ENVIRONMENTAL SCANNING FOR CUSTOMER COMPLAINT IDENTIFICATION IN SOCIAL MEDIA

Research-in-Progress

Ming Yang
School of Management, Harbin Institute of Technology,
Harbin, Heilongjiang, China
yangming.hit@gmail.com

Yijun Li
School of Management, Harbin Institute of Technology,
Harbin, Heilongjiang, China
liyijun@hit.edu.cn

Melody Kiang
Department of Information System, California State University,
Long Beach, California, USA
mkiang@csulb.edu

Abstract

Social media provides a platform for dissatisfied and frustrated customers to discuss matters of common concerns and share experiences about products and services. While listening to and learning from customer has long been recognized as an important marketing charge, how to identify customer complaints on social media is a nontrivial task. Customer complaint messages are highly distributed on social media, while non-complaint messages are unspecific and topically diverse. It is costly and time consuming to manually label a large number of customer complaint messages (positive examples) and non-complaint messages (negative examples) for training classification systems. Nevertheless, it is relatively easy to obtain large volumes of unlabeled content on social media. In this paper, we propose a partially supervised learning approach to automatically extract high quality positive and negative examples from an unlabeled dataset. The empirical evaluation suggested that the proposed approach generally outperforms the benchmark techniques and exhibits more stable performance.

Keywords: Social media analytics, partially supervised learning, document classification, customer compliant management

Introduction

The proliferation of Web 2.0 applications has given birth to a number of social media channels, such as Web forums, Web blogs, microblogs, and social networking, where consumers can discuss matters of common concerns and share their experiences with others. Through social media, dissatisfied and frustrated consumers can easily articulate their opinions and comments on products, services, brands, and firms. The mobile devices further equip the consumers participating in social media the ability to react at the moment an issue arises. Due to the chaotic nature of the Internet and the openness of social media, online complaint messages can be easily and rapidly distributed to a large volume of potential consumers. Moreover, the groups of unhappy customers can form their own forums on social media outside a firm's customer complaint management system using the brand name as a searchable keyword for the forum. Research evidence suggests that harmful consumers' negative word-of-mouth has negative impact on financial analyst's ratings of the firm (Luo 2007). As a result, firms recognize the importance of censoring negative online voice as one critical mission of their enterprise risk management.

In order to identify potential risks, it is important for firms to be able to collect and analyze online complaints on social media in a timely manner. However, assessing online opinions is a nontrivial task due to the high volume of messages, casual writing style, and the significant amount of noise (Zimbra et al. 2009). In addition, the information seeking process is heavily affected by the mental model of the individual manager. Managers need to understand their business environment so as to integrate it into their planning and decision-making process (Ontrup et al. 2009). The term Environmental Scanning (ES) refers to methods to achieve such integrative understanding.

Our research endeavors to develop a new process to scan large amount of text-based customer complaints collected from social media. The proposed method is able to augment the training data through partially supervised learning, which allows the construction of a classifier with a small number of labeled examples and mostly unlabeled data points. We test our method on an important form of social media, Web forum, and the preliminary results are encouraging. The proposed method is helpful for firms to monitor customer to customer dialogs, to correct inaccuracies, and to mitigate potential damages.

The remainder of the paper is organized as follows. Section 2 reviews the technical roots of the project. Section 3 summarizes current research gaps and the objectives of our research. In section 4, we present our method for autonomous identification of customer complaints on social media. In section 5 the preliminary findings from a case study based on a Wal-Mart related forum is discussed. We conclude the paper with a summary of our contributions and suggestions for future research directions in section 6.

Related Works

Customers commonly use plain text to express their complaints on social media. A typical complaint is a report of the failure of a product or service, followed by a narrative on the customer's attempts to resolve the issue (Galitsky et al. 2009). Due to the anonymous nature of the Internet activities, social media also can be misused for reputation manipulation. Reputation manipulation allows online marketers to inflate their own reputations or to sabotage the competitors' reputation through multiple identities (Abbasi et al. 2008c). Since it is almost impossible to verify the actual occurrences of failures about a product or service, observers and potential consumers have to interpret the complaints on the basis of the message itself. Previous research suggested that the receivers' perceptions depended heavily on the manner in which the negative word-of-mouth communication was conveyed (Laczniak et al. 2001). Lee et al. (2010) examined the impact of message characteristics on the observer's perception of a product or service provider, showing that observers who were exposed to online complaints with high level of vividness and high level of consensus were more likely to evaluate the company negatively. Therefore, a major task in customer complaint management is to distinguish those complaints that have high levels of vividness and consensus, and to handle them effectively.

