

Effective Blog Advertising by Understanding Blogger's Emotions & Needs

Wen-Hsiang Lu

Dept. of Computer Science & Information Engineering
National Cheng Kung University, Taiwan, R.O.C
whlu@mail.ncku.edu.tw

Yao-Sheng Chang

Dept. of Computer Science & Information Engineering
National Cheng Kung University, Taiwan, R.O.C
ys.chang1976@gmail.com

ABSTRACT

Recently, more and more advertising systems have been developed to deliver contextual ads on blog articles. To deliver contextual ads, most existing advertising systems conventionally adopt simple keyword-matching-based advertising methods, however, which may not be effective to match appropriate ads against affective blog articles without considering bloggers' emotions and needs. Different from informative blog articles, affective blog articles mostly describe bloggers' daily life events (activities), corresponding to emotions (feelings), and their personal emotional needs. Thus, in this paper, we first aim to investigate the interesting issue of what bloggers' emotions and needs respond to their daily life events by making comprehensive analyses on a collection of affective blog articles. Second, we propose a new event-driven emotion-need-based advertising model to effectively deliver ads to meet bloggers' emotional needs. The extensive experimental results show that our proposed event-driven emotion-need-based advertising (EENA) model can obtain better performance than conventional keyword-match-based advertising methods.

Categories and Subject Descriptors

H.3.3 [Information Search and Retrieval]: Information Search and Retrieval – User Need, User Emotion

General Terms

Documentation, Economics, Languages

Keywords

Blogger, Need, Emotion, Affective blog article, Online advertising

1. INTRODUCTION

1.1 Motivation

The fast development of Web is resulting in great growth of online advertising. At the same time, online advertising has become the major part of optimizing revenues for many free Web services in recent years [9]. Thus, how to design efficient ways to deliver online ads to consumers while they surf the Web has become an important issue. More and more advertising systems have been developed by Web service providers to display contextual ads in the visited pages or blogs to attract visitors to click on while they visit the pages/blogs. Generally, most existing advertising systems assume that topic-relevant ads are suitably shown up on the visited pages, and conventionally adopt simple keyword-matching-based advertising methods to recommend ads by matching some advertiser-bidder topic keywords against the visited pages. Consequently, these Web service providers still think that keyword-matching-based advertising methods are effective to apply to blog advertising. Therefore, Yih et al. [10] systemically analyzed the relevance between bidding phrase and page content by picking up some dominant keywords from pages.

As for vocabulary impedance problem of Ads [11], the authors proposed the idea of expanding the ad and the page with some related terms to reduce the vocabulary impedance. In fact, all of the above works have not considered more influential and reasonable features for affective blog articles Ads, e.g., emotions and emotional needs. According to the perspective of Ni et al. [4], blog articles can be roughly divided into two genres: the first genre is topic-oriented and related to bloggers' hobbies or interests, etc. Thus, this genre of articles like web pages is called informative blog articles. The second one is like online diary and considered as affective blog articles, which describe bloggers' daily life events, emotions, and their individual emotional needs. Recently, many researchers [12, 13] aim to understand bloggers' daily lives and emotions, and proposed various methods to automatically recognize or classify bloggers' emotions.

From a preliminary analysis on our collected 115,551 blog articles, bloggers usually share in blog articles about their life events which happened on their daily lives, and further express their emotions and even consumption needs. For examples, some bloggers post articles to complain when they were jobless, or to express how sad they feel after failures in love. Although keyword-matching-based advertising methods might be useful to match topic-relevant ads to informative blog articles, these methods are usually weak to deliver appropriate ads to affective blog articles without consideration of bloggers' emotions and needs. For the instance in Figure 1, the affective blog article described that the blogger originally want to go out to a bar on Sunday night while all bars are closed early, thus he hate weekends. And then he changed his mind to look for good songs. As a result, five unrelated contextual ads are delivered to the article from the advertising system (Google AdSense), including two horoscope services, one accommodation information and one event happening list in London, and an opera in Vienna.

