Sentiment-Oriented Contextual Advertising

Teng-Kai Fan, Chia-Hui Chang

Department of Computer Science and Information Engineering,
National Central University, Chung-Li, Taiwan 320, ROC
tengkaifan@gmail.com, chia@csie.ncu.edu.tw

Abstract. Web advertising (Online advertising), a form of advertising that uses the World Wide Web to attract customers, has become one of the world's most important marketing channels. This paper addresses the mechanism of Content-Oriented advertising (Contextual advertising), which refers to the assignment of relevant ads within the content of a generic web page, e.g. blogs. As blogs become a platform for expressing personal opinion, they naturally contain various kinds of expressions, including both facts and comments of both a positive and negative nature. In this paper, we propose the utilization of sentiment detection to improve Web-based contextual advertising. The proposed **SOCA** (Sentiment-Oriented Contextual Advertising) framework aims to combine contextual advertising matching with sentiment analysis to select ads that are related to the positive (and neutral) aspects of a blog and rank them according to their relevance. We experimentally validate our approach using a set of data that includes both real ads and actual blog pages. The results clearly indicate that our proposed method can effectively identify those ads that are positively correlated with the given blog pages.

Keywords: Web Advertising, Sentiment Detection, Marketing.

1 Introduction

With the rapid growth and popularization of the World Wide Web, it has become one of the most essential media channels for advertising. According to the IAB[1] (Interactive Advertising Bureau), Internet advertising revenues exceeded 5.2 billion USD for the third quarter of 2007, representing yet another historic quarterly record and a 25.3 percent gain over Q3 2006. The statistics clearly show that advertisers are increasingly using the Web since it not only produces good consumer interaction, but is also highly flexible in terms of both geography and time. Furthermore, multiple forms of Web advertising are available, including plain text, video, and e-mail (including spam). Generally, for text-based ads, there are two main categories [1], [2]:

Sponsored Search (*Keyword Targeted Marketing*): an advertiser bids a reasonable price using certain keywords or phrases in order to appear at a certain position in lists of advertisers. Then, a list of ranked ads are triggered by user's search keyword (query) and placed on the result pages from a search engine [17].

[1] http://www.iab.net/

Content-Oriented Advertising (*Contextual Advertising*): This uses an intermediate ad matching system that parses the content of a page and returns those ads that are most relevant to currently-viewed pages, either through ads placements or pop-ups. For instance, if a user is browsing a web page about mobile phones, ads related to the content of this page may be selected by Google's content-based advertising system[2] (Google AdSense). These ads can be arbitrarily displayed in predefined positions, as shown in Figure 1.

Some prior studies have suggested that strong relevance increases the number of click-throughs [4], [16]. However, existing content-oriented advertising strategies only attempt to match relevant ads to a given web page, which we refer to using the term *correlation*, but they neglect to distinguish *positive* and *negative* correlations between ads and the pages on which they are placed. For example, as shown in Figure 1, an ad agency system might place the three most relevant ads about mobile phone on the top of a web page. However, the content of this page may describe reasons why someone should consider not using a mobile phone. We hypothesize that such ads that conflict with the negative orientation of the page are less likely to trigger click-throughs. In other words, even if an ad is related to the content of a triggering page, the ad agency system should avoid placing ads on pages which discuss the product/service in a negative light. Moreover, it is likely that blog owners themselves may not be happy about advertisements that directly conflict with their opinions.

Fig. 1. Correlation ads conflicting with blog content.

Hence, in this paper, we proposed an ad matching mechanism, which we refer to as Sentiment-Oriented Contextual Advertising (**SOCA**), based on sentiment detection to associate ads with blog pages. Instead of traditional placement of relevant ads, SOCA emphasizes that the ad agency's system should provide relevant ads that are related to the positive (and neutral) aspects of the page in order to best attract consumers. To evaluate our proposed method, we used a real-word collection comprising ads and blog pages respectively from Google AdSense and Blog-Search Engine[3]. The results suggested that our proposed method can effectively match relevant ads to a given blog page.

The rest of this paper is organized as follows: Section 2 provides background information on current on-line advertising and sentiment detection. Section 3

[2] https://www.google.com/adsense/
[3] http://blogsearch.google.com

introduces our methodology. The experimental evaluations are presented in Section 4. Section 5 outlines some related work. Finally, we present conclusions and future directions in Section 6.

2 Background

In this section, we briefly describe current on-line content-oriented advertising and review the main concepts in opinion detection and sentiment classification.

