A NOVEL KEYWORD SUGGESTION METHOD TO ACHIEVE COMPETITIVE ADVERTISING ON SEARCH ENGINES

Dandan Qiao, School of Economics and Management, Tsinghua University, Beijing, China, qiaodd.12@sem.tsinghua.edu.cn

Jin Zhang (corresponding author), School of Business, Renmin University of China, Beijing, China, zhangjin@rbs.org.cn

Abstract

Search engine advertising is a popular business model for online advertising and recently a new strategy (i.e. competitive advertising) is emerging. Competitive advertising is helpful for organizations to expand market shares from competitors, which is crucial to sustain competitive advantage. To achieve the goal of competitive advertising, appropriate and fruitful competitive keywords should be provided to advertisers. However, existing keywords suggestion methods usually recommend general business keywords based on co-occurrence analysis. They not only fail to enable competitive advertising, but also limit advertisers to a small number of hot keywords, causing high bidding costs. As a response, this study proposes a competitive keywords suggestion method based on query logs. It uses the indirect associations between keywords and the hidden topic information captured by query logs to recommend competitive keywords. Through the method, massive competitive keywords are mined out to help organizations achieve competitive advertising and simultaneously broaden the choices of keywords for search engine advertising. Experiments are conducted to demonstrate that the proposed method could have a good performance than other methods, proving that it can help organizations well achieve the goal of competitive advertising.

Keywords: Competitive advertising, Keyword suggestion, Topic modeling, Factor graph model, Search engine advertising, Query logs

1 INTRODUCTION

With the assistance of internet development, search engine advertising (SEA) has been a fast-growing advertising channel for organization development (Fuxman et al. 2008; Jansen et al. 2013). Advertisers usually bid on specific keywords to have advertisements displayed along the search results of these keywords. When a user submits a query of the bidden keyword, the corresponding advertisements would be presented to him (Chen et al. 2008). As the advertisements are content correlated with the query keyword, it's highly probable for users to be interested with the advertisements and click them. This is an effective way for organizations to target potential customers and expand their market shares. Therefore, organization managers are vigorously investing in SEA.

Recently a new trend has been witnessed for search engine advertising, namely competitive advertising. A typical example to illustrate the strategy is that BMW may buy the keyword "Audi" on search engines. When a user type "Audi" into the search engine bar, advertisements with respect to "BMW" would appear on the organic results page. Then BMW would probably obtain more shares by free riding on the market created by its competitor Audi. This competitive advertising strategy enables an organization to "poach" the competitor's consumers by advertising on competitive keywords directly. Then more market share from competitors could be seized for the organizations, which attach much importance to business spread in the competitive market. Besides, the competition between organizations has a big impact on their market performance (Chen & Miller 2012; Clark 2011). It's necessary to adopt a competitive strategy to grab customers from competitors' shares (Fudenberg & Tirole 2000; Sayedi et al. 2011), especially when they want to sustain a competitive advantage in the market. Therefore, leveraging keywords to achieve competitive advertising on search engines is very useful and crucial for an organization's market development.

To realize the strategy of competitive advertising on search engines, advertisers should be provided with a large number of competitive keywords. However, traditional keyword suggestion methods are typical to produce general keywords related with the organization business (Schwaighofer et al. 2009; Fuxman et al. 2008; Huang et al. 2003; Jones 2011; Joshi & Motwani 2006; Kelly et al. 2009; Wu et al. 2009). For example, the keyword "hotel" may be bidden by a hotel company to locate the consumers searching for the service. As a result, many advertisers fight for a limited number of popular keywords with high bidding prices, greatly increasing the advertising budgets. Besides, these business-relevant keywords couldn't effectively satisfy the organization needs for competitive advertising, which ought to be supported with competitive keywords (Wei et al. 2013). Therefore, it's necessary to design a competitive keyword suggestion method, which can not only expand the keyword choices for managers, but also enable the competitive strategy to rival with peers on search engines.

To make up the gap, this paper proposes a novel method named IAT for competitive keyword suggestion to help organizations achieve the competitive strategy on search engines. Firstly, the indirect associations between keywords hidden in query logs are employed to find enough candidate keywords. Secondly, the topic structure of these keywords is mined based on the query corpus and further used to identify competitive keywords through factor graph modeling. Data experiments are conducted across various business domains to demonstrate the effectiveness of IAT. And it's also compared with other keywords suggestion methods to present the advantageous performance of IAT. The remainder of this paper is organized as follows. Section 2 reviews the state-of-art of keyword suggestion methods. The framework of the proposed competitive keyword suggestion method IAT is presented in Section 3. Section 4 gives the corresponding algorithmic details. And section 5 shows the experimental results which reveal the outperformance of the proposed method. Finally this paper is concluded in Section 6.

2 LITERATURE REVIEW

The popularity of sponsored search engine advertising has motivated a plethora of research efforts developed to solve the problem of keyword suggestion. According to the type of data source, three streams of literature are mainly reviewed, namely query log based, proximity based and meta-tag crawlers based keywords suggestion methods (Abhishek & Hosanagar 2007).