Ads by Google

Your 2009 Horoscope Free
Free predictions, lucky numbers and success secrets for 2009.
www.astro.com/astrofree

Visiting London?
Why Not Stay In A Self-Serviced Apartment In Central London?
www.citybaseapartments.co

Horoscope & Future 2009 ?
Free horoscope year 2009 birthdate required. You'll be surprised!
www.astro.com/horosc

What to See in Vienna?
Opera, Gala Concerts, Mozart, Waltz Free Anniversary CD Each Ticket!
Vienna Tourism Tickets Online

What's happening in London
Events happening in London and the UK. Comprehensive event listings.
www.whatshappening.co.uk

● ● ● **sunday afternoon**

it's sunday. i **hate weekends**. well, it's nice to have the day off but everything is closed! stupid germany. oberhausen is d e a d. nothing's happening. maybe i should go out to a bar tonight? uuuu no, they all close early, because everybody has to work monday morning. well, not everybody. i work monday night only. that's nice.

well i guess i'll keep **looking for good songs** to sing at our next karaoke night in two weeks.

so long!
Kai

Figure 1. An example of blog with contextual advertisements

According to the report in [4], more than 85% of blog articles belong to affective articles, but most of them are less attractive in

the blogosphere (less than 100 visitors per day). After posting affective articles, most bloggers would still expect some comments from other visitors. Basically, there are two kinds of the real-time users in blog, including writers and readers. According to our investigation, most writers may be the only visitors who usually surf their own blogs. In addition, the conventional advertising systems take both writers (bloggers) and readers (visitors) as the same advertising targets. However, according to [14], they pointed out that both writers and readers have different emotions. Based on some available evidences from the study in advertising [3], emotional advertising is more and more important to meet the needs of goal-oriented consumers than rational advertising. Thus, we think this is a good opportunity to meet bloggers' needs by delivering appropriate ads based on the life events happened to them and their response of emotions and needs. In this work, we particularly expand a new type of immediate online advertising on affective blog articles while taking into account of bloggers as the advertising targets rather than visitors.

Previous studies in advertising [1, 2] also showed that understanding the attitude of customers will make ads acutely stimulate consumption. Ducoffe argued that vast amount of advertising exposures reach consumers when they have no consumption need [2], i.e., most ads are simply irrelevant to consumer concerns at the time of exposure. Therefore, the importance of an advertising strategy that caters to consumers' needs has emerged [1]. It is clear that more effective strategies for advertising of affective blog articles are required. For that, it is necessary to take into account bloggers' emotions associated with their corresponding consumption needs.

In this paper, we first point out the importance of understanding bloggers' emotions and the corresponding emotional needs toward affective blog advertising. Also, we address the problem of matching suitable ads against affective blog articles.

1.2 Challenges and Objectives

As described before, in the past few years, huge amounts of blog articles are dramatically increased. To effectively solve the problems of delivering appropriate ads to affective blog articles, we propose a new idea to take advantages of bloggers' emotions and needs to recommend bloggers proper ads. Based on the idea, we carefully develop a feasible framework (see Section 3 for details). However, a number of challenges in implementing this framework will be suffered and described below.

1. How to detect affective blog article from any given blog article.
2. How to detect the terms of bloggers' life event, emotions and needs from the unstructured text data in a given affective blog article.
3. How to deliver appropriate ads to an affective blog article.

In this initial work, we first focus on the third challenge to investigate the feasibility of our new idea and will propose effective methods to tackle the other two challenges in the future. Especially, we aim to address two interesting issues in this paper. First, we try to provide advertisers more clear realization about bloggers' emotions and needs by making comprehensive analyses from a collection of blog articles. Second, we propose a novel event-driven emotion-need-based advertising model to effectively deliver ads to really meet emotional needs of bloggers.

2. IDEA AND OBSERVATION

2.1 Main Idea

In this paper, we addressed the problem of how to effectively match suitable ads against an affective blog article. As shown in Figure 1, conventional keyword-matching-based advertising methods didn't take bloggers' emotions and needs into account, and thus delivered five unrelated contextual ads to the article. To alleviate the above problem, we propose a new idea to make good use of a few significant implicit features in affective blog articles, such as emotions and needs. Figure 2 shows the difference between our proposed method (path 2) and conventional methods (path 1). In the next subsection, we will illustrate some interesting observations and statistical analyses towards further realization of influence of emotions and needs on affective blog advertising. We hopefully expect that these implicit features would be useful to enhance the performance of recommending suitable ads to affective blog articles.