2.1 Content-Oriented Advertising

Without a loss of generality, in content-oriented advertising, an ad generally features a *title*, a text-based *abstract*, and a *hyperlink*. For example, as shown in Figure 1, the first ad is selected by Google, the title is "**AT&T Wireless**," usually depicted in bold or a colorful font, the abstract is "Large Selection of Free Cell Phones visit AT&T Wireless Official Site." The latter is generally concise due to space limitations. The hyperlink is "www.att.com/wireless," which links to an ad web page, known as the *landing page*.

In general, a functional estimation is to calculate the number of click-throughs or judge whether a user's activity is consistent with the relevant ad guidelines. The most widespread model for contextual ads is **CPC** (Cost Per Click), also known as **PPC** (Pay Per Click). Advertisers pay every time a user clicks on their ads and is redirected to their web site. They do not actually pay for the ads, but instead they pay only when the ads are clicked. A number of studies have suggested that strong relevance definitely increases the number of ad clicks [4], [16]. Hence, in this study, we similarly assume that the probability of a click for a given ad and page is determined by the ad's relevance score with respect to the page. For simplicity, we ignored the positional effect of ad placement and pricing models, as in [1], [2], [9], [13].

2.2 Sentiment Detection

Recently, TREC[4] (**T**ext **Re**trieval **C**onference) developed a blog task (called TREC-BLOG) that focuses on information retrieval from blog documents. A core task is opinion retrieval that focuses on a specific aspect of blogs: the opinionated nature of the writing (e.g., products, movies and political candidates). The opinion retrieval task involves locating blog posts that express an opinion about a given target. It can be summarized as "*What do people think about* <***target***>." The **target** not only can be a named entity (e.g., name of a person, or organization) but also a concept (such as a type of technology), or an event. Another main task is *polarity opinion retrieval* (sentiment classification) that aims to determine the polarity of the opinions in the retrieved documents/sentences, namely whether the opinions are positive, negative or mixed.

3 SOCA Framework

In our sentiment-oriented contextual advertising framework (SOCA), the advertising system processes the content of the page, detects sentiment, and then searches the ad collection to find the best matching ads. Thus, we focused on how to construct a scoring function to determine the relationship between the ads and a blog page. Given a page $\boldsymbol{p}$, which we called a *triggering page*, and a set of ads $\boldsymbol{A}$, the task is to select ads $a_i \in \boldsymbol{A}$ related to the content of a page. We designed three processes to

[4] http://trec.nist.gov/

assign the relevant ads to a given page, as shown in Figure 2. The sentiment detection mechanism is adopted to detect the sentiment of the blog's contents. We subsequently removed those sentences that reflect negative sentiments. Then, due to limitations in information about the ads and a certain page, we selected some specific terms in the triggering page as a set of seed words that were then used for vocabulary expansion to enhance the likelihood of intersection with available ads. A linear combination function using the traditional cosine similarity and an ontology mapping function was deployed for the page-ad matching strategy to rank the ads.

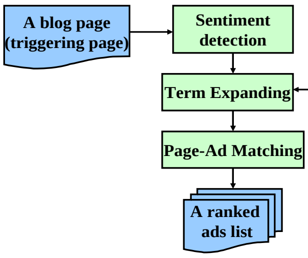

Fig. 2. The SOCA Framework.

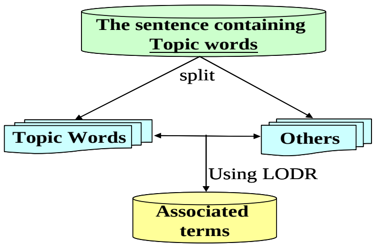

Fig. 3. Process for generating associated terms.

3.1 Sentiment Detection

Our aim is to apply a sentiment technique for recognizing the intent of a blog page. In this study, we regarded a sentence as the minimum unit and we tried to identify opinion-bearing sentences, and classify their sentiment into positive and negative semantic categories using machine learning algorithm and simple linear models.