2.1 Query Log based Keywords Suggestion

In the branch of query log based methods, keywords are mainly suggested by conducting association/co-occurrence analysis in search engine query logs (Schwaighofer et al. 2009). This co-occurring law has been widely used in many commercial keyword suggestion tools like Google AdWords[1] and Baidu Tuiguang[2]. The up-to-date property (Liu et al. 2012) and user intention based feature (Da et al. 2011) drive lots of studies to focus on query logs based analysis. For example, Bartz et al. (2006) used logistic regression and collaborative filtering models to recommend business-relevant terms which represent product features in the search logs. Fuxman et al. (2008) considered that strong associations exist between queries and the clicked URLs. They exploited such reinforcing relationships to mine queries that are related to the interests of the advertisers. Sarmento et al. (2009) proposed a keyword suggestion method by mining the logs of previous submitted ads to infer similarity relations among the associated keywords. To some extent, the existing query log based methods could recommend effective keywords to the advertisers. However, as they are based on co-occurrence analysis, it's probable that only a limited number of hot words are suggested to different advertisers (Szpektor et al. 2011). On the one hand, the bidding prices of these keywords are driven up, leading to a high advertising cost for merchants (Bartz et al. 2006). On the other hand, quite a few keywords in the long tail are ignored despite their high commercial value to the advertisers (Zhang et al. 2014). Recently, Wei et al. (2013) designed a notion of competitive keywords and found these competitive keywords based on query logs. Although providing a new perspective for keyword advertising, the keyword market needs to be further expanded to satisfy the vast needs from the advertisers.

2.2 Proximity based Keywords Suggestion

Proximity based keywords generation methods are to conduct similarity analysis based on query results with the seed keyword and append the seed with keywords found in the search results (Abhishek & Hosanagar 2007). By conducting similarity analysis between the seed keyword and keywords in advertisers' websites, Abhishek and Hosanagar (2007) proposed a method to recommend keywords with high relevance. Wu et al. (2009) employed the search results of the seed keyword to generate a large set of candidate terms and further filter out the irrelevant ones by leveraging the user relevance feedback information. Some researchers also built a system for keyword extraction from web pages (Joshi & Motwani 2006). To better understand the keyword meaning of the queries and boost the search engine ads, Broder et al. (2007) proposed a classification method of rare queries using the web search results to determine the topic of the given queries. In addition, some research is to calculate the proximity based on vocabulary dictionaries/corpus pre-constructed by domain experts (Chen et al. 2008; Amiri et al. 2008), e.g., thesaurus dictionary, Wikipedia, etc. The quality of the proximity based methods varies across the source of web pages or corpus. Besides, considering the quantity of the information on each webpage, it would take great computational efforts to conduct these proximity based methods. Furthermore, in many cases, due to the source quality (web pages, texts, and dictionaries), the suggested keywords cannot reflect search engine users' real intentions, which are the primary concern of advertisers.

---

[1] https://adwords.google.com/
[2] http://fengchao.baidu.com/

2.3 Meta-tag crawlers based Keywords Suggestion

The keyword suggestion methods based on meta-tag crawlers are to extract meta-tag words from highly ranked web pages that are returned by the search engines with the query of seed keywords. Some popular online tools like WordStream and Wordtracker use meta-tag crawlers to search meta-tag keywords and make suggestion of relevant keywords for advertisers. Though these techniques can suggest keywords from the meta-tag data of web pages, they have two aspects of limitations (Joshi & Motwani 2006). First, the number of relevant keywords generated by this kind of methods is still low and cannot provide sufficient choices for advertisers. Second, there is no guarantee to find good and relevant keywords, since the meta-tag keywords of web pages are sometimes too diverse to keep the focal meaning.

As discussed above, the existing three streams could to some extent suggest related keywords to advertisers. However, some obvious drawbacks like hot competition and ignoring user intentions limit their usage by advertisers. More importantly, they couldn't help organizations to achieve the strategy of competitive advertising. Therefore, a novel keyword suggestion method is proposed in this paper to make up the gap.

3 COMPETITIVE KEYWORDS SUGGESTION

This section introduces the framework of the proposed keyword suggestion method (i.e. IAT). As discussed above, IAT leverages query logs to mine competitive keywords for a given keyword, thus enabling the implementation of competitive advertising. To achieve the goal, IAT is designed to incorporate two steps: the first step is indirect association based analysis, and the second step is topic based competitive keywords suggestion.

3.1 Indirect Association Analysis

Supposing Q is a set of queries recording n query log data of the search engine within a period time. Each query q in Q consists of two elements, the query keywords $q.kw$ and the related number of search volume $q.vol$. Table 1 gives an example of the query logs about the keyword "Budweiser" within a day from Baidu[3].

q.kw	*q.vol*
The official website of Budweiser	70
Budweiser beer agent	50
Budweiser brewery	10
Budweiser wholesale	5
Prices of Budweiser	30
Budweiser beer degree	20
Budweiser beer recruitment	30
How about the taste of Budweiser?	10
…	…

Table 1. Examples of Query Log for "Budweiser" from Baidu

Definition 1 (volume of a keyword). Given a query log dataset Q within a certain time period, for any keyword i, its search volume in Q could be calculated as the aggregate volume of the queries that contain i.