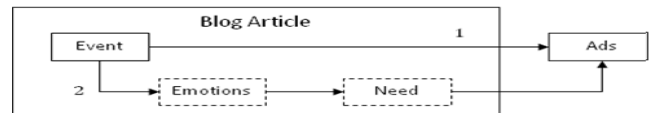


Figure 2. A novel idea of utilizing implicit features of emotions and needs to benefit blog advertising

2.2 Observations

Ni et al. [4] reported that affective articles occupy 85% of a collection of 99,059 blog articles from MSN space. We also made a statistical analysis about the distribution of affective articles on 115,551 blog articles collected from several popular Chinese blog spaces, and found similar results that 93% of the collected blog articles contain emotional terms. This statistics show the evidence that affective articles still play a major part in blog space.

Unlike the definition of "event" in the research fields of Topic Detection and Tracking [5] and New Event Detection [6]. In this paper, we focus on the daily life events (activities) in affective blog articles which usually happen to bloggers everyday, such as work, travel, chat, date, etc. To understand which life events more often appear in affective blog articles, we made a preliminary statistical analysis on the 115,551 collected blog articles by utilizing a part-of-speech tagging tool, the CKIP tagger¹, to assign POS (part-of-speech) labels for each blog article. In our initial work, The POS labels with active intransitive verbs (means action verbs) are taken as possible life event terms. Further, we are developing an efficient method to enhance the quality of event detection. Table 1 shows the distribution of five frequent life events with their occurrence count on the collected articles.

After understanding the frequent life events occurred with bloggers, we also want to investigate the most frequent emotional terms in the collected blog data based on the collected emotional term list (section 4.1.1). Table 2 shows a few high-frequency emotional terms and their corresponding occurrence count and coverage rate (the ratio of the occurrence count of an emotional term versus total number of articles).

Also, to further investigate which emotional terms and need terms co-occurred frequently with the frequent life events, we conduct some analysis for each of the top five frequent life events about a few frequent positive and negative emotional terms, and

¹ <http://ckipvr.iis.sinica.edu.tw/>

need terms on randomly selected 30 blog articles. Some interesting analysis about association among frequent life events and their corresponding emotions and needs is shown in Table 3. According to our observation, user needs usually can be represented as common nouns or actions (e.g., travel, shopping or drink).

Table 1. The distribution of five frequent life events on the 115,551 collected blog articles.

Rank	Emotional Terms	Count	Coverage Rate
1	喜歡 (enjoy)	21670	0.188
3	開心 (joyful)	14574	0.126
7	快樂 (happy)	8580	0.074
14	擔心 (worry)	5868	0.050
27	難過 (sad)	4013	0.035
33	害怕 (afraid)	3579	0.030

Table 2. The coverage rates of a few high-frequency emotional terms on the 115,551 collected blog articles.

Frequent Life Events	Count
回家 (back home)	11021
出門 (leave home)	5464
休息 (take a break)	4944
上課 (attend class)	4832
上班 (go to work)	4430

3. THE PROPOSED METHOD

In order to deal with the problem of delivering suitable ads to affective blog articles, we are the first to propose an event-driven emotion-need-based advertising model, which will be explained in detail in Section 3.2. However, we will face many difficulties since our novel idea of recommending ads to bloggers based on their emotions and needs expressed in their affective blog articles is a very complex and challenge task. Thus, we carefully develop an event-driven emotion-need-based advertising framework shown in Figure 3.

3.1 Emotion-Need-Based Advertising System Framework

There are two major parts in our event-driven emotion-need-based system framework.

3.1.1 Training Part

In the training part, we divide it into two major parts, including: (1) semi-automatically label events, emotions, and

Table 3. The analysis of emotions and needs on the randomly selected 30 blog articles for five frequent life events.