Information Retrieval and Feature Selection

Information Retrieval (IR) can be used to identify relevant customer complaint content in social media. A powerful approach in IR research is the vector space model, which is often employed to convert original documents into vectors in a feature space based on weighted term frequencies (Aasheim et al. 2006;

Coussement et al. 2008; Salton et al. 1986). The discourse in complaint messages can manifest in the form of various information types, including topics, events, opinions, emotions, styles, and interactions. Converting the information types into a feature vector that makes its most salient feature available is an important part of IR. Previous studies have shown that adding stylistic features into a classification model can significantly improve the identifiability of customer complaint emails (Coussement et al. 2008) and the detection of aliasing messages in online reputation systems (Abbasi et al. 2008c).

Stylistic features are the attributes or writing style markers that are the most effective discriminators in statistical analysis of writing style (Abbasi et al. 2008a). The vast array of stylistic features includes lexical, syntactic, structural, content-specific, and idiosyncratic style markers. *Lexical* features are word or character-based statistics of lexical variations (Abbasi et al. 2008c). *Syntactic* features include function words, punctuation, and part-of-speech tag *n*-grams (Abbasi et al. 2008b). *Structural* features, which are especially useful for online text mining, include attributes relating to text organization and layout (Zheng et al. 2006). *Content-specific* features are important key words and phrases pertaining to certain topics. Idiosyncratic features include misspellings, grammatical mistakes, and other usage anomalies (Abbasi et al. 2008c).

Learning with Positive and unlabeled Data

Supervised learning has long been the dominant approach for autonomous identification of customer complaints (Coussement et al. 2008; Galitsky et al. 2009). Supervised learning algorithms require high-quality labeled training data in order to construct an accurate classifier. However, customer complaint messages are highly distributed on social media. It is often a mentally exhausting, if not infeasible, process to manually acquire and label complaint messages for training a classifier. In addition, managers tend to handle the complaint messages related to their firm's interest while ignoring non-complaint messages. Previous research also suggested that firms should apply an adaptive complaint handling approach to avoid misallocation of attention, energy, and resources (Homburg et al. 2010). One way to overcome the difficulties is to dynamically augment the training data through a partially supervised learning, which constructs classifiers based on mostly unlabeled data and a small number of labeled positive examples that are of interest to the users (Zhu 2005). In many information retrieval applications, positive examples refer to the data points that are of interest to the researchers in a binary classification problem. In this research, we assign customer complaint messages as positive examples and non-complaint messages as negative examples to train the classifiers.

The positive class is usually more specific than the negative class (Zhou et al. 2010). Hence, it is possible to identify more potentially positive examples from the unlabeled data through exploiting the inherent structures in the set of positive examples. Ko et al. (2005) proposed a technique called EAT (Example Adaption for Text categorization) for automatically seeking more representative positive examples from the unlabeled documents. This approach consists of two steps: first, extracting a set of potentially positive examples from an unlabeled dataset; second, generating a set of classifiers iteratively through gradually increasing the number of positive examples until the classifier reaches its local maximum accuracy level. The effectiveness of the EAT is based on the content-specific features that capture the characteristics of the positive class. However, the content-specific features for the EAT are manually crafted from a very small number of sample documents. Thus, the effectiveness of the classifier is highly dependent on the quality of the content-specific features given.

Nevertheless, it is not an easy task to extract a proper set of positive examples due to the diversity of topics exhibited in unlabeled messages (Fung et al. 2005). In order to solve this problem, Fung et al. (2006) proposed an approach called PNLH (Positive examples and Negative examples Labeling Heuristics) using partition-based heuristics that iteratively extract reliable positive and negative examples from an unlabeled dataset (Fung et al. 2006). The effectiveness of this approach depends on the core vocabularies of the positive examples (i.e., positive features). To be more specific, the underlying assumption of this approach is that the positive features are sufficient to capture the characteristics of the positive class. When this assumption fails, that is, when the available positive features are not representative of the true positive class, the performance of this method may deteriorate.

Research Gaps and Research Objectives

Based on our review of the related works, we have identified several important research gaps. Due to the fact that customer complaint messages on social media is highly distributed and non-complaint content is very diverse in topics, it is costly and time consuming to label a large number of training data for the supervised learning algorithms to identify customer complaint content. One way to overcome such difficulties is to dynamically augment the training data through a partially supervised learning process. The method requires that the positive features identified are sufficient to capture the vividness and the consensus of the customer complaints, because the potential consumers' attributions depend on the levels of vividness and consensus of the negative word-of-mouth communications conveyed. However, there has been limited use of textual features to characterize the vividness of customer complaint messages in previous studies. Moreover, few studies have provided approaches to detect the community's consensus of customer complaints.