$$i.vol = \sum_{q \in Q, i \in q.kw} q.vol \tag{1}$$

[3] http://tuiguang.baidu.com/

Definition 2 (associative keyword). Given a query log dataset *Q* within a certain time period and a keyword *i*, for any keyword *a*, it could be defined as an associative keyword of *i*, if there exists at least one piece of query *q* in *Q* satisfying $i\in q.kw$ and $a\in q.kw$ simultaneously. Obviously, there is surely more than one keyword associated with *i*. All the keywords occurring with *i* compose an associative keyword set, denoted by *i.AK*. Obviously the keyword *i* is also an associative keyword of the keyword *a*, indicating the symmetry of the associative keyword relationship.

$$i.AK = \{a \mid \exists q \in Q, i \in q.kw \wedge a \in q.kw\} \quad (2)$$

Traditional keywords suggestion methods usually recommend co-occurring keywords to advertisers. But we go further than that simple relationship. For any seed keyword *s*, the associative keyword set *s.AK* could to some extent be considered as a characteristic profile. It reflects topics that users are concerning about the seed keyword *s*. By extending the relationship a step further, it's easy to understand that the associative keywords sometimes play an intermediate role, and are associated with the seed keyword's competitors as well. These intermediate associative keywords feature the common topics that users discuss about the pair of competitive keywords, which are connected by an indirect association. Therefore, indirect associations in the query data provide an important clue to mine competitive keywords, which constitute the foundation for candidate generation in IAT.

Given the query logs *Q* and the seed keyword *s*, we can retrieve all its associative keywords, *s.AK*, by traversing *Q*. Further, for each element *a* in *s.AK*, it can also derive a set of associative keywords, denoted by *a.AK*. The second-order associative keywords in *a.AK* share common intermediates with the seed keyword *s*. Projected to search engines, they compete for the attention of a common group of users. Therefore, items in *a.AK* could be roughly considered as the candidate competitive keywords of the seed keyword *s*. The whole process of candidate generation could thus be modelled as follows,

$$\begin{aligned} &Find && Cand = \{c_1, c_2, \ldots, c_n\} \\ &s.t. && Cand \subseteq K \\ & && \forall c_x \in Cand, \exists a \in s.AK \cap c_x.AK \end{aligned} \quad (3)$$

Where the seed keyword is denoted by *s* and the set *K* represents all the keywords occurring in the query logs *Q*. For any seed keyword *s*, this model discovers a set of candidates, denoted by *s.Cand*.

3.2 Topic Based Competitive Keywords Suggestion

Given any seed keyword *s*, the indirect association analysis could derive a broad set of candidate competitive keywords, denoted by *s.Cand*. But there may be some noises existing in the candidate set, making it not effective to advertise directly. For a specific item represented by a keyword, there are various topics cared by consumers for the corresponding item. The associative keyword set right serves as a good description of the various topics. And practically competition between items is mainly driven by topic commonalities recognized by consumers. Therefore, if the topic structure is extracted, it would benefit us to further identify competitive keywords effectively. For this purpose, Latent Dirichlet Allocation (LDA) model (Blei et al. 2003; Blei 2012), which is an unsupervised machine learning technique to identify latent topic information from large document collection, would be adopted. LDA is a probabilistic generative model assuming that every document is a distribution over topics and every topic is a distribution over words. Let the associative keyword set of all the candidates constitute the training corpus. Thus, it can enable topic distillation through LDA modeling.

Formally, suppose that given a seed keyword *s*, we have obtained its candidate competitive keyword set, denoted by *s.Cand*. The seed *s* and these candidates make up a new keyword set, i.e $I = \{s\} \cup s.Cand$. Each keyword *i* in the set *I* owns an associative keyword set, *i.AK*, which is a characteristic file for *i*. Keywords in the characteristic profile are affiliated with a value of search volume, as defined in *Definition 1*. Let $C= \{i.AK|i\in I\}$ be the collection of associative keyword sets for topic modeling. According to the LDA framework, the associative profile *i.AK* of each keyword *i* is interpreted as a multinomial distribution $Mult(\theta)$ over a series of topics $T= \{t\}$. And each topic *t* is assigned a multinomial distribution $Mult(\varphi)$ over all the words $W=\{w\}$ in C. θ and φ represent two dirichlet distributions with hyper-parameters α and β respectively, denoted by $\theta \sim Dir(\theta|\alpha)$ and $\varphi \sim Dir(\varphi|\beta)$. For each word w in profile *i.AK*, sample a topic t from the multinomial distribution $Mult(\theta)$ specific with the keyword profile *i.AK*. And subsequently sample the observed word w from the multinomial distribution $Mult(\varphi)$ associated with the topic t. Therefore, the generation probability of each word w in the corpus C can be formulated as,

$$\begin{aligned} p(w \mid i.AK) &= p(w \mid t, \beta)\, p(t \mid i.AK, \alpha) \\ &= \int p(w \mid \varphi) Dir(\varphi \mid t, \beta) d\varphi \int p(t \mid \theta) Dir(\theta \mid i.AK, \alpha) d\theta \end{aligned} \tag{4}$$

Repeating the sampling process above for all the words in a keyword profile and then for all the keywords in I would finally give the observed corpus C, whose generation probability can be expressed as,

$$p(C) = \prod_{i \in I} \prod_{w \in i,AK} p(w \mid i.AK) \tag{5}$$

To estimate the parameters θ and φ as well as the latent variables t, the Gibbs sampling (Griffiths 2002), a fast and effective algorithm for approximate inference, is applied to infer the model. After parameter estimation, information about the latent topics of these keywords could be obtained. Each keyword i is projected into a topic distribution, denoted by $i=\{p_t\}$, where p_t stands for the probability that the particular keyword i belongs to the topic t. From the perspective of users, it means to what extent the item represented by i could compete with alternative items on the topic t.