Frequent Life Events	Frequent Emotions Terms		Frequent Needs
	Positive	Negative	
回家(go home)	開心(joyful)、喜歡(enjoy)、可愛(cute)	害怕(afraid)、小心(careful)、不好意思(embarrassing)	感情(affection)、麻將(mahjong)、回台灣(return to Taiwan)
出門(leave home)	喜歡(enjoy)、可愛(cute)、開心(joyful)	害怕(afraid)、可惜(regretted)、懷疑(doubt)	治安(public security)、餐廳(restaurant)、環島(around the Island)
上課(attend class)	開心(joyful)、喜歡(enjoy)、可愛(cute)	可怕(fearful)、無聊(bored)、恐怖(terror)	旅行(travel)、學分(credit points)、功課(homework)
上班(go to work)	喜歡(enjoy)、希望(hope)、有趣(funny)	擔心(worry)、討厭(dislike)、無聊(bored)	旅遊(travel)、書籍(book)、換工作(change job)
休息(take a break)	喜歡(enjoy)、不錯(not bad)、開心(joyful)	緊張(nervous)、擔心(worry)、可惜(regretted)	旅行(travel)、演唱會(concert)、餐廳(restaurant)

needs for each training affective blog article, and then train a emotion model and a need model; (2) manually label an ad relevance score for each match ad to a blog article, and then train an ads model.

3.1.2 Testing Part

In the testing part, for each testing affective blog article, there are two major steps. In this initial work, the testing events adopt life events “生日” (birthday), “分手” (break up), “結婚” (get marry). In the future, we will develop automatic methods for the extraction of events, emotions, and needs. Second, the system utilizes a trained event-driven emotion-need advertising model to pick up the most matching ads from the large collection of ads for

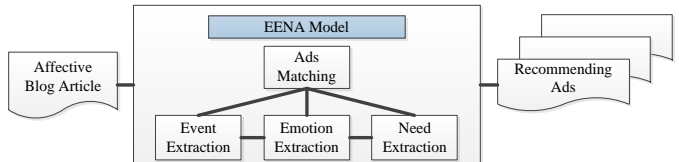


Figure 3. The framework of our proposed event-driven emotion-need-based advertising system.

the testing article.

3.2 Problem Formulation

3.2.1 Event-Driven Emotion-Need-Based Advertising Model

DEFINITION 1. Assume that each blog article b contains a life event e , and some implicit emotions terms $m_i \in M$, and needs $n_j \in N$, i.e., a blog article can be represented as a triple $b = (e, m_i, n_j)$, where $M = \{m_1, m_2, \dots, m_k\}$ is the emotion set, and $N = \{n_1, n_2, \dots, n_h\}$ is the need set.

The problem we addressed in this paper is that given an affective blog article b and an advertisement set A , some appropriate ads $a \in A$ can be recommended to the blogger who posted the article. According to our observation described in Section 2, different bloggers experience the same life events, but might evoke different emotions and needs. In addition, a blogger possibly expresses herself/himself with more than one emotional term for an event. Thus, we deal with ads recommendation by introducing emotion as an implicit feature to model a multiple relationships between an event and emotions as Equation (1).

$$P(a|e) = \sum_{m_i \in M} P(a, m_i|e) = \sum_{m_i \in M} P(m_i|e)P(a|e, m_i) \quad (1)$$

where $P(m_i|e)$ is called the *emotion model*, which is the probability of the blogger’s emotion m_i responding to a given event e . In Equation (1), $P(a|e, m_i)$ would be unreliable without the consideration of user needs. To meet blogger’s real needs

while recommending ads, we consider the blogger’s need n_j as a latent feature. Thus, Equation (2) can be derived as

$$P(a|e, m_i) = \sum_{n_j \in N} P(n_j | e, m_i) P(a|e, m_i, n_j), \quad (2)$$

where $P(n_j|e, m_i)$ is called the *need model*, and $P(a|e, m_i, n_j)$ is the *advertising model*.

Finally, Equations (1) and (2) are combined to obtain Equation (3) as follows:

$$P(a|e) = \sum_{m_i \in M} P(m_i|e) \sum_{n_j \in N} P(n_j|e, m_i) P(a|e, m_i, n_j). \quad (3)$$

We will describe in the following how to estimate the parameters for each submodel.

3.2.2 Parameter Estimation

● Emotion Model

To estimate $P(m_i|e)$, we exploit the association information of life event e and emotion m_i in the collection of affective blog articles.

$$P(m_i|e) = \frac{BCount(e, m_i)}{BCount(e)}, \quad (4)$$

where $BCount(x)$ means the times of term x occurring in the collection of affective blog articles.