Based on these research gaps, we propose the following research objectives:

1) To characterize the vividness of customer complaints using textual features.

2) To detect the community's consensus of customer complaint messages.

3) To propose the partially supervised learning approach for identifying customer complaint messages on social media based on the vividness and consensus of the complaints.

Research Design

This study is aimed at designing and examining a new approach to identify customer complaints on Web forums. Specifically, we develop and evaluate informatics tools and frameworks using a partially supervised learning algorithm to monitor customer complaints at targeted Web forums. The proposed system has three components – vividness characterization, consensus detection, and classifier construction (as shown in Figure 1). In order to improve text categorization and analysis capabilities, we use an extended stylistic feature set to characterize the discourse structure of a message. Then we propose a partially supervised learning algorithm to evaluate the level of consensus corresponding to certain customers' concerns and subsequently identify more related customer complaints. The input of the algorithm is a set of labeled positive examples (customer complaints) and an unlabeled dataset. The output of the approach is a classifier which is capable of identifying customer complaints covering diverse topics with high level of consensus.

Vividness Characterization

According to Nisbett and Ross (1980, p.45), vividness refers to information capacity to "attract and hold attention to excite imagination." Information is vivid if it contains a degree of "emotionally interesting, concrete and imagery-provoking, and/or proximate in a sensory, temporal, spatial way". Previous research has shown that vivid information has more influence on consumers' attitudinal judgment than non-vivid information (Kim et al., 1991). Vividly presented information is more persuasive when the information is personally relevant and produces emotionally arousing response (Taylor & Thompson, 1982; Taylor & Wood, 1983). An important part of message characterization is the vividness of the writing style. In this research, we use the extended stylistic features to characterize the vividness of a complaint message. Previous studies have shown that by adding stylistic features to a classification model can significantly improved the identifiability of customer complaint emails (Coussement et al. 2008) and the detection of aliasing messages in online reputation systems (Abbasi et al. 2008c).

During the feature extraction phase we derive both static and dynamic features from the messages and the result is an extended feature set. The static feature is simply the statistics of the feature usage across text; while the dynamic feature categories such as *n*-grams require indexing and feature selection. The feature extraction procedure for the extended feature set is described in details below. Table 1 provides a description of the stylistic features included. For dynamic feature categories, the number of attributes varies depending on the indexing and the feature selection processes. In this study, the dynamic features incorporated in the extended feature set are the *n*-gram feature groups, including word, character, part-of-speech tag, and digit-level *n*-grams. These categories require indexing with the number of initially indexed features varying by the dataset. The resulting indexed features are then forwarded to the next step: the feature selection phase.