The topic structures derived from LDA model indicate competition between items, which is crucial to identify effective competitive keywords. But we should also understand that peer items compete with each other in the market, which means the competition is contextual rather than being isolated. To model the interactions between competitive relationships, factor graph model would be adopted to combine with the topic structure information for further competitive keywords identification. As a popular method used in many applications (Colavolpe & Germi 2005; Kschischang et al. 2001; Loeliger 2004), factor graph model interprets how a global function of many variables factors into a product of several local functions. The factorization structure gives important information about statistical dependencies among these variables (Kschischang et al. 2001; Loeliger 2004). Therefore, mapping the identification of competitive keywords into a factor graph could model the interactions between item competitions. This is consistent with the natural competitive context, thus ensuring the effectiveness of the recommended keywords.

For the seed keyword s, candidate competitive keywords in *s.Cand* are mixed with noises. Therefore, we need to judge whether each candidate c constitutes an effective competitive relationship with the seed s. In the framework of factor graph model, the problem could be transformed to label all the relations like $<s, c>$ as a binary value, where "1" represents competitive and "0" otherwise. For each pair of keywords $<s,c>$, let $X = \{x_1, x_2 \ldots x_n\}$ be some features associated with the pair relation. These features can be defined heterogeneously in different contexts. For example, the volume of a keyword can be used as a good feature, which indicates how popular the corresponding keyword is among the users. Given the defined features, the labels of all the relations, denoted by $L = \{l_1, l_2 \ldots l_n\}$ can be modelled as a conditional probability, i.e., $P = \{L|X\}$. The relation between keywords is determined by several factors, which are functions over features X. Thus the joint probability of keyword relations could be modelled as,

$$p(L \mid X) = \prod_j F(l_j, X) H(l_j, L) \tag{6}$$

Where j is the relation index and l represents the label of the relation. The function $F(l_j,X)$ denotes the feature-specific factors, represented as the posterior probability of l_j given the feature vector X. The function $H(l_j,L)$ reflects the correlations between keyword relations. By learning the factor graph model, we could get a predictive function to infer the relationship between a pair of keywords. As the topic structure is a good indicator of item competitions, we could expand the factor graph model with LDA modeling process as follows,

$$p(L \mid X, Q) = \prod_j F(l_j, X) T(l_j, p_{j1}, p_{j2}) H(l_j, L) \quad (7)$$

Where Q represents the query logs and the added factor $T(l_j, p_{j1}, p_{j2})$ represents the posterior probability of l_j based on the topic structure of the two keywords composing the relation l_j. These factors could be formalized in different ways according to the needs. Here we use exponential-linear functions to define these factors as follows,

$$F(l_j, X) = \frac{1}{z_1} \exp\{\sum_m \mu_m f_m(l_j, X)\} \quad (8)$$

$$H(l_j, L) = \frac{1}{z_2} \exp\{\sigma h(l_j, L)\} \quad (9)$$

$$T(l_j, p_{j1}, p_{j2}) = \frac{1}{z_3} \exp\{rt(l_j, p_{j1}, p_{j2})\} \quad (10)$$

Where z_1, z_2, z_3 are normalization factors, $f_m(l_j, X)$ represents the m_{th} feature function, $h(l_j, L)$ denotes an indicator function to represent the correlations between keyword relationships and $t(l_j, p_{j1}, p_{j2})$ could be defined as the proximity of topic structure for the two keywords in the relation l_j. To learn the model, some labelled data is needed in advance. The learning process incorporates two steps. Use the Gibbs sampling (Griffiths 2002) to maximize $p(C|\alpha, \beta)$ in equation (5). Based on the results of LDA modeling, the next step is to maximize the joint probability $p(L|X, Q)$ through the sum-product algorithm. The two steps are iteratively computed until the joint probability in equation (7) converges. We could see that the topic structure mined from query logs is timely updated into the subsequent prediction of relationships. Finally the relations between each pair of keywords in the factor graph are labelled as competitive or non-competitive.

4 ALGORITHMIC DETAILS

Following the framework of IAT introduced in Section 3, a set of effective competitive keywords could be obtained for the seed keyword. To better present the whole process of the topic based competitive keyword recommendation method, the pseudo-code is given as ALGORITHM 1.

The whole process of the method IAT consists of two major steps. The first step is to generate candidates for the given seed keyword s. It incorporates two rings of co-occurrence analysis to find the associative keyword set in query logs. Therefore, the time complexity is mainly determined by two parameters, namely the length of query logs and the number of keywords associated with the seed keyword. For each associative keyword, it needs again to find the associative keywords in query logs Q. The process should be conducted query by query. Denote that the length of query logs Q is l and the number of associative keywords for the seed keyword is n. The time taken for the first step could be expressed as $O(ln)$.