The first sentiment detection method uses the SVM algorithm with different feature sets, namely uni-gram and opinion-bearing words (pre-selected opinion-bearing words from [5]). We adopted the LibSVM developed by Lin *et al*. [3]. To implement this learning algorithm on our dataset, we used the standard bag-of-features framework. Let $\{f_1,\ldots,f_m\}$ be a predefined set of m feature that can appear in a document. Let $w_i(d)$ be the weight of f_i that occurs in document $\boldsymbol{d}$. Then, each document $\boldsymbol{d}$ is represented by the document vector. $\vec{d} = (w_1(d), w_2(d),\ldots, w_m(d))$. As for the weighting value, here we used the *tf-idf* value. Besides, to build an efficient learning model, we divided the task of detecting the sentiment of a sentence into two steps, each of which belongs to a binary classification problem. The first is an identification step that aims to identify whether the sentence is subjective or objective. The second is a classification step that classifies the subjective sentences as positive or negative.

We applied a dictionary from [5] to assign each word a weight in either a positive, negative or objective direction for designing our second sentiment approach, as follows:

$$\text{Model 1: } \prod_{w=1}^{n} (\text{the sign of a word}) \tag{1}$$

Model 1 simply considers the polarities of the sentiments. The intuition here is based on English grammar like "double negation". For example, consider a positive sentiment, "I will never be unmoved whenever I see the film." Although this sentence contains two negative words *never* and *unmoved*, it will be reversed into a positive sentiment by the above formula. As for Model 2, it mainly assesses the strength of sentence $sentence_{strength}$; hence, we can sum the strength of each word and the $sentence_{strength}$ can then be normalized by its length, as follows:

$$\text{Model 2}: \sum_{w=1}^{n} (\text{the strength of a word}) / \text{the length of sentence} \tag{2}$$

In model 2, we defined a threshold ε (default value of 0.35) to determine the sentence categories according to iterative experiments.

3.2 Term Expansion

In general, a blog page can be about any theme while the advertisements are concise in nature. Hence, the intersections of terms between ads and pages are very low. If we only consider the existing terms included in a triggering page, an ad agency may not accurately retrieve relevant ads, even when an ad is related to a page. According to [1], considering the ads' abstracts and titles is not sufficient to perform page-ad matching. Thus, a term expansion of the keywords in the triggering page as well as the ads is conducted to increase the overlap. For a triggering page, because not all the words included in a page are useful to carry out term expansion, we simply took the terms tagged as nouns (NN & NNS) as candidate terms from which we generate a set of seed terms according to the following rules.

$T_{Capitalization}$: Whether a candidate term is capitalized is an indication of it being a proper noun, or an important word.

$T_{hypertext}$: Whether a candidate term is part of the anchor text of a hypertext link.

T_{title}: Whether a candidate term is part of the post's title.

$T_{frequency}$: Consistent with term frequency, we considered the three most frequently occurring candidate terms as a subset of the seed terms.

Subsequently, the set of seed terms ($Seed_{Term} = T_{Capitalization} \cup T_{hypertext} \cup T_{title} \cup T_{frequency}$) undergoes three term expanding methods.

For the first method, we submitted each seed term to WordNet[5] to acquire its synonyms. For instance, a noun *car* has the synonym *automobile*. However, since many product names and acronyms are not covered by WordNet, we introduced another open collaborative dictionary (Wikipedia[6]) to further expand the term. For example, the term *Nokia* has no corresponding information stored in WordNet; on the other hand, Wikipedia contains an entry page about this term. By ranking the words (excluding stop-words) in the entry page based on frequencies, we can select the top five terms to be part of our list of expanded terms. In this example, those terms would be *mobile*, *phone*, *network*, *company*, and *telecommunication*.

Co-occurrence in the linguistic sense can be interpreted as an indicator of semantic proximity. Certain words often appear together with other words and phrases. We assumed that the Web documents ***D*** similar to the triggering page would share common topics ***t***. Then by inspecting the terms in these documents we can construct a co-occurrence list for topic ***t***. Hence, the final term expansion technique we applied is a web-based method. For efficiency, we simply took the $T_{title} \cup T_{frequency}$ to represent the blog's topics ***t***. For each topic, we used this to retrieve the top three ranked documents ***D*** from the search engine. Then, we adopted the LODR (Logarithm of the Odds Ratio) formula to recognize topic-related words in ***D***, as shown in Figure 3. We first obtained a sentence set containing the topic words of ***D***. Then, the

[5] http://wordnet.princeton.edu/
[6] http://en.wikipedia.org

sentence set is split into two subsets comprising topic words and other words, respectively. Finally, the LODR formula is used to generate associated terms that occur frequently together with a topic word, as follows:

$$LODR(w_i, T) = \log \frac{p/(1-p)}{q/(1-q)} = \log \frac{p(1-q)}{q(1-p)} \tag{3}$$

where, T is the topic word, w_i is any word in a set of other words. Let p be the probability that a word, w_i, co-occurs with any of the topic words. That is,

$$p = P_r\,(w_i \mid \text{sentence containing topic words}) \tag{4}$$

Also, let q be the probability that w_i co-occurs with non-topic words. The formula for the occurrence probability of the word w_i with non-topic words is shown below:

$$q = P_r\,(w_i \mid \text{sentence excluding topic words}) \tag{5}$$

If LODR (w_i, T) > δ (where, δ is a threshold), then w_i is considered an associated term that topic words co-occur with.