● Need Model

In this initial work, we haven’t developed an automatic method to extract needs from affective blog articles, thus adopted manually labeling data in the training stage. $P(n_j|e, m_i)$ is computed using Equation (5).

$$P(n_j|e, m_i) = \frac{BCount(e, m_i, n_j)}{BCount(e, m_i)}, \quad (5)$$

where $BCount(x)$ is the same as the above definition.

● Advertising Model

To estimate $P(a|e, m_i, n_j)$ from our collection of ads, we exploited a simple smoothing approach to mitigate the zero probability problem. $P(a|e, m_i, n_j)$ is computed as follows:

$$P(a|e, m_i, n_j) = \begin{cases} P(a|e, m_i, n_j) & \text{if } ACount(e, m_i, n_j) > 0 \\ P(a|m_i, n_j) & \text{elseif } ACount(m_i, n_j) > 0, \\ P(a|n_j) & \text{otherwise } ACount(n_j) > 0 \end{cases} \quad (6)$$

where $ACount(x)$ means the frequency of term x occurring in the collection of ads.

4. EXPERIMENTS

We conducted several experiments towards two major goals in this work. First, we will show that making good use of emotions corresponding to life events in a affective blog article can really infer a blogger’s needs. Second, we show that our event-driven emotion-need-based advertising model can effectively recommend appropriate ads and outperforms the conventional keyword-matching-based advertising method.

4.1 Experimental Setup

4.1.1 Data Set

Blog Articles

In this work, we randomly collect 115,551 blog articles from a Chinese blog host² in Taiwan, while filtering out some too short or non-Chinese articles.

Ads

² <http://www.pixnet.net/>

Kijiji³ is a popular advertising platform containing 14 categories of ads covering seven subareas in Taiwan. A lot of stores post abundant ads to the platform. Thus, we employ a crawling program to collect over 94,000 ads and filter out some pornographic ads, and then about 61,424 ads are taken as our ad corpus in this work.

Emotional Terms

We collected 458 Chinese emotion words from a Chinese website⁴, and then extended them with an additional 2,248 emotion words using Chinese Synonyms Thesaurus⁵ via these 458 basic emotion words. After filtering a number of nonsense words, the rest 1,216 emotion words are regarded as the emotional term list.

To enhance the reliability of our emotion model (Equation (4)), we classify 1216 emotional terms into seven categories by referring to Chinese medical theory⁶, including “happy” (喜), “angry” (怒), “worry” (憂), “think” (思), “sad” (悲), “scare” (恐) and “frighten” (驚). For instance, the terms “開心” (joyful), “喜歡” (enjoy) and “可愛” (cute) are classified to the category happy, “可怕” (fearful) and “恐怖” (terrible) belong to scare category.

4.1.2 Training Data

In order to remove the influence of non-affective (informative) articles, in this work, we in advance filter out non-affective articles by adopting the method of detecting affective article proposed by Ni et al. [4]. The method used collocations in articles as features for training a SVM classifier [8], which can effectively separate blog articles into affective articles and informative articles. We finally obtained 107,462 affective articles from the total 115,551 articles. Then we randomly selected 100 articles as dataset respectively for the three life events “生日” (birthday), “分手” (break up), “結婚” (get marry). We use the evaluation technique of 5-fold cross validation which divides the whole dataset into 80 percent of training set and 20 percent of testing set. Three volunteers are invited to manually label suitable life events, emotions and needs for the selected training articles. For each ad, they give relevance score for the relationship between ads and blog articles. They decide the primary need after reading a blog article, and then judge manually whether the ad candidates picked based on the primary need are related to the blog article or not. The most related ad gained two points; on the contrary the most unrelated ad gained zero point. For example, a blogger wrote an article which is related to birthday, but his real need is to get a bike as his birthday present. In this situation, the birthday cake ad will gain zero point, and the bike for sale ad will gain two points.

4.1.3 Testing Data

Since we use 5-fold cross validation evaluation, we use the 20 percent of dataset as a testing data sets and will have 5 testing sets.

4.1.4 Evaluation Metrics and Baselines

Metrics

The problem of recommending a suitable ad can be seen as a classic IR problem, given a blog article as a query, we should find the most relevant ads. Therefore, in this work, we utilize the classic IR metrics, including Precision, Recall, F-measure, and top-k inclusion rate to evaluate our proposed models.

