitude is to identify words with positive/negative semantic orientation. The semantic orientation or polarity of a word

indicates the direction the word deviates from the norm (Lehrer, 1974). Past work has demonstrated that polarized words are very good indicators of subjective sentences (Hatzivassiloglou and Wiebe, 2000; Wiebe et al., 2001). We use a Random Walk based method to identify the semantic orientation of words (Hassan and Radev, 2010). We construct a graph where each node represents a word/part-of-speech pair. We connect nodes based on synonyms, hypernyms, and similar-to relations from WordNet (Miller, 1995). For words that do not appear in WordNet, we use distributional similarity (Lee, 1999) as a proxy for word relatedness.

We use a list of words with known polarity (Stone et al., 1966) to label some of the nodes in the graph. We then define a random walk model where the set of nodes correspond to the state space, and transition probabilities are estimated by normalizing edge weights. We assume that a random surfer walks along the word relatedness graph starting from a word with unknown polarity. The walk continues until the surfer hits a word with a known polarity. Seed words with known polarity act as an absorbing boundary for the random walk. We calculate the mean hitting time (Norris, 1997) from any word with unknown polarity to the set of positive seeds and the set of negative seeds. If the absolute difference of the two mean hitting times is below a certain threshold, the word is classified as neutral. Otherwise, it is labeled with the class that has the smallest mean hitting time.

3.2 Identifying Attitude from Text

The first step toward identifying attitude is to identify words with positive/negative semantic orientation. The semantic orientation or polarity of a word indicates the direction the word deviates from the norm (Lehrer, 1974). We use OpinionFinder (Wilson et al., 2005a) to identify words with positive or negative semantic orientation. The polarity of a word is also affected by the context where the word appears. For example, a positive word that appears in a negated context should have a negative polarity. Other polarized words sometimes appear as neutral words in some contexts. Hence, we use the method described in (Wilson et al., 2005b) to identify the contextual polarity of words given their isolated polarity. A large set of features is used for that purpose including words, sentences, structure, and other features.

Our overall objective is to find the direct attitude between participants. Hence after identifying the semantic orientation of individual words, we move on to predicting which polarized expressions target the addressee and which are not.

Sentences that show an attitude are different from subjective sentences. Subjective sentences are sentences used to express opinions, evaluations, and speculations (Riloff and Wiebe, 2003). While every sentence that shows an attitude is a subjective sentence, not every subjective sentence shows an attitude toward the recipient. A discussion sentence may display an opinion about any topic yet no attitude.

We address the problem of identifying sentences with attitude as a relation detection problem in a supervised learning setting (Hassan et al., 2010). We study sentences that use second person pronouns and polarized expressions. We predict whether the second person pronoun is related to the polarized expression or not. We regard the second person pronoun and the polarized expression as two entities and try to learn a classifier that predicts whether the two entities are related or not. The text connecting the two entities offers a very condensed representation of the information needed to assess whether they are related or not. For example the two sentences "you are completely unqualified" and "you know what, he is unqualified ..." show two different ways the words "you", and "unqualified" could appear in a sentence. In the first case the polarized word unqualified refers to the word you. In the second case, the two words are not related. The sequence of words connecting the two entities is a very good predictor for whether they are related or not. However, these paths are completely lexicalized and consequently their performance will be limited by data sparseness. To alleviate this problem, we use higher levels of generalization to represent the path connecting the two tokens. These representations are the part-of-speech tags, and the shortest path in a dependency graph connecting the two tokens. We represent every sentence with several representations at different levels of generalization. For example, the sentence your ideas are very inspiring will be represented using lexical, polarity, part-of-

speech, and dependency information as follows:

LEX: "*YOUR ideas are very POS*"
POS: "*YOUR NNS VBP RB JJ_POS*"
DEP: "*YOUR poss nsubj POS*"

3.2.1 A Text Classification Approach

In this method, we treat the problem as a topic classification problem with two topics: having positive attitude and having negative attitude. As we are only interested in attitude between participants rather than sentiment in general, we restrict the text we analyze to sentences that contain mentions of the addressee (e.g. name or second person pronouns). A similar approach for sentiment classification has been presented in (Pang et al.,).

We represent text using the popular bag-of-words approach. Every piece of text is represented using a high dimensional feature space. Every word is considered a feature. The *tf-idf* weighting schema is used to calculate feature weights. *tf*, or term frequency, is the number of time a term t occurred in a document d. *idf*, or inverse document frequency, is a measure of the general importance of the term. It is obtained by dividing the total number of documents by the number of documents containing the term. The logarithm of this value is often used instead of the original value.