3.3 Page-Ad Matching

We can regard the sentiment-oriented contextual advertising issue as a traditional information retrieval problem, that is, given a user's query ***q***, the IR system returns relevant documents ***d*** according to the query content. Hence, we modeled a triggering page ***p*** and relevant ads ***a*** with a user's query ***q*** and corresponding documents ***d***, respectively. In general, the most frequently used data representation in text mining is the bag-of-words approach. Thus, for our data representation, we used the vector space model (VSM), which is a way of representing documents through the words that they contain. Pages and advertisements are represented as weights in an n-dimension space. Let $w_{i,p}$ be a weight associated with a term t_i on a page ***p*** and let $w_{i,a}$ be a weight associated with a term t_i in an ad a_j. Then, the page vector $\vec{p}$ is defined as $\vec{p} = \{w_{1,p}, w_{2,p},\ldots,w_{t,p}\}$ and the vector for an ad a_j is defined as $\vec{a} = \{w_{1,j}, w_{2,a},\ldots,w_{t,a}\}$. Moreover, the weight of each vector can be assigned either a boolean value or a *tf-idf* (term frequency – inverse document frequency) value. Here we adopted the *tf-idf* weight. The vector model evaluates the degree of similarity between two documents in terms of the correlation between two vectors. Hence, the ranking of the page ***p*** with regard to the ad ***a*** is computed by the cosine similarity function, that is, the cosine of the angle between the vector $\vec{p}$ and the vector $\vec{a}$:

$$Sim_{Cos}\,(a_j, p) = \frac{\vec{a}_j \bullet \vec{p}}{|\vec{a}_j| \times |\vec{p}|} = \frac{\sum_{i=1}^{t} w_{i,a} \times w_{i,p}}{\sqrt{\sum_{i=1}^{t} w^2{}_{i,a} \times w^2{}_{i,p}}} \tag{6}$$

Generally, the goal of ontology mapping is to compute the domain of two ontologies which are semantically related at a conceptual level. Hence, in addition to the cosine similarity score, we also considered the terms ontology mapping between a page and an ad to estimate the degree of similarity. Since the page and the advertisement may match many categories, it is difficult to design a common ontology that is appropriate in every case. Therefore, we simply took the WordNet that contains general terms as an intermediate ontology to map the nouns included in the page and in the ad. Our proposed mapping function mainly considers the relationships and distances between terms. The relationship can be divided into symmetric and asymmetric relationships. An asymmetric relationship Rel_{asym} is one whose source and target synonym sets have lineages with a definite divergence point.

The common parent index is the index of the node in the relationship that represents this divergence point. For example, in finding a hypernym relationship between *dog* and *cat*, the relationship is dog -> canine -> carnivore; cat -> feline -> carnivore. Both the ancestry of "dog" and "cat" diverge at "carnivore," so the common parent index is 2; and the distance between them is 4, so we used the inverse distance for the Rel_{asym} value. For a symmetric relationship Rel_{sym}, an example of a relationship would be the synonym. For example, the term *gondola* is one of the synonyms for a *car*, which occurs at a depth of four consistent with the conceptual hierarchy. The distance is 4 and the inverse depth is chosen for the Rel_{sym} value. Given a page-ad pair, the degree of similarity of term ontology mapping is calculated through:

$$Sim_{Onto}(a,p) = \sum_{i=1}^{n}\sum_{j=1}^{m} \alpha Rel_{sym}(t_i,t_j) + (1-\alpha) Rel_{asym}(t_i,t_j) \quad (7)$$

where, t_i and t_j is any noun included in the page and the ad respectively, and the parameter α (default value of 0.2) determines the relative weight of the symmetric and asymmetric relationships.