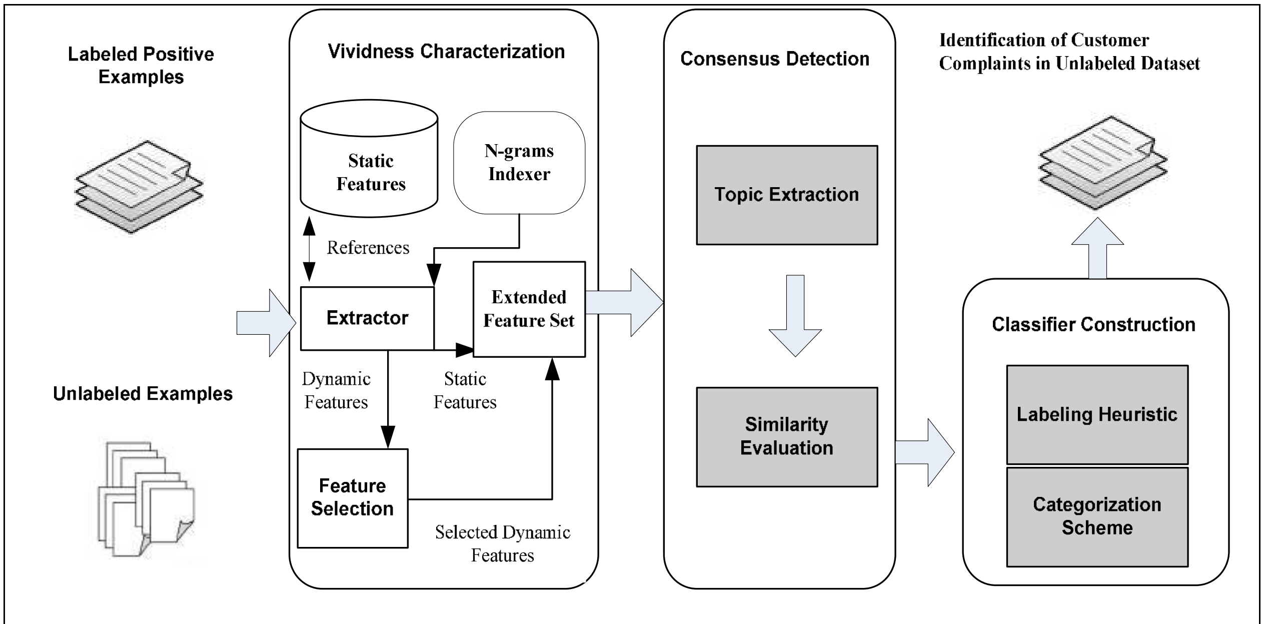

Figure 1. The System Architecture

Table 1. The Extended Feature Set

Group	Category	Quantity	Description/examples
Lexical	Word level	5	Total words, percent characters per word
	Character level	5	Total characters, percent characters per message
	Character n-grams	<16,247	Count of letter n-grams (e.g., a, at, att)
	Digit n-grams	<1,020	Count of digital n-grams (e.g., 1, 12, 123)
	Word-length distri.	20	Frequency distribution of 1-20 letter words
	Vocabulary richness	8	Ratio of words which occur only once
	Special characters	21	Occurrences of special characters (e.g., @, ^, $)
Syntactic	Function words	300	Frequency of function words (e.g., of, for, to)
	Punctuation	8	Occurrence of punctuation marks (e.g., ! , : ; ?)
	POS tag n-grams	Varies	Part-of-speech tag n-grams (e.g., NNP, NNP JJ)
Structural	Message level	6	Ex. requoted content, has URL, more…
	Paragraph level	8	Number of paragraphs, sentences per paragraph
	Technical structure	50	Ex. fonts, use of multimedia
Content specific	Word n-grams	Varies	Bag-of-word n-grams (e.g., “seller”, “bad sale”)

The feature selection is applied to all the *n*-gram feature groups using information gain (IG) heuristic. Information gain is an effective text feature selection technique and has been used in many text categorization studies (Abbasi et al. 2008a; Abbasi et al. 2008c; Abbasi et al. 2008d). It is computationally efficient compared to the search-based techniques and works well for multi-class text problems (Yang et al. 1997). The information gain for feature *j* across a set of classes *c* is derived as $IG(c, j) = H(c) - H(c \mid j)$, where *H(c)* is the overall entropy across all classes and *H(c|j)* is the conditional entropy for feature *j*. For each topic of a customer complaint, information gain is applied using a two-class (one-against-all) setup (size of $c=2, c_1=\text{topic A}, c_2=\text{rest}$). Thus, each feature set of a certain topic comprises the set of dynamic features that can best distinguish that particular customer complaint topic against the others.

Consensus Detection

When people encounter negative information of certain product or service, they oftentimes would first consider the reactions of others prior to making their own judgment regarding the cause of the problem (Lapinski et al. 2005). We define consensus as broad declarative statements reflecting a majority of certain group of consumers' opinions. Consensus in the communication can provide great impact on individual's causal judgment when the communication involves negative information (Conway et al. 1990). However, on social media the customer complaint messages that are of interest to the users usually consist of diverse topics and are highly distributed. Hence, it is difficult to detect the community's consensus for further analysis. In the proposed method, we detect the consensus of the customer complaint messages via two steps, i.e., topic extraction and similarity evaluation.

To extract topics from the labeled positive training set *S*, we first employ domain experts to identify representative messages as seeds from the set *S*. Then, using the seeds, we perform text clustering on the entire set *S* based on the nearest neighbor clustering algorithm. After the clustering process, the clusters contain more than τ documents are considered informative clusters. For each identified informative cluster, we apply a Mutual Information (MI)-based noun phrase extractor, Arizona Noun Phraser (Tolle et al. 2000), to extract major terms that can represent the topic.