We used Support Vector Machines (SVMs) for classification. SVM has been shown to be highly effective for traditional text classification. We used the SVM Light implementation with default parameters (Joachims, 1999). All stop words were removed and all documents were length normalized before training.

The set of features we use are the set of unigrams, and bigrams representing the words, part-of-speech tags, and dependency relations connecting the two entities. For example the following features will be set for the previous example:

YOUR_ideas, YOUR_NNS, YOUR_poss, poss_nsubj, ..., etc.

We use Support Vector Machines (SVM) as a learning system because it is good with handling high dimensional feature spaces.

3.3 Extracting the Signed Network

In this subsection, we describe the procedure we used to build the signed network given the components we described in the previous subsections. This procedure consists of two main steps. The first is building the network without signs, and the second is assigning signs to different edges.

To build the network, we parse our data to identify different threads, posts and senders. Every sender is represented with a node in the network. An edge connects two nodes if there exists an interaction between the corresponding participants. We add a directed edge $A \rightarrow B$, if A replies to B's posts at least n times in m different threads. We set m, and n to 2 in most of our experiments. The interaction information (i.e. who replies to whom) can be extracted directly from the thread structure.

Once we build the network, we move to the more challenging task in which we associate a sign with every edge. We have shown in the previous section how sentences with positive and negative attitude can be extracted from text. Unfortunately the sign of an interaction cannot be trivially inferred from the polarity of sentences. For example, a single negative sentence written by A and directed to B does not mean that the interaction between A and B is negative. One way to solve this problem would be to compare the number of negative sentences to positive sentences in all posts between A and B and classify the interaction according to the plurality value. We will show later, in our experiment section, that such a simplistic method does not perform well in predicting the sign of an interaction.

As a result, we decided to pose the problem as a classical supervised learning problem. We came up with a set of features that we think are good predictors of the interaction sign, and we train a classifier using those features on a labeled dataset. Our features include numbers and percentages of positive/negative sentences per post, posts per thread, and so on. A sentence is labeled as positive/negative if a relation has been detected in this sentence between a second person pronoun and a positive/negative expression. A post is considered positive/negative based on the majority of relations detected in it. We use two sets of features. The first set is related to A only or B only. The second set

Participant Features
Number of posts per month for A (B)
Percentage of positive posts per month for A (B)
Percentage of negative posts per month for A (B)
gender
Interaction Features
Percentage/number of positive (negative) sentences per post
Percentage/number of positive (negative) posts per thread
Discussion Topic

Table 1: Features used by the Interaction Sign Classifier.

Class	Pos.	Neg.	Weigh. Avg.
TP Rate	0.847	0.809	0.835
FP Rate	0.191	0.153	0.179
Precision	0.906	0.71	0.844
Recall	0.847	0.809	0.835
F-Measure	0.875	0.756	0.838
Accuracy	-	-	**0.835**

Table 2: Interaction sign classifier evaluation.

is related to the interactions between A and B. The features are outlined in Table 1.

4 Data

Our data consists of a large amount of discussion threads collected from online discussion forums. We collected around $41,000$ threads and 1.2M posts from the period between the end of 2008 and the end of 2010. All threads were in English and had 5 posts or more. They covered a wide range of topics including: politics, religion, science, etc. The data was tokenized, sentence-split, and part-of-speech tagged with the OpenNLP toolkit. It was parsed with the Stanford parser (Klein and Manning, 2003).

We randomly selected 5300 posts (having approximately 1000 interactions), and asked human annotators to label them. Our annotators were instructed to read all the posts exchanged between two participants and decide whether the interaction between them is positive or negative. We used Amazon Mechanical Turk for annotations. Following previous work (Callison-Burch, 2009; Akkaya et al., 2010), we took several precautions to maintain data integrity. We restricted annotators to those based in the US to maintain an acceptable level of English fluency. We also restricted annotators to those who have more than 95% approval rate for all previous work. Moreover, we asked three different annotators to label every interaction. The label was computed by taking the majority vote among the three annotators. We refer to this data as the *Interactions Dataset*.

The kappa measure between the three groups of annotations was 0.62. To better assess the quality of the annotations, we asked a trained annotator to label 10% of the data. We measured the agreement between the expert annotator and the majority label from the Mechanical Turk. The kappa measure was 0.69.

We trained the classifier that detects sentences with attitude (Section 3.1) on a set of 4000 manually annotated sentences. None of this data overlaps with the dataset described earlier. A similar annotation procedure was used to label this data. We refer to this data as the *Sentences Dataset*.