³ <http://taipei.kijiji.com.tw/>

⁴ <http://blog.lanyue.com/view/79/1479975.htm>

⁵ <http://www.chinese-tools.com/tools/synonyms.html>

⁶ http://city.udn.com/61613/3484022?tpno=-1&cate_no=0

Baseline

- **Need Inference:** Take blogger’s need inference as classification problem, thus each need is considered as a class. We adopt the SVM classifier as our baseline, with bag of words as features.
- **Ad Matching:** In order to validate the efficiency of our EENA model, we use conventional keyword-matching-based advertising method as a baseline. We choose event terms as keywords to match suitable ads from the collected ad corpus.

4.2 Need Inference: Performance of the Proposed Event-Driven Emotion-Need Model

In this section, we conduct several experiments to determine the effectiveness of inferring bloggers’ needs using our proposed event-driven emotion-need model. Table 4 shows that the inclusion rate of correct needs for the top 20 of recommended needs are over 50% in all testing events using our proposed model. Besides, Table 5 shows that our event-driven emotion-need model actually outperforms the SVM baseline on all three testing events. It is promising that our event-driven emotion-need model is actually effective to predict bloggers’ needs.

After we examined a few positive examples, the correct needs are usually recommended at the top three ranks. For the first example in Table 7, the article describe that a flood disaster happened in the blogger’s birthday and our model can predict correct result of his need “賑災” (disaster relief) at the first rank. Another example (#3 in Table 7) also shows that our system can predict the need “戒指” (ring) for the event “結婚” (get marry) at the second rank. These two positive examples show again that our event-driven emotion-need model is robust.

4.3 Ad Matching: Performance of the

Table 4. The top-*k* inclusion rate and precision of Need Inference by our event-driven emotion-need model

Life Event	Top k Result	Inclusion Rate	Precision
生日 (birthday)	TOP 1	0.3191	0.3191
	TOP 5	0.4680	0.0978
	TOP 20	0.6170	0.0329
分手 (break up)	TOP 1	0.2745	0.2745
	TOP 5	0.5098	0.1215
	TOP 20	0.5294	0.0313
結婚 (get marry)	TOP 1	0.4523	0.4523
	TOP 5	0.4761	0.1333
	TOP 20	0.5476	0.0404

Table 5. The Need inference comparison of our EENA model and SVM (baseline) with top-1 precision

Life Event	Need Model	SVM
生日(birthday)	0.3191	0.229
分手(break up)	0.2745	0.278
結婚(get marry)	0.4050	0.323

Proposed Event-Driven Emotion-Need-Based Advertising Model

To determine the effectiveness of our proposed event-driven emotion-need advertising model, we conducted the following experiments to compare our model with conventional keyword-matching-based advertising methods. Table 6 shows the

experimental results. From the table, our proposed EENA model performs better than the baseline for most testing life event. Furthermore, in the event “get marry”, the top-1 inclusion rate of our EENA model outperforms the baseline by 14.96% (0.2095 vs 0.0600). However, the precision of the life event “birthday” is lower than that of baseline. After our analysis, need for birthday is too diverse to lead to good results. Besides, the number of training data is not enough and thus make the recall value is lower than baseline.

The ads recommended by our system can satisfy the blogger’s need in a blog article. For example, in Table 7, the #1 blog article which is related to the event “birthday”, the blogger mentioned that there was a flood disaster in her birthday. Therefore, the blogger didn’t want any cakes or birthday parties to celebrate. She felt sad and wanted to help the victims who were suffering from the flood disaster. The predicted top one need is “災區” (disaster area) and the top one ads that our system recommended, in which is about the activities of rescuing victims from the flood calamity. Another example, #2 (Table 7) is related to the event “break up”. A blogger wants to die after broke up with her boyfriend. The predicted need “諮商” (counseling) has included at the third rank and our model thus offered suitable ads about a solidarity and making friends.

