AFFECTIVE ALGORITHM TO POLARIZE CUSTOMER OPINIONS

Domenico Consoli, Claudia Diamantini and Domenico Potena
Dipartimento di Ingegneria Informatica, Gestionale e dell'Automazione
Universitá Politecnica delle Marche, Via B. Bianche, 60131, Ancona, Italy

Keywords: Customer Opinions, Web 2.0, Sentiment Analysis, Affective Analysis, Opinion Mining, Opinion Polarity Ekman Emotional Indexes.

Abstract: Human interact with other people and exchange reviews and ideas via web. With the explosion of Web 2.0 platforms such as blogs, discussion forums, peer-to-peer networks, and various other types of social media, consumers share, on the web, their opinions regarding any product/service. Opinions give information about how product/service and reality in general is perceived by other people. Emotional needs are associated with the psychological aspects of product ownership. The customer when writes his reviews on a product/service transmits emotions in the message that he/she feels first and after purchasing the product. For the enterprise understanding customer emotional needs is vital for predicting and influencing their purchasing behaviour. In this paper, we polarize, with original algorithm, customer opinions basing on emotional indexes that are used for decipher, in affective key, facial expressions and emotional lexicon.

1 INTRODUCTION

Emotions are treated largely in affective computing (Picard, 1997) that focuses on improving the interaction between user and computer. The affective computing aims to build emotional machines (Khashman, 2008) that can recognize and express emotions. The computer, in future, will understand the mood of operator and they'll act accordingly; also the computer, with its behavior, will arouse emotions in the operator. We'll have a *friendly computer* with technological and advanced emotional interfaces. The user affective state is present in many ways, through facial expressions, speech and text; also the text brings an intrinsic emotional content. We must not forget that today computers, for the most part, processing only textual data and that the text analysis, based on the sentiments, has increased with the availability of large amount of web pages.

In this work, we focus on the analysis of the opinions that customers express over the Web. Customers exploit web 2.0 tools (chats, forums, blogs, and so on) for expressing their opinions about a product and suggesting solutions for improving it. There are various web sites that collect free customer reviews: *epinions.com, cnet.com, complaints.com, planetfeedback.com, ecomplaints.com, dooyoo.it.*

Customer opinions constitute a gold mine for an enterprise, both for the improvement of products and for the reinforcement of the customer loyalty. In order to full exploit the information contained in customer opinions, it is important to polarize opinion about (part of) product and in particular to return whether the opinion is positive or negative.

In our paper we present an original algorithm, based on emotional indexes, to polarize customer opinions. This paper is organized as follows: in the next section we give a brief description of the emotional purchase process, while in the third section, related works for sentiment analysis are discussed. The fourth and fifth sections are devoted to present our approach to customers opinions polarization and the results. Finally some conclusions are drawn.

2 EMOTIONS IN CUSTOMER BEHAVIOR

Emotions play a crucial role in all phases of purchase behavior. The customer when writes a review about a product/service, expresses also the same emotions that she felt first and after purchasing the product; so from a enterprise point of view it is important to understand what drives customer to choose a particular product/service. Understanding customer emotional

Consoli D., Diamantini C. and Potena D. (2009).
AFFECTIVE ALGORITHM TO POLARIZE CUSTOMER OPINIONS.
In *Proceedings of the 11th International Conference on Enterprise Information Systems - Human-Computer Interaction*, pages 157-160
DOI: 10.5220/0001851601570160

needs is vital for predicting and influencing their purchasing behavior.

Many words express a direct affective sentiment (*"joy"*, *"sorrow"*, *"happiness"*) other indirect (*"cry"*, *"smile"*, *"monster"*, *"abandoned"*). Some words (e.g. *"nail"*, *"ball"*, *"table"*), if decontextualized, don't express neither directly or indirectly emotions, also if the sentence, where these words belong to, has a specific emotional polarity. For example, the word *"nail"* means an object used to fix a picture to a wall and can remind us positive or negative emotions; it depends from the context. The *"ball"*, that easily we associate with a game, has likely an emotional positive contribution.

Many times the emotional/symbolic traits are highly representative of the specific identity and brand. For example, in *luxury goods*, the emotional aspects as brand, uniqueness and prestige for purchasing decisions, are more important than rational aspects such as technical, functional or price. For the seller would be important to measure accurately the value generated by emotions to determine so an appropriate price.

Another factor that influence purchasing customer is, for example, the *disgust* that plays a key role in the inverse relationship between attitude and intention to purchase. Customer don't buy products disgusting. The disgust is a repugnance toward any object, action or person. The disgusting is an index of variation of the intention to purchase.

The relationship that establishes between a brand and a customer is the the same among people. Between company and customer establishes an emotional relationship. The creation and the excitement of a series of positive emotions into customers helps company in the sell. The enterprise can involve customers to increase their sense of belonging to a community that shares the same brand and the same values with other people. The enterprises often encourage exchange of opinions, by making available virtual communities, e.g. Italian Nikon's Camera forum, where people review Nikon products (http://www.nital.it/forum/), the blog on Benetton products (http://benettontalk.com), and so on.