5 Results and Discussion

We performed experiments on the data described in the previous section. We trained and tested the sentence with attitude detection classifiers described in Section 3.1 using the *Sentences Dataset*. We also trained and tested the interaction sign classifier described in Section 3.3 using the *Interactions Dataset*. We build one unsigned network from every topic in the data set. This results in a signed social network for every topic (e.g. politics, economics,etc.). We decided to build a network for every topic as opposed to one single network because the relation between any two individuals may vary across topics. In the rest of this section, we will describe the experiments we did to assess the performance of the sentences with attitude detection and interaction sign prediction steps.

In addition to classical evaluation, we evaluate our results using the structural balance theory which has been shown to hold both theoretically (Heider, 1946) and empirically (Leskovec et al., 2010b). We validate our results by showing that the *automatically* extracted networks mostly agree with the theory.

5.1 Identifying Sentences with Attitude

We tested this component using the *Sentences Dataset* described in Section 4. In a 10-fold cross validation mode, the classifier achieves 80.3% accuracy, 81.0% precision, %79.4 recall, and 80.2% F1.

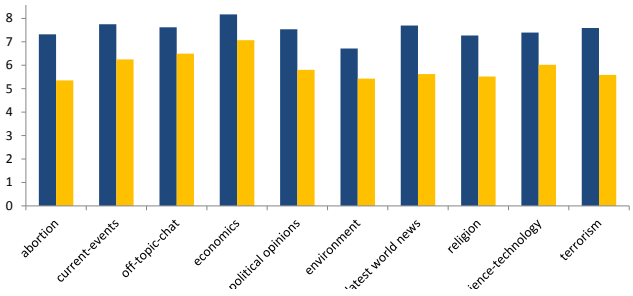

Figure 1: Percentage of balanced triangles in extracted network vs. random network.

5.2 Interaction Sign Classifier

We used the relation detection classifier described in Section 3.2 to find sentences with positive and negative attitude. The output of this classifier was used to compute the the features described in Section 3.3, which were used to train a classifier that predicts the sign of an interaction between any two individuals.

We used Support Vector Machines (SVM) to train the sign interaction classifier. We report several performance metrics for them in Table 2. All results were computed using 10 fold cross validation on the labeled data. To better assess the performance of the proposed classifier, we compare it to a baseline that labels the relation as negative if the percentage of negative sentences exceeds a particular threshold, otherwise it is labeled as positive. The thresholds was empirically evaluated using a separate development set. The accuracy of this baseline is only 71%.

We evaluated the importance of the features listed in Table 1 by measuring the chi-squared statistic for every feature with respect to the class. We found out that the features describing the interaction between the two participants are more informative than the ones describing individuals characteristics. The later features are still helpful though and they improve the performance by a statistically significant amount. We also noticed that all features based on percentages are more informative than those based on count. The most informative features are: percentage of negative posts per tread, percentage of negative sentences per post, percentage of positive posts per thread, number of negative posts, and discussion topic.

5.3 Structural Balance Theory

The structural balance theory is a psychological theory that tries to explain the dynamics of signed social interactions. It has been shown to hold both theoretically (Heider, 1946) and empirically (Leskovec et al., 2010b). In this section, we study the agreement between the theory and the *automatically* extracted networks. The theory has its origins in the work of Heider (1946). It was then formalized in a graph theoretic form in (Cartwright and Harary,). The theory is based on the principles that "the friend of my friend is my friend", "the enemy of my friend is my enemy", "the friend of my enemy is my enemy", and variations on these.

There are several possible ways in which triangles representing the relation of three people can be signed. The structural balance theory states that triangles that have an odd number of positive signs (+ + + and + - -) are balanced, while triangles that have an even number of positive signs (- - - and + + -) are not.

Even though the structural balance theory posits some triangles as unbalanced, that does not eliminate the chance of their existence. Actually, for most observed signed structures for social groups, exact structural balance does not hold (Doreian and Mrvar, 1996). Davis (1967) developed the theory further into the *weak structural balance theory*. In this theory, he extended the structural balance theory to cases where there can be more than two such mutually antagonistic subgroups. Hence, he suggested that only triangles with exactly two positive edges are implausible in real networks, and that all other kinds of triangles should be permissible.

In this section, we connect our analysis to the structural balance theory. We compare the predictions of edge signs made by our system to the structural balance theory by counting the frequencies of different types of triangles in the predicted network. Showing that our automatically constructed network agrees with the structural balance theory further validates our results.