Since the number of labeled positive examples is much smaller than that of the unlabeled examples, in order to detect the consensus of the topics extracted from the informative clusters, we need to evaluate the similarity of all messages in the unlabeled set *P* for each topic extracted from the informative clusters. If the similarity of a message in *P* with at least one topic pattern of the informative cluster is greater than a given threshold ε , the message is selected and introduced into the available repertoire *A*. At the end of the process, the available repertoire *A* is merged with the original positive training set *S*, to form the final positive training set that contains customer complaints with high level of consensus; while the rejected messages form the negative training set. Through this method, customer complaints with high level of consensus are identified, in the meantime more positive examples and negative examples for training can be obtained. Figure 2 outlines the process of similarity evaluation.

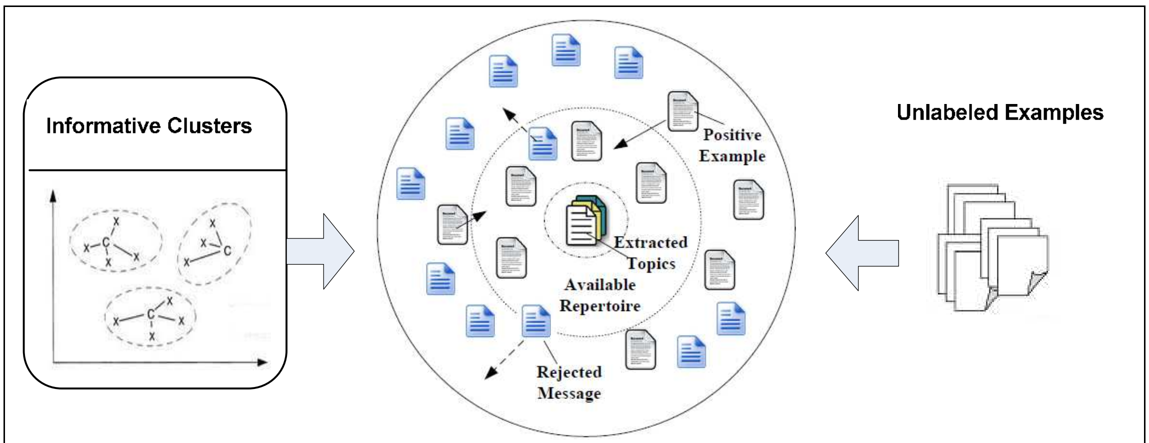

Figure 2. The Similarity Evaluation Process

The algorithm for similarity evaluation aims at selecting messages from unlabeled set (*P*) into either available repertoire or negative examples set (*N*). It takes two inputs, unlabeled set (*P*) and informative clusters (S^M). For each $S_j^M \in S^M$, we let all messages in *P* be negative examples and those messages in S_j^M be positive examples. Then two prototype vectors *x* and *y*, corresponding to positive and negative prototype respectively, are learned by the Rocchio algorithm. For each $p_i \in P$, we calculate its similarity with S_j^M as follows:

$$\varphi(p_i, S_j^M) = \frac{p_i \cdot x}{\|p_i\| \cdot \|x\|} - \frac{p_i \cdot y}{\|p_i\| \cdot \|y\|}$$

The Rocchio algorithm is utilized to build a classifier using each informative cluster and the unlabeled set. The classier is then applied to select more positive examples into available repertoire and delete them from unlabeled set. The idea is to identify more positive examples in unlabeled dataset in a localized manner. In terms of similarity measurement, p_i may be similar to both S^M and N_i. We only extract p_i which is significantly similar to either S^M or N_i . The output of the similarity evaluation process is available repertoire and negative examples set.

Classifier Construction

Previous studies have found that the positive examples extracted through comparing the differences of the feature distributions between the positive class and the unlabeled dataset does not always guarantee all the extracted positive examples are reliable (Fung et al. 2006). Moreover, when too many labeled examples were extracted, it may cause over-fitting problem and result in performance degradation (Cohen et al. 2004). In order to solve these problems, we select reliable positive and negative examples to fit the distribution of the positive class and the negative class, respectively. The process is controlled by iteratively running a classification scheme C_m. We choose Support Vector Machine (SVM) as the text classifier due to its popularity and superb performance in text classification (Fung et al. 2006). To guarantee the quality of positive examples in available repertoire in a stable state, we keep the change of the threshold $\Delta\varepsilon$ during *m*th iteration within a limited range. We used the following updating rule: $\Delta\varepsilon = e^{-m/\sigma}$, where *m* is the number of iteration, σ is the parameter that control its decay.