According to above the description of our similarity formula, we formally defined the relevance score of an ad and a page as a linear combination of the cosine function score and ontology mapping score:

$$Score(a,p) = \beta Sim_{Cos} + (1\text{-}\beta) Sim_{Onto} \quad (8)$$

where the parameter β determines the relative importance of the cosine similarity and ontology mapping score.

4 Experimental Results

In this section, we focus on our two experiments, namely, sentiment detection and page-ad matching. We begin by describing the dataset and metrics used, and we then proceed to a discussion of the experimental results.

4.1 Datasets

To evaluate sentiment detection performance, we collected data from *epinions.com* and used this as our training dataset for building a learning classifier model. The dataset from *epinions.com* expresses positive and negative user experiences on specific fields of web pages. Hence, we adopted the labeling method proposed in [8] to automatically mark the positive, negative and neutral sentences. We examined 32,304 reviews, with an average number of sentences in per review of 29. For page-ad matching, because of the lack of large-scale ad databases, we first chose general topic words as query terms to request web pages from search engine. About 10,000 pages were retrieved and we placed these pages on an ad-crawler platform to obtain the corresponding ads assigned by Google AdSense. Our ad-crawler was similar to a generic blog website that can be embedded in a JavaScript module and we totally collected 104,094 ads. Unfortunately, we were unable to obtain hidden information about the ads, such as the ad keywords. Hence, for each ad and its *landing page*, we first generated a set of seed terms consistent with the rules discussed in Section 3.2; we then submitted them to keyword suggestion toolkits (e.g., Google AdWords[7]) to obtain the top 5 most popular terms and thereby approximate the likely advertiser keywords.

[7] https://adwords.google.com/select/KeywordToolExternal

Our triggering page collection comprised 150 blog pages on various topics. It included a range of opinions and comprised various subjective articles (100 positive (neutral) and 50 negative articles). We selected triggering pages according to the ratio of positive and negative sentences, that is, if the ratio of positive to negative sentence was over 4:1, we regarded a triggering page as positive, and vice versa. In our experiments, we tried not to restrict triggering to specific domains. To acquire the POS tag, we adopted the GENIA Tagger developed by the University of Tokyo[8]. In addition, we preprocessed the full text of a triggering page with its expanded terms and the ads (including ad's abstract and its title) with their expanded terms by removing stop words and one-character words, followed by stemming [12].

4.2 Evaluation of Sentiment Detection

The goals of this section was similar to [8]: the first is to explore how well the positive and negative detection model with different approaches on the data collected from *epinions.com*; and the second is to investigate how well the trained model performed on a different data source (150 triggering pages).

For the *epinions.com* data, we compared various methods as described in Section 3.1 for sentiment detection. We adopted a ten-fold cross validation mechanism for learning algorithm. Table 1 shows subjective sentence identification results. The baseline system assigned all sentences as opinions and achieved a precision of 31.6%. Our results show that better precision (64.40%) and recall (65.25%) performance appeared using SVM with unigram features.

Table 1. Subjective sentence identification results.

Approach	Precision	Recall	F-measure
SVM with unigram	**64.40 %**	**65.25 %**	**64.66%**
SVM with opinion-bearing words	63.83 %	65.18 %	64.50 %
Baseline	31.60 %		

For the sentiment classification experiment, we used the results of sentiment identification as input. The goal of this experiment was to classify the subjective sentence into a suitable class (i.e., positive or negative). Since our two linear models are not classified into two steps in the task of sentiment detection, the meaning of final results in linear models is identical to that of sentiment classification.

Table 2. Positive and negative sentence classification results.

Approach	Precision	Recall	F-measure
SVM with unigram	**52.03%**	56.11%	54.0%
SVM with opinion-bearing words	51.20%	**57.81%**	**54.3%**
Linear model 1	39.10%	29.52%	33.64%
Linear model 2	32.90%	32.53%	32.71%
Baseline	42.1%		

Table 2 shows results for sentiment classification experiment. The baseline system determined that all sentences featured negative opinions and achieved 42.1% precision. The results clearly show that better precision (52.03%) and recall (57.81%) performance are respectively produced by the SVM with unigram and by an SVM with opinion-bearing words. Moreover, the SVM approach outperformed the baseline system around 10% in terms of precision. However, the results indicate that the F-measures of 33.64 % and 32.71 % are respectively produced by linear model 1 and

[8] http://www-tsujii.is.s.u-tokyo.ac.jp/GENIA/tagger/

linear model 2. Although the two linear models produced unsatisfactory results with performance of only about 30%, they still approximated the baseline system.