5. ONGOING WORK – LIFE EVENT

Table 6. The Ads matching performance comparison of our EENA model and keyword-matching-based advertising method (baseline)

		Inclusion Rate		F-measure	
		EENA	Baseline	EENA	Baseline
生日 (birthday)	TOP 1	0.0590	0.0654	0.0037	0.0089
	TOP 5	0.1196	0.1969	0.0143	0.0309
	TOP 20	0.3259	0.4893	0.0518	0.0940
分手 (break up)	TOP 1	0.3429	0.1385	0.0667	0.0456
	TOP 5	0.3429	0.2483	0.0512	0.0667
	TOP 10	0.8143	0.5269	0.1708	0.1136
結婚 (get marry)	TOP 1	0.2095	0.0600	0.0196	0.0236
	TOP 5	0.3552	0.2406	0.0488	0.0549
	TOP 20	0.6702	0.4839	0.0930	0.0776

EXTRACTION MODEL

In Section 2.2, we manually select the life event terms from blogger’s affective articles. However, manually labeling life events is time-consuming with rapidly growing affective articles on the web. To automatically extract life events, we are developing a life event extraction model. Furthermore, the model will be integrated into the proposed EENA model and enhance the quality of recommending Ads.

According to HowNet (<http://www.keenage.com/>) and Kuo’s work [16], a life event is defined as several kinds of combinations of POS tags, like NV, VN, NVN and so on. Also, we found that verb is the most important constituent of life events. Furthermore, to extract completely semantic life events, we proposed a life event extraction model, which utilizes two effective features. The first feature is E_{vn} which represents a VN pair consisting of a hint verb e_v and a hint noun e_n , i.e., $e = (e_v, e_n)$, and the other is $E_{distance}$ which considers the affection of the distance between e_v and e_n . Since Conditional Random Fields (CRF) [15] have great flexibility to include a wide variety of arbitrary, non-independent features, we adopt a CRF model to maximize the probability of life event e given an affective blog article B comprised of a set of paragraphs $\{p_1, \dots, p_j, \dots, p_{|B|}\}$. We assume that e_v and e_n of a life

Table 7. The correct examples of recommended ads

#	Life Event	Blog Article (partial translation)	Top-3 Predicted Needs	Top-1 Ad (partial translation)
1	生日 (birthday)	... 8 8 節是父親節, 也是我的生日, 但也是 8 8 水災的日子... 只希望政府趕快救助南部災區的民眾... (...father's day is my birthday, too, but there was a flood disaster happened in the same day)	災區 (disaster area) 蛋糕 (cake) 禮物 (gift)	..., 為此基金會成員不斷的開會..., 就是希望南下災區時讓災區的學校有所幫助... (a disaster relief ad about flood disaster in south of Taiwan)
2	分手 (break up)	..., 未神麼每次分手後妳討厭我好像罵句髒話我活在世上不是來給你糟蹋的如果真的是那我寧願死一死算了, ... (...why do you insult me every time we broke up...)	張老師 (Teacher Chang) ⁷ 情歌 (love song) 諮商 (counseling)	中壢市美滿服務中心~~婚友~聯誼~交友~婚姻~戀愛~愛情~ (a solidarity and making friends ad)
3	結婚 (get marry)	以為自己會比大家先結婚, ... 現在的我不但沒了男友, 還被那條不小心搞出來的人命追著跑... (I thought I would get married before the others, ... but now I not only lose boyfriend, but also with a baby...)	男朋友 (boyfriend) 戒指 (ring) 小孩 (child)	出租貼心男友 (a boyfriend rental ad)

event e must appear in a paragraph simultaneously. Thus, any candidate e is selected from the event candidate set $\bigcup_{i=1}^n E_i$, where $E_i = V \times N$ and $V = \{v | \text{any verb } v \text{ in } p_j\}$, $N = \{n | \text{any noun } n \text{ in } p_j\}$. The model is formulated as follows:

$$P(e|B) \cong Z(B) \times \exp(\sum_{m=1}^2 \beta_m w_m(e, B)), \quad (7)$$

where $Z(B)$ is a normalization function and β_m is the weight corresponding to a feature function $w_m(e, B)$. The details of individual feature function will be introduced in the following.