3 RELATED WORKS

In the literature, various methods have been proposed for opinion mining and sentiment analysis(Liu et al., 2003).

In the *Keyword Spotting* (Boucouvalas and Zhe, 2002) approach, text is classified into affective categories based on the presence of fairly unambiguous affective words like *"distressed*, *"happiness* and *"anger*.

All terms that describe emotional states represent the most direct way to communicate emotion by text. The simplest and most used analysis is based on the search for keywords (like *"happy*, *"sad*, *"angry*, etc).

The *Lexical Affinity* (Valitutti et al., 2004) method assigns to words a probabilistic affinity (trained from linguistic corpora) for a particular emotion. For example, *"accident"* might be assigned a 80% probability of being indicating a negative affect, as in *"car accident"*, *"hurt by accident"*.

Esuli and Sebastiani (Esuli and Sebastiani, 2006) have created SentiWordNet, a lexical resource for opinion mining, where they assign to each synset (set of synonyms) of WordNet a sentiment scores: positivity, negativity and objectivity (i.e. neutral). The opinion is positive if the positivity of its terms is higher than negative and objective scores and viceversa for negative opinion. WordNet Affect (Strapparava and Valitutti, 2004) is a linguistic resource for a lexical representation of affective knowledge. In WordNet Affect each synset of WordNet is labeled by one or more affective-labels, representing the affective meaning of the synset. Examples of affective-labels are emotion, mood, trait, cognitive state, physical state, etc.

The original algorithm that we have developed mainly focuses on six Ekman emotional indexes (Ekman, 2007): *"anger"*, *"disgust"*, *"fear"*, *"happiness"*, *"sadness"*, and *"surprise"*. We use these components to study emotional lexicon.

4 ALGORITHM TO POLARIZE CUSTOMER OPINIONS

The goal of our approach is to polarize a customer opinion about a topic, that is a characteristic of a (part of) product/service. We use this approach in the conceptual framework CeC (Consoli et al., 2008) for gathering customer opinion and improving the product/service.

We first consider customer opinions concerning a some topic, then we break them into homogenous sentences, next each sentence is split in words. The polarity of a sentence is given by estimating the polarity of any its words.

To divide a sentence in words we make a preprocess: elimination of stop-words (articles, conjunctions, prepositions) and division of the sentence into single words with lemmatization (Berry and Castellanos, 2007). In the preprocessing we have used the GATE library with some modification in order to reduce the complexity of the analysis. The output of the preprocessing phase is a dataset of significant words of each opinion.

Analysing this dataset, we create a matrix of word-document occurrences containing the number of times a word appears in a document. This matrix is then compressed by SVD algorithm (Okša and Vajteršic, 2001) to obtain the words-words co-occurrence matrix W, where each word is represented by a n dimensional statistical vector. Each dimension of such statistical vector represents the co-occurrence of the single word with other words of the matrix.

Words in W can be divided in affectives and non-affectives. The former can be further classified into direct affective words, that directly refers to affective states (e.g. ”*fear*”, ”*cheerful*”, ”*disgusted*”), and indirect affective words, having an indirect affective impact that depends on the context, i.e. words representing moods, situations eliciting emotions, or emotional responses like ”*monster*” or ”*cry*”.

Affective words will be exploited to build up the model that will allow us to assign a polarity value to non-affective words. For this reason we call training set, with an abuse of notation, the set of affective words, and test set the non-affective words. In the following, it is described the algorithm for building the polarity model and to assign a polarity value to a non-affective word:

1) Assign an affective vector to each word of the training set.

The components of the affective vector represent the contribution of the following states in the emotional definition of the word: *“happy”*, *“sad”*, *“angry”*, *“surprise”*, *“fear”*, and *“disgust”*. In our approach, the affective components have an integer value between 0 and 10. For example, if we associate to the word ”*stench*” the affective vector (0, 2, 2, 0, 0, 6), it means that in the definition of the affective meaning of the word stench, the elements sad and angry contribute with a small value, the element disgust with a high value and other elements don't produce any contribution.

2) For each word t of test set, select the word w of the training set, maximizing the normalized scalar product k :

$$k = \frac{t_s . w_s}{\|t_s\| . \|w_s\|}$$

where t_s and w_s are the statistical vectors of t and w respectively. The resulting w represents the most similar word to t in the co-occurrence space. Our idea is that word with similar statistical vector have also similar affective meaning.

3) Compute the affective vector t_a of t:

$$t_a = k \cdot w_a$$

where w_a is the affective vector of w and the function int(x) gives the integer part of x.

4) Compute the polarity value of t:

The word has a positive value if the sum of components related to positive affective states (happy, surprise) is greater than the sum of components of negative affective states (sad, angry, fear, disgusted); the word polarity is negative in the other case. Finally, the sum of the affective vectors of all words of an sentence defines its affectivity and, consequently, its polarity value.

5 CASE STUDY

In order to test the validity of our algorithm on the opinion polarity we have used almost 1000 customer opinions about a resort in Sharm el-Sheikh and in particular we selected opinions on some services: Kitchen, Restaurant, Room Service, and Administration.