In the second step, the method IAT attempts to use the hidden topic structure of the items to recommend competitive keywords. It transfers the problem of keywords mining into the prediction of relationships between items over a factor graph model. To complete such a task, it needs to train a predictive function which could be achieved by maximizing the joint probability in equation (7) iteratively. In each circle of the iterative process, it needs to conduct one round of LDA modeling for the whole query logs. Therefore, the time of LDA modeling is crucial to determine how long the second step lasts for. The LDA process is to repeatedly sampling the dimension and the keyword for the whole corpus. If denoting the average number of the associative keywords for an item as n, the size of the corpus is about n^2. When there are m hidden topics for these items, the time complexity of LDA modeling thus equals $O(mn^2)$.

The above analysis shows that the overall time complexity of IAT could be expressed as $O(ln)+O(mn^2)$. Compared with the length of query logs, the size of associative keywords and the number of hidden dimensions are usually too little to consider. Therefore, the time complexity of the proposed IAT is mainly determined by one factor, namely the length *l* of query logs.

ALGORITHM 1.

IAT (Indirect Association and Topic Based Competitive Keyword Suggestion)

```
Input: Query logs Q, The seed keyword s.
Output: The competitive keyword set CK = {ck_1, ck_1… ck_n}.

Begin:
Preprocessing (Q) by Stanford Word Segmenter
Step1:                  /* Indirect association analysis*/
CAND = Φ                /* To store the candidate competitive keywords*/
RELATION = Φ            /* To store the relation pairs for the factor graph modeling*/
TRIPLE = Φ              /* To store the triple information for the factor graph modeling*/
CORP = Φ                /* Collection of associative keywords set AK for all the items */
s.AK = Find_Associative_Keyword_Set (Q, s)
CORP = CORP U {s.AK}
for each keyword a in s.AK do
        a.AK = Find_ Associative _Keyword_Set (Q, a)
        for each keyword c in a.AK do
                CAND = CAND U {c}
                RELATION = RELATION U {<s, c >}
                X = Calculate_Feature_Vector (Q, <s, c >)
                TRIPLE = TRIPLE U {<s, c, X>}
                c.AK = Find_ Associative _Keyword_Set (Q, c)
         CORP = CORP U c.AK
        End
End
ITEM = CAND U s

Step2:          /*Topic based competitive keywords suggestion through factor graph model*/
repeat
     repeat
        for each topic t do
        Gibbs_Sample_Mixture φ ~ Dir(φ,β)
        end
        for each item i in ITEM do
                i.AK = Get_ Associative _Keyword_Set (CORP, i)
                Gibbs_Sample_Mixture θ ~Dir(θ|α)
                for each word w in i.AK do
                        Gibbs_Sample_Dimension t~Mult(t|θ,α, i.AK)
                        Gibbs_Sample_Word w~Mult(w|φ, β,t)
                end
        end
     until Convergence
     TD = Update_Topic_Distribution (ITEM, θ, φ)        /* Topic distribution for items*/
     Conduct_Factor_Graph_Modeling (Q, TTIPLE, TD)
     Update_Parameters (μ, σ, r)
until Convergence
L = Label_Relaitonships (TRIPLE, μ, σ, r)
CK = Get_Effective_Competitive_Keyword_Set (TRIPLE, L)
OutPut(CK)
```

---

5 EXPERIMENTAL RESULTS

In this section, a series of experiments are presented centring on the proposed method IAT for competitive keywords suggestion. The following gives a detailed analysis about the experimental results.

5.1 Experiment Setup

Evaluation data. To perform experiments, we collect data from Baidu Tuiguang, a search advertising tool, which supports free download of daily query log data for a given keyword. The query log is characterized by features including query keywords and the corresponding daily search volume, as represented in Table 1. Fifteen different seed keywords are used to conduct the experiments. They are motivated from the mainstream product/service category (Wu et al. 2009) of Taobao.com, thus ensuring the diversity of the domains. The set of seed keywords are listed below in Table 2. Based on these seed keywords, approximately 6000 query logs are downloaded for the experiment.

Index	Seed keyword	Index	Seed keyword	Index	Seed keyword
1	Skype	6	Dove	11	PingAn Insurance
2	Columbia Sportswear	7	Pantene	12	Meizu
3	Wedome	8	Sony	13	HaiDian Fahrschule
4	Tide	9	Tuniu.com	14	Midea
5	Budweiser	10	Yoshinoya	15	Princeton University

Table 2. Seed keywords

Evaluation methdology. Furthermore a TREC-type evaluation methodology will be used to show the effectiveness of IAT. TREC-type is one of the classical evaluation methods for information retrieval and search engine performance (Can et al. 2004), and has also been adopted in the emerging UGC-based competitive intelligence (Ma et al. 2011; Bao et al. 2008). Ground truth labelling of the recommended competitive keywords was provided by 6 human evaluators, who are experienced at online shopping and familiar with search engine ads. Each recommended competitive keyword was paired with its corresponding seed keyword, for a total of 9000 pairs. Each pair was assigned randomly to three annotators to evaluate. For each pair (the seed keyword and the recommended keyword), the annotators were asked to give a judgment, whether the latter could be a competitive keyword for the seed keyword. Only agreement was achieved by at least two annotators, the recommended keyword could be considered as an effective result. As this is a semi-supervised method, 20 percent of the labelled data is used as the training data.