Table 2. Classier Construction Algorithm

Input: A_0 (available repertoire) and N_0 (rejected messages set), *P* (potential repertoire), C_0 (a classifier built by A_0 as the positive set and N_0 as the negative set), and F_0 (the F-measure of C_0 for the positive class);
Output: A_m (updated available repertoire) and N_m (updated rejected messages set);
$A_m \leftarrow A_0; N_m \leftarrow N_0;$
Initialize $F_{\max} = F_0; m = 1;$
repeat
    $\varepsilon = \varepsilon_0 * (1 + \Delta\varepsilon_m)$
    Obtain *A* and *N* through the similarity evaluation;
    $A_m \leftarrow A; N_m \leftarrow N;$
    Construct classifier C_m using A_m as the positive training set and N_m as the negative training set;
    Let F_m be the F-measure of C_m for the positive class;
    if $F_m \geq F_{\max}$ **then**
        $F_{\max} = F_m;$
    end if
    m++;
until $F_{m-1} < F_{\max}$
return A_m and N_m ;

Table 2 shows the steps for the classifier construction. To build classifier C_m, we let A_m be the positive training set that consists of customer complaints with high level of vividness and consensus, and let N_m be the negative training set that consists of messages that could not match any topic extracted from the informative clusters. The classifier C_m is selected based on its local maximum F_1 score (for positive class) on the dynamically updated training data (in Lines 8-13). The F_1 score is defined as follows:

$$F_1(i) = \frac{2 \times \text{precision}(i) \times \text{recall}(i)}{\text{precision}(i) + \text{recall}(i)}$$

where $$\text{precison}(i) = \frac{\text{number of correctly classified cases for class } i}{\text{total number of cases classified as class } i}$$

$$\text{recall}(i) = \frac{\text{number of correctly classified cases for class } i}{\text{total number of cases in class } i}$$

Hence, customer complaints (positive examples) which have high level of similarity with the topic of interest will be used to construct the classifier first. This process iterates until it finds the best classifier that renders the highest local maximum F_1 score, in other words the iterative process stops when the F_1 score starts to decrease. This allows us to construct a classifier that can identify customer complaint messages that cover diverse topics with high level of consensus.

Experiment Design

Dataset

For proof of concept, we conduct a case study on a Wal-Mart related discussion Web forum: the "Wal-Mart Sucks" board of the Wal-Mart blows forum (http://www.warmart-blows.com/forum/viewforum.php?f=3). The dataset consists of 1,354 threads with 19,624 postings by 1,885 authors between November 2003 and November 2008. Discussions on this forum mainly express negative opinions on Wal-Mart. The online discussions are first collected by Web crawlers, then parsed and stored in a relational database for analysis. A domain expert manually goes through each posting and determines if it belongs to complaint or non-complaint message. Among the entire collected corpus, 3140 (16.6%) postings are complaint messages, and 16,484 (83.4%) postings are non-complaint messages.

Experiment Setup

In this study, we assume only positive and unlabeled examples are available for training purpose, so we randomly pick 40% of the customer complaint messages as the positive training examples, and another 40% of the customer complaint messages and 40% of the non-complaint messages as the unlabeled examples. To maintain the same class distribution in both the training and testing sets, we create our testing set from the remaining 20% of the customer complaint messages and a randomly selected 20% of non-complaint messages from the remaining pool.

To minimize the potential bias from the randomized sampling process and obtain more reliable performance estimates, we perform the same experiment process 30 times and report the average performance from the 30 individual trials. We preprocess all messages in the dataset, including stop word removal and finding the word stems.

System Evaluation

To assess the effectiveness of the proposed customer complaint identification method, we implement and compare its performance with that of two benchmark labeling heuristics, the EAT and the PNLH (discussed in Related Works). The three labeling heuristics are comparable in the following aspects: first, they are all independent from the classifiers to be implemented; second, they all employ the common idea of enlarging the positive training examples during the learning process. As the implementations of the

EAT and the PNLH are not publicly available, we implemented the two systems based on the descriptions in (Ko et al. 2005) and (Fung et al. 2006), respectively.

An objective of the experiment is to compare the sensitivity of the proposed approach and the benchmark techniques to the ratio of the positive training examples. Table 3 shows the performance of the three labeling heuristic methods when the relative size of the positive training examples varies from 20% to 40% in the training set. We also calculated the *p*-values of the paired-sample Wilcoxon signed-rank test between the two benchmark labeling heuristics and the proposed approach.