Opinions were collected from various Internet sites, like alpharooms.com, realholidayreports.com. After the pre-processing phase, we derive 11900 significant words, divided in 2300 affective words and 9600 non-affective ones.

In order to show the results of our algorithm we consider two typical sentences taken from separate posts: the first expresses globally positive opinion while second a negative one. The sentences that we have taken into account are follows:

- **Sentence 1.** *Food was excellent and a great variety, especially if you are fond of sweets.*(positive opinion)
- **Sentence 2.** *Beware of the dangerous on the bad beach and disgusting and terrible pasta.*(negative opinion)

The resulting affective vectors of any words in these sentences, are shown in Table 1 and Table 2. In Tables words already belong to the training set are in italic. In the tables there are several zeros. If a word has a higher value of *happy* probably it has a low value (zero) as *angry* or *sad* and viceversa.
In Table 1, despite fond (”cherished with great affection” from the Merriam-Webster dictionary) has a

Table 1: Affective vectors of words in Sentence 1.

	Happy	Sad	Angry	Surprise	Fear	Disgust
food	7	0	0	0	0	0
excellent	9	0	0	0	0	0
great	10	0	0	0	0	0
variety	5	0	0	2	0	0
especially	5	0	0	0	0	0
fond	0	6	0	0	4	0
sweets	8	0	0	0	0	0
Phrase 1	44	6	0	2	4	0

positive emotional valence, the algorithm associates to term a negative affective vector. This does not happen for other words, whose the affective vectors are correctly estimated. The sum of happy and surprise (46) is greater than the sum of other components (12), thus the sentence globally returns a positive opinion.

Table 2: Affective vectors of words in Sentence 2.

	Happy	Sad	Angry	Surprise	Fear	Disgust
beware	0	6	0	0	4	0
dangerous	0	0	0	0	8	0
bad	0	0	0	0	7	0
beach	9	0	0	0	0	0
disgusting	0	0	0	0	9	0
terrible	0	6	0	0	9	0
pasta	5	0	0	0	0	0
Phrase	14	0	0	0	24	10

In Table 2, the algorithm associates to the adjective *terrible* a high score of *fear* index. Since it refers to pasta we would associate it with disgusted; fear would be appropriate in the case of a terrible percept in the presence of a bleak scene. The disambiguation of the true sentiment in response to the context is a limit of the present work and will be investigated in.

6 CONCLUSIONS

Nowaday, for the enterprise, it is important gathering a large amount of customer opinion from the web 2.0 tools. In this paper we illustrate an original approach to polarize (positive or negative) opinion based on Ekman indexes to evaluate emotional value of review about product/service. In our opinion, these indexex, allows to better capture the emotional state of customers about purchase.

The approach, based on the affective value of each single word, produces good results on documents medium-large dimensions but often fails on individual sentence. For resolving the question, in further works, we'll try to understand the semantics of the contents of a sentence using a knowledge base of common sense.

For use these results in the emotional recognition it is necessary focuses on concepts rather than on terms, in order to relate the words to emotional states through a conceptual representation. To this end we'll consider a conceptual semantics.

Another improvement to characterize and to structure emotional concepts, it is to build a detailed taxonomy from basic emotions (fear, anger, joy, disgust, sadness). This can help more easy to capture the emotional sense of a generic word.

REFERENCES

Berry, M. and Castellanos, M., editors (2007). *Survey of Text Mining II: Clustering, Classification, and Retrieval*. Springer.

Boucouvalas, A. and Zhe, X. (2002). Text-to emotion engine for real time internet communication. In *Proceedings of International Symposium on CSNDSP*, pages 164–168, UK. Staffordshire University.

Consoli, D., Diamantini, C., and Potena, D. (2008). Improving the customer intelligence with customer enterprise customer model. In *Proceedings of the 10th International Conference on Enterprise Information Systems*, pages 323–326, Barcelona, Spain.

Ekman, P. (2007). *Emotions Revealed: Recognizing Faces and Feelings to Improve Communication and Emotional Life*. OWL Books, NY.

Esuli, A. and Sebastiani, F. (2006). SENTIWORDNET: A publicly available lexical resource for opinion mining. In *Proceedings of LREC-06, 5th Conference on Language Resources and Evaluation*, pages 417–422, Genova, IT.

Khashman, A. (2008). A modified backpropagation learning algorithm with added emo-tional coefficients. In *Transaction on neural networks*, volume 19 of *11*, pages 1896–1909. IEEE.

Liu, H., Lieberman, H., and Selker, T. (2003). A model of textual affect sensing using real-world knowledge. In *IUI '03: Proceedings of the 8th international conference on Intelligent user interfaces*, pages 125–132, New York, NY, USA. ACM.

Okša, G. and Vajteršic, M. (2001). Systolic block-jacobi svd algorithm for processor meshes. 185:211–235.

Picard, R. W. (1997). *Affective computing*. MIT Press, Cambridge, MA, USA.

Strapparava, C. and Valitutti, A. (2004). Wordnet-affect: An affective extension of wordnet. In *The 4th International Conference On Language Resources and Evaluation*, pages 1083–1086.

Valitutti, A., Strapparava, C., and Stock, O. (2004). Developing affective lexical resources. 2(1):61–83.