According to the above results, it seems reasonable to infer that the SVM with unigram feature model has no significant effect to one which uses opinion-bearing words. Furthermore, using an SVM algorithm to detect sentiment seems more appropriate than using linear models.

For the dataset of triggering pages, owing to a lack of training data, we chose SVM with unigram features to train a model on the *epinion.com* dataset. We then applied this model to our triggering page dataset. In order to get a reasonable evaluation, human experts had to annotate the entire triggering page set. The subjective sentences contained 976 positive and 458 negative sentences. We conducted similar sentiment identification and classification step and the results are shown in Table 3. On average, we respectively achieved an F-measure of about 65% and 58% in sentiment identification and classification step. Although the triggering pages covered various topics (few intersected with *epinions.com*'*s*), the results are consistent with the epinions.com data. It seems reasonable to conclude that our trained sentiment classifier models are not limited to a few specific domains.

Table 3. System results for triggering pages.

Detection Task	Precision	Recall	F-measure
Identification	64.76%	66.10%	65.42%
Classification	56.17%	62.12%	58.99%

4.3 Evaluation of Page-Ad Matching

The goal of this section is to investigate to what extent the ad placements are actually related to the positive (and neutral) aspects of the triggering pages. To evaluate our page-ad matching framework, we compared the top-10 ranked ads provided by three different ranking methods, namely our proposed approach with sentiment detection (here we used the classifier learned from SVM with unigram features as sentiment detection model), our proposed approach without sentiment detection, and Google AdSense. No more than 30 ads were retrieved and inserted into a pool for each triggering page. All the advertisements in each pool were manually judged by experts. By comparing with Google AdSense, we only measured the accuracy for this experiment. The experts regarded an ad related to the positive (neutral) aspects of the triggering pages as a target. Table 4 shows the results for three page-ad matching methods, it indicated that the proposed approach with sentiment mechanism can yield best performance (74.1%) than other approaches (60.0% for our method without sentiment detection and 52.5% for Google AdSense). According to Table 4, these results lead us to the conclusion that our sentiment-oriented contextual advertising framework can place ads that are related to the positive (and neutral) content of triggering pages.

Table 4. Accuracy of page-ad matching.

Method	Accuracy
With Sentiment Detection	**74.1%**
Without Sentiment Detection	60.0%
Google AdSense	52.5%

The goal of our next experiment was to explore our SOCA in detail. We selected the top 10 ranked ads provided by three matching mechanisms, namely cosine, ontology function and a combination of cosine and ontology (Cos+Onto). We thereby ensured that no more than 30 ads would be retrieved and inserted into a pool for that triggering page. All the advertisements in each pool were manually judged by experts. To quantify the precision of our results, we applied an 11-point averaged figure. Since it is quite difficult to evaluate our entire ad collection, our recall values were only relative to the set of judged advertisements. The results are shown in Figure 4. Each data point corresponds to the precision value calculated at a certain percentage of recall. The results clearly indicated that using the Cos+Onto with default β = 0.8 can achieve better performance than the use of cosine (β = 1) and ontology (β = 0) alone.

In addition to the precision-recall curve, we also used another presentation involving two quality measures (*Precision@K* and mean average precision (MAP)) to assess matching results. The results are displayed in Figure 5. It is clear that the cosine and ontology approach can generate MAP of around 39% and 29%, respectively; besides, improved performance (of around 43%) can be produced by Cos+Onto. As shown in these figures, the results based on cosine similarity are apparently very powerful and consistently superior than using ontology alone. However, the combination of cosine and ontology does offer a slightly positive effect.

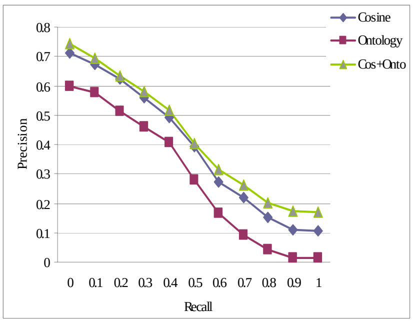

Fig. 4. Precision-Recall 11-point curve.