Baselines. In order to better verify the performance of the proposed method IAT, empirical comparison experiments are conducted with the baselines. In the framework of IAT, the query log based topic information is essential to recommend competitive keywords. To demonstrate the effectiveness, we compare it with the pure factor graph model, denoted by FGM. The same data is used in FGM except for excluding LDA modeling in the learning process. Despite of lack for free commercial tools to suggest competitive keywords, quite a few tools have been developed to support keywords advertising. For example, Baidu, a very popular online application among business practices and research work, has promoted a mature marketing tool named BaiduTuiguang to help advertisers select keywords. Therefore, this will be used as a second benchmark in the following experiments.

Evaluation metrics. Typical metrics, namely Precision, Recall, F_1-measure, are considered in the experiments to evaluate the effectiveness of the proposed method. It’s commonly used in the performance measurement of information retrieval (Powers 2011), keyword suggestion (Bartz et al. 2006; Chen et al. 2008; Joshi & Motwani 2006), and recommendation (Lathia et al. 2010). Given the seed keyword s, denote the collection of keyword sets as $K = \{K_1, K_2...K_n\}$, in which the set K_i is a list of keywords detected by the method P_i correspondingly. In addition, E_i is a subset of K_i and contains just the effective competitive keywords for the seed *s*. The metric F_1-measure could be formulated as follows,

$$F_1 - measure(P_i) = 2 \times Precision(P_i) \times Recall(P_i) / (Precision(P_i) + Recall(P_i)) \quad (11)$$

Where,

$$Precision(P_i) = |E_i| / |K_i| \quad (12)$$

$$Recall(P_i) = |E_i| / \sum_{j \in \{1,2,\ldots,n\}} |E_j| \quad (13)$$

In the Equation (12) and (13), $|E_i|$ and $|K_i|$ represent the size of E_i and K_i respectively. Precision, by comparing the number of effective competitive keywords to all the recommended ones, captures the suggestion accuracy of the corresponding method. Recall, by calculating the ratio between recommended effective competitive keywords and the universal effective ones, measures the power of the method to recommend effective competitive keywords. The metric F_1-measure, calculated as their harmonic mean, can measure the performance of a keyword suggestion method from the two perspectives simultaneously. Therefore, we would also use F_1-measure to compare the performance of IAT and other benchmark methods.

The environment for experiments and performance analysis is a Windows 7 system on a PC with Intel Core i3-2100 CPU (3.1 GHz) and 4G RAM. The programs of IAT and FGM are implemented with the basic routines in java.

5.2 Comparative Results

FGM and Baidu are chosen as benchmarks to compare with our method IAT. The measurements mainly include Precision, Recall and F_1-measure, which are defined in Section 5.1. Firstly, the recall values of three methods for all the 15 seed keywords are shown in Figure 1, the numbers of horizontal coordinates represent the indexes of these seed keywords that are introduced in Table 2. The results in Figure 1 indicate that IAT has the best recall values among the three methods. The statistical results tested by paired T-test and Friedman test are shown in Table 3 to further verify the findings.

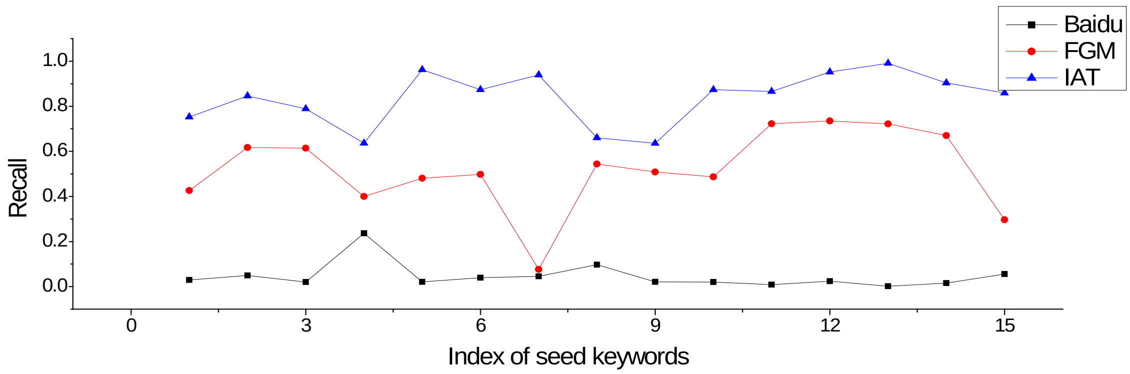

Figure 1. Comparison of recall values of the three methods

The testing results in Table 3 reveal that the recall value of IAT was significantly larger than the other comparative methods. As defined in Section 5.1, recall is the certain measure for the ratio of the number of the suggested competitive keywords to the number of all the competitive keywords. It accounts for the richness and the number of competitive keywords suggested by a certain method. Large recall values of IAT indicate that it can recommend more competitive keywords than the other methods and provide more potential choices for advertisers to achieve competitive advertising.

Methods	Hypothesis	t-value	Significance
Paired t-test	Recall value of Baidu< Recall value of IAT	-19.303	***
	Recall value of FGM < Recall value of IAT	-6.153	***
	Hypothesis	χ^2 value	Significance
Friedman test	Recall value of Baidu< Recall value of IAT	15.000	***
	Recall value of FGM < Recall value of IAT	15.000	***

*Notes: *: p<0.05; **: p<0.01; ***: p<0.001; --: no significance*

Table 3. Paired t-test and Friedman test on recall values of the three methods

Figure 2 shows the comparison of precision values for the three methods. The numbers of horizontal coordinates also represent the indexes of the chosen seed keywords. In the testing results, it can be found that IAT has far better precision values than Baidu. And for most of the 15 seed keywords the precision values of IAT also surpass FGM.