- *Hint verb and hint noun* ($f_1 = E_{vn}$, without considering the order of the tags)

Initially, we consider a paragraph has only one event. To avoid the incorrect VN pair, e_v and e_n have to appear in the same paragraph p_j . The first feature function is as follows:

$$w_1(e, B) = \sum_{j=1}^{|B|} \delta(e, p_j), \quad (8)$$

$$\text{where } \delta(e, p_j) = \begin{cases} 1, & \text{if } e \text{ appears in } p_j \\ 0, & \text{otherwise.} \end{cases} \quad (9)$$

- *VN distance* ($f_2 = E_{distance}$)

Although e_v and e_n are limited to appear in the same paragraph, the smaller the distance of e_v and e_n is, the more semantically suitable the life event e is. The second feature function is defined as

$$w_2(e, B) = f_{vn-distance}(e) = \frac{1}{\sqrt{2\pi}\sigma} \exp[-\frac{1}{2\sigma^2}(d - \mu)^2], \quad (10)$$

where $e = (e_v, e_n)$ appearing in the paragraph p_j and d is the distance between e_v and e_n , μ is the distance mean, and σ is the standard deviation of distance.

6. CONCLUSIONS AND FUTURE WORK

In this paper, we propose a new idea to take advantages of bloggers' emotions and needs to recommend bloggers proper ads. Based on the idea, we carefully propose an event-driven emotion-need-based advertising model and develop a feasible framework to solve problems of conventional keyword-matching-based advertising approach which often recommends unsuitable ads. In the experiments of need inference, our event-driven emotion-need model outperforms the SVM. Besides, for the evaluation of ads matching, our proposed event-driven emotion-need-based advertising model performs better than the keyword-matching-based approaches. Also, another contribution of this paper is that we have made a comprehensive analysis of author's emotions and needs towards affective blog advertising.

7. REFERENCES

- [1] C. Wang, P. Zhang, R. Choi, and M. D. Eredita. Understanding Consumers Attitude Toward Advertising. In Eighth Americas Conference on Information Systems, 2002
- [2] R.H. Ducoffe. How consumers assess the value of advertising. Journal of Current Issues and Research in Advertising, 1996, pp.1-18.
- [3] A. Mehta and S. C. Purvis. Reconsidering Recall and Emotion in Advertising. In Journal of Advertising Research, 2006.
- [4] X. Ni, G.R. Xue, X. Ling, Y. Yu, and Q. Yang. Exploring in the Weblog Space by Detecting Informative and Affective Articles. In WWW '07, 2007.
- [5] C. C. Chen and M. C. Chen. TSCAN: A Novel Method for Topic Summarization and Content Anatomy. In SIGIR'08, 2008.
- [6] G. Kumaran and J. Allan. Text Classification and Named Entities for New Event Detection. In SIGIR'04, 2004.
- [7] S. E. Robertson, S. Walker, S. Jones, M. Hancock-Beaulieu, and M. Gatford. Okapi at TREC-3. In TREC '94, 1994.
- [8] N. Cristianini and J. Shawe-Taylor. An Introduction to Support Vector Machines and other kernel-based learning methods. Cambridge University Press, 2000. ISBN 0-521-78019-5
- [9] Y. Zhu, G. Wang, J. Yang, D. Wang, J. Yan, J. Hu, and Z. Chen. Optimizing Search Engine Revenue in Sponsored Search. In SIGIR '09, 2009.
- [10] W. Yih, J. Goodman, and V. Carvalho. Finding Advertising Keywords on Web Pages. In WWW '06, 2006.
- [11] G. Leshed, and J. Kaye. Understanding How Bloggers Feel: Recognizing Affect in Blog Posts. In CHI'06, 2006.
- [12] P. Melville, W. Gryc, and R. D. Lawrence. Sentiment Analysis of Blogs by Combining Lexical Knowledge with Text Classification. In KDD'09, 2009.
- [13] K. H. Lin, C. Yang, and H. Chen. Emotion Classification of Online News Articles from the Reader's Perspective. In IEEE'08, 2008.
- [14] J. J. Kuo and H. H. Chen. Multi-Document Summary Generation Using Informative and Event Words. ACM Transactions on Asian Language Information Processing, 7(1), 3:1~3:23, 2008.
- [15] J. Lafferty, A. McCallum and F. Pereira. Conditional Random Fields: Probabilistic Models for Segmenting and Labeling Sequence Data. In ICML'01, 2001.

⁷ A free 24hour telephone counseling service in Taiwan.