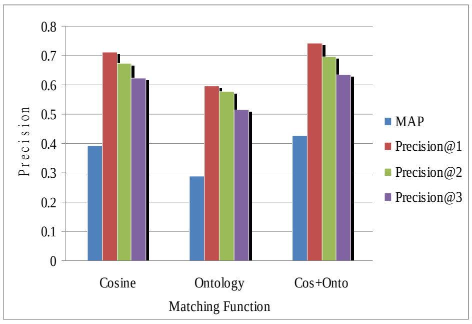

Fig. 5. The performance of the three matching strategies

5 Related Work

Several studies pertaining to advertising research have stressed the importance of relevant associations for consumers [10] and how irrelevant ads can turn off users and relevant ads are more likely to be clicked [4]. They show that advertisements that are presented to users who are not interested can result in customer annoyance. Thus, in order to be effective, the authors conclude that advertisements should be relevant to a consumer's interests at the time of exposure. As for contextual advertising, Ribeiro-Neto *et al*. [13] proposed a number of strategies for matching pages to ads based on extracted keywords. The winning strategy required the bid phrase to appear on the page, and then ranked all such ads using the cosine of the union of all the ad sections and the page vectors. In follow-up research [9], the authors proposed a method to learn the impact of individual features using genetic programming to generate a matching function. The function is represented as a tree comprised of arithmetic operators and functions as internal nodes, and different numerical features of the query and ad terms as leaves. The results show that genetic programming can identify matching functions with significantly improved performance compared to the best method proposed in [13]. However, due to the vagaries of phrase extraction, and the lack of context, approaches based on “bid phrases” leads to many irrelevant ads. To overcome this problem, Andrei *et al*. [2] proposed a system for contextual ad matching based on a combination of the keyword (syntactic) and classification (semantic) score. The keyword score is defined as the cosine of the angle between the page and the ad vectors. For the semantic score, it relies on the classification of pages and ads into about 6000 nodes within a commercial advertising taxonomy to determine topical distance. Even if several prior studies have proved that the relevance has a definite impact on contextual advertising and on the proposed effective ranking function that matches ads with pages, they neglect sentiment in the assignment of relevant ads. In this study, in addition to considering general syntactic (keyword) and semantic (ontology) matching, we further investigate the importance of sentiment analysis for improved contextual advertising.

Sentiment classification has been pursued in multiple ways. While most researchers use a supervised approach [14], [18], others use an unsupervised approach [15], [20]. They can be classified into three different levels: words, sentences and documents [19]. Hatzivassiloglou and McKeown [6] described an unsupervised learning method for identifying positively and negatively oriented adjectives with an accuracy of over 90%. Turney [15] showed that it is possible to use only a few of those semantically oriented words (namely, “excellent” and “poor”) to label other phrases co-occurring with them as positive or negative. These phrases were subsequently used to automatically separate positive and negative movie and product reviews, with accuracies of 66-84%. Pang *et al*. [11] adopted supervised machined learning with words and n-grams as features to predict orientation at the document level; they achieved up to 83% precision. Kim and Hovy [8] presented a system that, given a topic, automatically finds those people who hold opinions about this topic and the sentiment of each opinion. Their system used a sentiment words list and WordNet to classify the opinions at the word and sentence level. In their subsequent research [7], they focused on the identification of pro and con reasons in online reviews. Lexical (n-grams), positional and opinion-bearing word feature sets were used in a maximum entropy model to extract pros and cons from review sites. In this paper, we adopted Kim’s concept to quickly build a sentiment detection model. The major difference between Kim’s model and ours is that we adopted the SVM learning algorithm because it is a well-known learning method that is widely used for classification and regression. Besides, we further compared the effects by using learning algorithm with different feature sets and linear models.

6 Conclusion

In this study, we proposed and evaluated a novel framework for associating ads with blog pages based on sentiment analysis. Prior work to date has only examined the extent of content relevance between pages and ads. In this paper, we investigated the sentiments of blog pages and utilized this information to demonstrate sentiment-oriented contextual advertising. For sentiment detection, we compared machine learning-based algorithm with different feature sets and two linear models. Besides, we used the *epinions.com* data source for training. The results showed that using SVM can outperform linear models around 10 %. The best performance on sentiment detection may be as much as 54% in terms of F-measures. As for page-ad matching, we evaluated our framework using 150 blog pages and over 100,000 ads sampled from Google AdSense. First, we compared SOCA with Google AdSense and found that our proposed method with sentiment detection can achieve superior performance (74.1% precision). To analyze our SOCA in detail, we evaluated three matching strategies (i.e., cosine similarity, ontology similarity and the combined approach). The results indicated that the three proposed approaches can assign relevant ads to the positive (and neutral) aspects of a blog page. Besides, the combined approach has a better performance than cosine and ontology independently.