Table 3 shows the proposed approach outperforms the two benchmark techniques in all cases. This indicates that the proposed approach can provide better identification power for the detection of customer complaints in Web forums than the benchmark methods. When the proportion of the labeled positive examples increases, the performances of all labeling heuristics improve as well. Hence, all three methods are capable of improving the identification power of the classifiers via making use of the labeled positive examples.

Table 3. Experiment Results (F_1 score %) from the Case Study

% of positive examples	20%	25%	30%	35%	40%
Our Approach	**70.14%**	**73.27%**	**79.61%**	**84.23%**	**91.51%**
EAT	49.13%*	54.67%*	59.26%*	68.97%*	81.45%*
PNLH	53.17%*	58.31%*	67.19%*	75.98%*	85.73%**

**p-values significant at alpha*<0.01

***p-values significant at alpha*<0.05

The *p*-values of the significance test results in Table 3 show that the proposed partially supervised learning approach significantly outperforms the benchmark techniques in all cases. In addition, we find that the performance differences between the proposed approach and the two benchmark techniques increases as the percent of positive examples decreases.

Conclusions and Future Research Directions

In this research, we propose a social media monitoring system that combines an information retrieval and a partially supervised learning algorithm to identify the customer complaint messages on social media. Through this method, we can automatically augment the training data, thus build a more robust classifier for customer complaint messages identification on social media. Using a Wal-Mart related Web forum for our case study, we find that the proposed approach generally outperforms the benchmark techniques and exhibits more stable performance, in particular, when existing knowledge about certain complaints are scarce. The goal of this study is to contribute to the development of online social media monitoring system for customer complaint management. First of all, the proposed method integrates a set of techniques that can handle diverse topics in online discussion. Secondly, the proposed method can capture the characteristics of positive class with diverse topics that are of interest to the company. Thirdly, the proposed method can also extract reliable customer complaints with high level of vividness and consensus, which tends to exhibit high level of influence on the behaviors of the recipients.

In order to better validate the proposed system, we are in the process of collecting more datasets and conducting more comprehensive experiments and analysis. We have started collecting forum discussions reflecting customer opinions on Carrefour, Wal-Mart's largest competitor in China, to control the possible performance differences across companies. We are also in the process of examining the effectiveness of the proposed system on datasets with different characteristics (i.e., the reply network in discussion and recommendation network). We will also extend our comparative experiments to include the implementation of different topic extraction and consensus detection techniques to assess the effectiveness of the different techniques on different datasets.

References

Aasheim, C., and Koehler, G.J. 2006. "Scanning World Wide Web documents with the Vector Space Model," *Decision Support Systems* (42:2), pp 690-699.

Abbasi, A., and Chen, H. 2008a. "Writeprints: A Stylometric Approach to Identity-level Identification and Similarity Detection in Cyberspace," *ACM Transactions on Information Systems* (26:2), pp 7-35.

Abbasi, A., and Chen, H.C. 2008b. "Cybergate: A Design Framwork and System for Text Analysis of Computer-mediated Communication," *MIS Quarterly* (32:4), pp 811-837.

Abbasi, A., Chen, H.C., and Nunamaker, J.F. 2008c. "Stylometric Identification in Electronic Markets: Scalability and Robustness," *Journal of Management Information Systems* (25:1), pp 49-78.

Abbasi, A., Chen, H.C., and Salem, A. 2008d. "Sentiment Analysis in Multiple Languages: Feature Selection for Opinion Classification in Web Forums," *ACM Transactions on Information Systems* (26:3), pp 12-45.

Cohen, I., Cozman, F.G., Sebe, N., Cirelo, M.C., and Huang, T.S. 2004. "Semisupervised Learning of Classifiers: Theory, Algorithms, and Their Application to Human-computer Interaction," *IEEE Transactions on Pattern Analysis and Machine Intelligence* (26:12), pp 1553-1567.

Conway, M., Difazio, R., and Bonneville, F. 1990. "Consensus and Causal Attributions for Negative Affect," *Journal of Social Psychology* (130:3), pp 375-384.

Coussement, K., and Van den Poel, D. 2008. "Improving Customer Complaint Management by Automatic Email Classification Using Linguistic Style Features as Predictors," *Decision Support Systems* (44:4), pp 870-882.

Fung, G.P.C., Yu, J.X., Hongjun, L., and Yu, P.S. 2005. "Text Classification without Labeled Negative Documents," in *Proceedings. 21st International Conference on Data Engineering*, IEEE Computer Society, Washington, DC, USA, pp. 594-605.