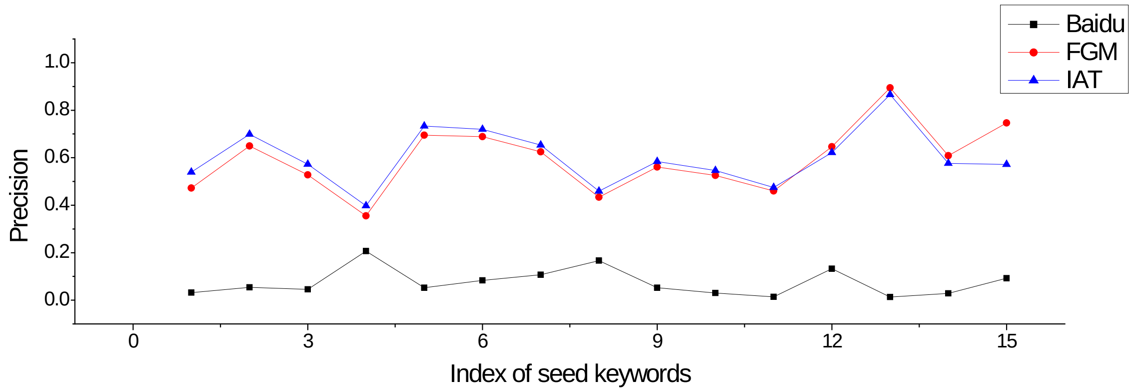

Figure 2. Comparison of precision values of the three methods

The statistical results for precision values tested by paired T-test and Friedman test are shown in Table 4. It further verifies that the precision value of IAT is significantly larger than Baidu. Although not significant, the t-value reveals that average precision value of IAT is larger than FGM.

Methods	Hypothesis	t-value	Significance
Paired t-test	Precision value of Baidu< Precision value of IAT	-13.243	***
	Precision value of FGM < Precision value of IAT	-0.550	--
	Hypothesis	χ^2 value	Significance
Friedman test	Precision value of Baidu< Precision value of IAT	15.000	***
	Precision value of FGM < Precision value of IAT	3.267	--

*Notes: *: p<0.05; **: p<0.01; ***: p<0.001; --: no significance*

Table 4. Paired t-test and Friedman test on precision values of the three methods

Precision and recall values are commonly used to measures the quality of suggested keywords from two different angels. To comprehensive and synthetically measure the quality of suggested results, harmonic mean of precision and recall is often adopted as F_1-measure, which is defined in Section 5.1. Figure 3 presents the F_1-measure values of all the 15 seed keywords for the three comparative methods. The results demonstrate that IAT has the best F_1-measure values than other two methods. The statistical results tested by paired T-test and Friedman test shown in Table 5 further verify the findings.

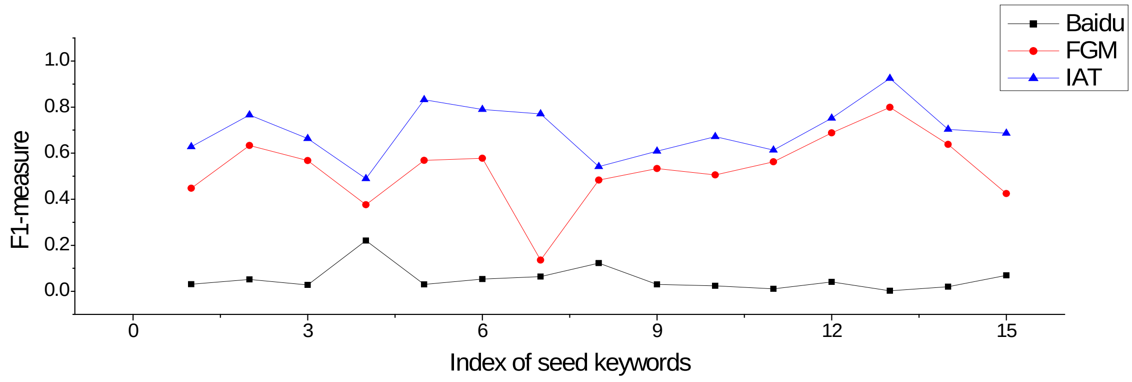

Figure 3. Comparison of F_1-measure values of the three methods

Table 5 demonstrates that the F_1-measure value of IAT is significantly larger than Baidu and FGM. As a comprehensive measure, F_1-measure combines the perspectives of precision and recall. The testing results in Table 5 reveal that compared with the other two methods, the proposed method IAT can recommend competitive keywords of high quality and effectiveness to help achieve competitive advertising.