In the future, we intend to conduct a more comprehensive analysis of our model and explore the effectiveness of sentiment detection using different machine learning algorithms (such as HMM and CRF). We may also apply the concept of topic analysis to pages and ads to enhance their performance.

References

[1] Anagnostopoulous, A., Broder, A.Z., Gabrilovich, E., Josifovski, V., Riedei, L.: Just-in-Time Contextual Advertising. In: 16th conference on CIKM, pp. 331--340. ACM (2007)

[2] Broder, A., Fontoura, M., Josifovski, V., Riedel, L.: A semantic approach to contextual advertising. In: 30th conference on SIGIR, pp. 559--566. ACM Press (2007)

[3] C.-C. Chang and C.-J. Lin, LIBSVM : a library for support vector machines, 2001.

[4] Chatterjee, P., Hoffman, D.L., Noval, T.P.: Modeling the Clickstream: Implications for Web-Based Advertising Efforts. Marketing Science 22, pp. 520--541. (2003)

[5] Esuli, A., Sebastiani, F.: SENTIWORDNET: A Publicly Available Lexical Resource for Opinion Mining. In: 5th Conference on Language Resources and Evaluation, pp. 520--541. (2006)

[6] Hatzivassiloglou, V., Mckeown, K.R.: Predicting the Semantic Orientation of Adjectives. In: 35th conference on ACL, pp. 174--181. ACM Press, New York (1997)

[7] Kim, S.-M., Hovy, E.: Determining the sentiment of opinions. In: 20th international conference on COLING, pp. 1367--1373. (2004)

[8] Kim, S.-M., Hovy, E.: Automatic Identification of Pro and Con Reasons in Online Reviews. In: Proceedings of the COLING/ACL on Main conference poster sessions, pp. 483-490, ACM Press, New York (2006)

[9] Lacerda, A., Cristo, M., Gonçalves, M.A., Fan, W.g., Ziviani, N., Ribeiro-Neto, B.: Learning to advertise. In: 29th conference on SIGIR, pp 549--556. ACM Press (2006)

[10] Langheinrich, M., Nakamura, A., Abe, N., Kamba, T., Koseki, Y.: Unintrusive customization techniques for Web advertising. J. Computer and Telecommunications Networking 31, 1259--1272. (1999)

[11] Pang, B., Lee, L., Vaithyanathan, S.: Thumbs up? Sentiment Classification Using Machine Learning Techniques. In: Conference on EMNLP, pp 79--86. (2002)

[12] Porter, M.F.: An algorithm for suffix stripping. Program: electronic library & information systems, Vol. 40, pp. 211--218. (2006)

[13] Ribeiro-Neto, B., Cristo, M., Golgher, P.B., Moura, E.S.d.: Impedance coupling in content-targeted advertising. In: 28th conference on SIGIR, pp. 496-503. ACM (2005)

[14] Riloff, E., Wiebe, J.: Learning Extraction Patterns for Subjective Expressions. In: Proceedings of the 2003 conference on EMNLP, Vol. 10, pp. 105--112. (2003)

[15] Turney, P.: Thumbs up or thumbs down? Semantic orientation applied to unsupervised classification of reviews. In: 40th Annual Meeting of the ACL, pp. 417--424. (2002)

[16] Wang, C., Zhang, P., Choi, R., D'Eredita, M.: Understanding consumers attitude toward advertising. In: 8th Americas Conference on Information Systems, pp. 1143-1148. (2002)

[17] Weideman, M., Haig-Smith, T.: An investigation into search engines as a form of targeted advert delivery. In: Proceedings of the 2002 annual research conference of the South African institute of computer scientists and information technologists on Enablement through technology, pp. 258--258. (2002)

[18] Wiebe, J., Riloff, E.: Creating Subjectives and Objective Sentence Classifiers from Unannotated Texts. CICLing (2005)

[19] Wilson, T., Wiebe, J., Hoffmann, P.: Recognizing Contextual Polarity in Phrase-Level Sentiment Analysis. In: Proceedings of the conference on HLT/ EMNLP, pp. 347--354. (2005)

[20] Yu, H., Hatzivassiloglou, V.: Towards answering opinion questions: separating facts from opinions and identifying the polarity of opinion sentences. In: Proceedings of the 2003 conference on EMNLP, pp. 129--136. ACM Press, New York (2003)