We compute the frequency of every type of triangle for ten different topics. We compare these frequencies to the frequencies of triangles in a set of random networks. We shuffle signs for all edges on every network keeping the fractions of positive and

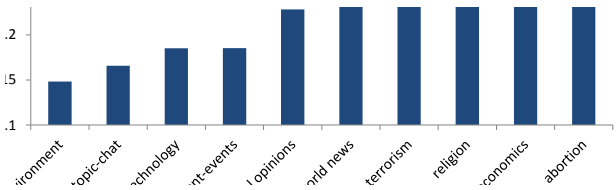

Figure 2: Percentage of negative edges across topics.

negative edges constant.

We repeat shuffling for 1000 times. Every time, we compute the frequencies of different types of triangles. We find that the all-positive triangle $(+++)$ is overrepresented in the generated network compared to chance across all topics. We also see that the triangle with two positive edges $(++-)$, and the all-negative triangle $(---)$ are underrepresented compared to chance across all topics. The triangle with a single positive edge is slightly overrepresented in most but not all of the topics compared to chance. This shows that the predicted networks mostly agree with the structural balance theory. In general, the percentage of balanced triangles in the predicted networks is higher than in the shuffled networks, and hence the balanced triangles are significantly overrepresented compared to chance. Figure 1 compares the percentage of balanced triangles in the predicted networks and the shuffled networks. This proves that our *automatically* constructed network is similar to explicit signed networks in that they both mostly agree with the balance theory.

6 Application: Dispute Level Prediction

There are many applications that could benefit from the signed network representation of discussions such as community finding, stance recognition, recommendation systems, and disputed topics identification. In this section, we will describe one such application.

Discussion forums usually respond quickly to new topics and events. Some of those topics usually receive more attention and more dispute than others. We can identify such topics and in general measure the amount of dispute every topic receives using the extracted signed network. We computed the percentage of negative edges to all edges for every topic. We believe that this would act as a measure for how disputed a particular topic is. We see, from Figure 2, that "environment", "science", and "technology" topics are among the least disputed topics, whereas "terrorism", "abortion" and "economics" are among the most disputed topics. These findings are another way of validating our predictions. They also suggest another application for this work that focuses on measuring the amount of dispute different topics receive. This can be done for more specific topics, rather than high level topics as shown here, to identify hot topics that receive a lot of dispute.

7 Conclusions

In this paper, we have shown that natural language processing techniques can be reliably used to extract signed social networks from text correspondences. We believe that this work brings us closer to understanding the relation between language use and social interactions and opens the door to further research efforts that go beyond standard social network analysis by studying the interplay of positive and negative connections. We rigorously evaluated the proposed methods on labeled data and connected our analysis to social psychology theories to show that our predictions mostly agree with them. Finally, we presented potential applications that benefit from the automatically extracted signed network.

References

Cem Akkaya, Alexander Conrad, Janyce Wiebe, and Rada Mihalcea. 2010. Amazon mechanical turk for subjectivity word sense disambiguation. In *CSLDAMT '10*.

Carmen Banea, Rada Mihalcea, and Janyce Wiebe. 2008. A bootstrapping method for building subjectivity lexicons for languages with scarce resources. In *LREC'08*.

Michael J. Brzozowski, Tad Hogg, and Gabor Szabo. 2008. Friends and foes: ideological social networking. In *SIGCHI*.

Chris Callison-Burch. 2009. Fast, cheap, and creative: evaluating translation quality using amazon's mechanical turk. In *EMNLP '9*, EMNLP '09.

Dorwin Cartwright and Frank Harary. Structure balance: A generalization of heiders theory. *Psych. Rev.*

J. A. Davis. 1967. Clustering and structural balance in graphs. *Human Relations*.

Patrick Doreian and Andrej Mrvar. 1996. A partitioning approach to structural balance. *Social Networks*.

Patrick Doreian and Andrej Mrvar. 2009. Partitioning signed social networks. *Social Networks*.

David Elson, Nicholas Dames, and Kathleen McKeown. 2010. Extracting social networks from literary fiction. In *ACL 2010*, Uppsala, Sweden.

Anatoliy Gruzd and Caroline Haythornthwaite. 2008. Automated discovery and analysis of social networks from threaded discussions. In *(INSNA)*.

Ahmed Hassan and Dragomir Radev. 2010. Identifying text polarity using random walks. In *ACL'10*.

Ahmed Hassan, Vahed Qazvinian, and Dragomir Radev. 2010. What's with the attitude?: identifying sentences with attitude in online discussions. In *Proceedings of the 2010 Conference on Empirical Methods in Natural Language Processing*.