Fung, G.P.C., Yu, J.X., Lu, H.J., and Yu, P.S. 2006. "Text Classification without Negative Examples Revisit," *IEEE Transactions on Knowledge and Data Engineering* (18:1), pp 6-20.

Galitsky, B.A., Gonzalez, M.P., and Chesnevar, C.I. 2009. "A Novel Approach for Classifying Customer Complaints through Graphs Similarities in Argumentative Dialogues," *Decision Support Systems* (46:3), pp 717-729.

Homburg, C., Furst, A., and Koschate, N. 2010. "On the Importance of Complaint Handling Design: a Multi-level Analysis of the Impact in Specific Complaint Situations," *Journal of the Academy of Marketing Science* (38:3), pp 265-287.

Kim, J., Kardes, F. R., and Herr, P. M. 1991. “Consumer Expertise and the Vividness Effect: Implication for Judgment and Inference,” *Advances in Consumer Research*, 18, pp 90–93.

Ko, H.M., and Lam, W. 2005. "A New Approach for Partially supervised Online News Classification," in *Web and Communication Technologies and Internet -Related Social Issues*, S. Shimojo, S. Ichii, T.W. Ling and K.H. Song (eds.), Springer-Verlag Berlin, Berlin, pp. 238-247.

Laczniak, R.N., DeCarlo, T.E., and Ramaswami, S.N. 2001. "Consumers' Responses to Nnegative Word-of-mouth Communication: An Attributions Theory Perspective," *Journal of Consumer Psychology* (11:1), pp 57-73.

Lapinski, M.K., and Rimal, R.N. 2005. "An Explication of Social Norms," *Communication Theory* (15:2), pp 127-147.

Lee, Y.L., and Song, S. 2010. "An Empirical Investigation of Electronic Word-of-mouth: Informational Motive and Corporate Response Strategy," *Computers in Human Behavior* (26:5), pp 1073-1080.

Luo, X.M. 2007. "Consumer Negative Voice and Firm-idiosyncratic Stock Returns," *Journal of Marketing* (71:3), pp 75-88.

Nisbett, R., and Ross, L. 1980. *Human Inference: Strtegies and Shortcomings of Social Judgment*, Englewood Cliffs, NJ: Prentice-Hall.

Ontrup, J., Ritter, H., Scholz, S.W., and Wagner, R. 2009. "Detecting, Assessing, and Monitoring Relevant Topics in Virtual Information Environments," *IEEE Transactions on Knowledge and Data Engineering* (21:3), pp 415-427.

Salton, G., and McGill, M.J. 1986. *Introduction to Modern Information Retrieval*, New York, NY: McGraw-Hill, Inc.

Taylor, S.E., and Wood, J.V. 1983. “The Vividness Effect: Making a Mountain out of a Molehill,” *Advances in Consumer Research*, 10, pp 540–542.

Taylor, S.E., and Thompson, S.C. 1982. “Stalking the Elusive ‘Vividness’ Effect,” *Psychological Review*, 89, pp 155–181.

Tolle, K.M., and Chen, H.C. 2000. "Comparing noun phrasing techniques for use with medical digital library tools," *Journal of the American Society for Information Science* (51:4), pp 352-370.

Wilcoxon, F. 1945. "Individual Comparisons by Ranking Methods," Biometrics Bulletin (1:6), pp 80-83.

Yang, Y., and Pedersen, J. 1997. "A Comparative Study on Feature Selection in Text Categorization," in *Proceedings of the 14th International Conference on Machine Learning*, Morgan Kaufmann Publishers Inc., pp. 412-420.

Zheng, R., Li, J.X., Chen, H.C., and Huang, Z. 2006. "A Framework for Authorship Identification of Online Messages: Writing-style Features and Classification Techniques," *Journal of the American Society for Information Science and Technology* (57:3), pp 378-393.

Zhou, K., Xue, G.R., Yang, Q., and Yu, Y. 2010. "Learning with Positive and Unlabeled Examples Using Topic-Sensitive PLSA," *IEEE Transactions on Knowledge and Data Engineering* (22:1), pp 46-58.

Zhu, X. 2008. "Partially supervised Learning Literature Survey." http://pages.cs.wisc.edu/~jerryzhu/research/ssl/ semisupervised.html.

Zimbra, D., Fu, T., and Li, X. 2009. "Assessing Public Opinions Through Web 2.0: A Case Study on Wal-Mart," in *Proceedings of the 30th International Conference on Information Systems*, Phoenix, AZ, Paper 67.