V. Hatzivassiloglou and K. McKeown. 1997. Predicting the semantic orientation of adjectives. In *EACL'97*.

Vasileios Hatzivassiloglou and Janyce Wiebe. 2000. Effects of adjective orientation and gradability on sentence subjectivity. In *COLING*.

Fritz Heider. 1946. Attitudes and cognitive organization. *Journal of Psychology*.

J. Huang, M. Zhou, and D. Yang. 2007. Extracting chatbot knowledge from online discussion forums. In *IJCAI'07*.

Thorsten Joachims, 1999. *Making large-scale support vector machine learning practical*. MIT Press.

Dan Klein and Christopher D. Manning. 2003. Accurate unlexicalized parsing. In *ACL'03*.

Jérôme Kunegis, Andreas Lommatzsch, and Christian Bauckhage. 2009. The slashdot zoo: mining a social network with negative edges. In *WWW '09*.

Lillian Lee. 1999. Measures of distributional similarity. In *ACL-1999*.

A. Lehrer. 1974. Semantic fields and lezical structure.

Jure Leskovec, Daniel Huttenlocher, and Jon Kleinberg. 2010a. Predicting positive and negative links in online social networks. In *WWW '10*.

Jure Leskovec, Daniel Huttenlocher, and Jon Kleinberg. 2010b. Signed networks in social media. In *CHI 2010*.

Jure Leskovec, Daniel Huttenlocher, and Jon Kleinberg. 2010c. Signed networks in social media. In *Proceedings of the 28th international conference on Human factors in computing systems*, pages 1361–1370, New York, NY, USA.

Chen Lin, Jiang-Ming Yang, Rui Cai, Xin-Jing Wang, and Wei Wang. 2009. Simultaneously modeling semantics and structure of threaded discussions: a sparse coding approach and its applications. In *SIGIR '09*.

Andrew McCallum, Xuerui Wang, and Andrés Corrada-Emmanuel. 2007. Topic and role discovery in social networks with experiments on enron and academic email. *J. Artif. Int. Res.*, 30:249–272, October.

George A. Miller. 1995. Wordnet: a lexical database for english. *Commun. ACM*.

Tetsuya Nasukawa and Jeonghee Yi. 2003. Sentiment analysis: capturing favorability using natural language processing. In *K-CAP '03*.

J. Norris. 1997. Markov chains. Cambridge University Press.

Bo Pang, Lillian Lee, and Shivakumar Vaithyanathan. Thumbs up?: sentiment classification using machine learning techniques. In *EMNLP*. Association for Computational Linguistics.

A. Popescu and O. Etzioni. 2005. Extracting product features and opinions from reviews. In *HLT-EMNLP'05*.

E. Riloff and J. Wiebe. 2003. Learning extraction patterns for subjective expressions. In *EMNLP'03*.

Dou Shen, Qiang Yang, Jian-Tao Sun, and Zheng Chen. 2006. Thread detection in dynamic text message streams. In *SIGIR '06*, pages 35–42.

Philip Stone, Dexter Dunphy, Marchall Smith, and Daniel Ogilvie. 1966. The general inquirer: A computer approach to content analysis. *The MIT Press*.

Hiroya Takamura, Takashi Inui, and Manabu Okumura. 2005. Extracting semantic orientations of words using spin model. In *ACL'05*.

P. Turney and M. Littman. 2003. Measuring praise and criticism: Inference of semantic orientation from association. *Transactions on Information Systems*.

Janyce Wiebe, Rebecca Bruce, Matthew Bell, Melanie Martin, and Theresa Wilson. 2001. A corpus study of evaluative and speculative language. In *Proceedings of the Second SIGdial Workshop on Discourse and Dialogue*, pages 1–10.

Janyce Wiebe. 2000. Learning subjective adjectives from corpora. In *AAAI-IAAI*.

Theresa Wilson, Paul Hoffmann, Swapna Somasundaran, Jason Kessler, Janyce Wiebe, Yejin Choi, Claire Cardie, Ellen Riloff, and Siddharth Patwardhan. 2005a. Opinionfinder: a system for subjectivity analysis. In *HLT/EMNLP*.

Theresa Wilson, Janyce Wiebe, and Paul Hoffmann. 2005b. Recognizing contextual polarity in phrase-level sentiment analysis. In *HLT/EMNLP'05*.

Bo Yang, William Cheung, and Jiming Liu. 2007. Community mining from signed social networks. *IEEE Trans. on Knowl. and Data Eng.*, 19(10).

Hong Yu and Vasileios Hatzivassiloglou. 2003. Towards answering opinion questions: separating facts from opinions and identifying the polarity of opinion sentences. In *EMNLP'03*.