Methods	Hypothesis	t-value	Significance
Paired t-test	F_1-Measure value of Baidu< F_1-Measure value of IAT	-16.333	***
	F_1-Measure value of FGM < F_1-Measure value of IAT	-4.378	***
	Hypothesis	χ^2 value	Significance
Friedman test	F_1-Measure value of Baidu< F_1-Measure value of IAT	15.000	***
	F_1-Measure value of FGM < F_1-Measure value of IAT	15.000	***

*Notes: *: p<0.05; **: p<0.01; ***: p<0.001; --: no significance*

Table 5. Paired t-test and Friedman test on F_1-measure values of the three methods

Comprehensively, IAT has a good performance on F_1-measure than any other benchmark method. It means that IAT is effective to recommend high quality competitive keywords for advertisers in search advertising market. Especially compared with Baidu, significant advantages could be seen for IAT in terms of precision, recall and F_1-measure. This is consistent with the discussions in the introduction section. Commercial keyword suggestion tools usually base on co-occurrence analysis to recommend general business relevant keywords, thus being unable to win competitive effect on search engines. When compared with the FGM, in spite of non-significance for the precision, apparent superiority could be demonstrated for IAT in terms of recall and F_1-measure. This demonstrates the crucial role of topic structure for identifying competition context, which is the focus of our method IAT. Therefore, conclusions about the effectiveness of IAT can be received as follows. From the three metrics (i.e. Precision, Recall and F_1-measure), IAT performs significantly better than the other comparative methods, which means that IAT is effective to recommend competitive keywords with high quality for advertisers.

6 CONCLUSIONS

It is deemed meaningful and desirable for organization development to conduct competitive advertising on search engines. Traditional methods based on co-occurrence analysis tend to mine a limited number of relevant keywords, and can hardly apply to meet the needs of competitive advertising. To respond to the organization needs for competitive advertising, this study proposes a topic based competitive keyword suggestion method IAT, which employs the indirect associations to find candidate competitive keywords and leverages their hidden topic information to further recommend massive novel competitive keywords. Extensive experiments have been conducted on the comparison between IAT and two other commonly used keyword suggestion methods, demonstrating that IAT is more effective on the suggestion of competitive keywords and can greatly help organizations achieve competitive advertising to expand their market shares.

Future studies can be explored to enrich the above research. One is to expand the comparison experiments to understand the proposed method comprehensively. The other is to measure the competitiveness of the recommended keywords and rank them for the advertisers to select. And furthermore, it may extend the proposed method to some other useful context, such as competitor identification in online reviews and patents.

Acknowledgments

This work was partly supported by the National Natural Science Foundation of China (71402186, 71372044, 71302158, 71331007, and 71110107027) and the MOE Project of Key Research Institute of Humanities and Social Sciences at Universities of China (12JJD630001).

References

Abhishek, V. and Hosanagar, K. (2007). Keyword generation for search engine advertising using semantic similarity between terms. In Proceedings of the 9th International Conference on Electronic Commerce, 89-94.

Amiri, H., AleAhmad, A., Rahgozar, M. and Oroumchian, F. (2008). Keyword suggestion using concept graph construction from Wikipedia rich documents, In Proceedings of the 30th European Conference on Information Retrieval, 63-9, Springer, Berlin.

Bao, S., Li, R., Yu, Y. and Cao, Y. (2008). Competitor mining with the web. IEEE Transactions on Knowledge and Data Engineering. 20(10), 1297-1310.

Bartz, K., Murthi, V. and Sebastian, S. (2006). Logistic regression and collaborative filtering for sponsored search term recommendation. In Proceedings of 2nd Workshop on Sponsored Search Auctions, 61-70.

Blei, D. M. (2012). Probabilistic topic models. Communications of the ACM, 55(4), 77-84.

Blei, D. M., Ng, A. Y., and Jordan, M. I. (2003). Latent Dirichlet Allocation. Journal of Machine Learning Research, 3, 993-1022.

Broder, A.Z., Fontoura, M., Gabrilovich, E., Joshi, A., Josifovski, V. and Zhang, T. (2007). Robust classification of rare queries using web knowledge. In Proceedings of the 30th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR’ 07), 231–238.

Can, F., Nuray, R. and Sevdik, A.B. (2004). Automatic performance evaluation of web search engines. Information Processing and Management, 40(3), 495-514.

Chen, M. and Miller, D. (2012). Competitive dynamics: themes, trends, and a prospective research platform. The Academy of Management Annals, 6(1), 135-210.

Chen, Y.F., Xue, G. R., and Yu, Y. (2008). Advertising keyword suggestion based on concept hierarchy. In Proceedings of the International Conference on Web Search and Web Data Mining (WSDM’ 08), 251-260.

Clark, B. H. (2011). Managerial identification of competitors: accuracy and performance consequences. Journal of Strategic Marketing, 19(3), 209-227.

Colavolpe, G. and Germi, G. (2005). On the application of factor graphs and the sum-product algorithm to ISI channels. IEEE Transactions on Communications, 53(5), 818-825.

Da, Z., Engelberg, J. and Gao, P. (2011). In search of attention. The Journal of Finance, 66(5), 1461-1499.

Fudenberg, D. and Tirole, J. (2000). Customer poaching and brand switching. RAND Journal of Economics, 634-657.

Fuxman, A., Tsaparas, P., Achan, K. and Agrawal, R. (2008). Using the wisdom of the crowds for keyword generation. In Proceedings of the 17th International Conference on World Wide Web (WWW’ 08), 21–25.

Griffiths, T. (2002). Gibbs sampling in the generative model of Latent Dirichlet Allocation, Technical Report, Stanford University.

Huang, C. K., Chien, L. F. and Oyang, Y. J. (2003). Relevant term suggestion in interactive web search based on contextual information in query session logs. Journal of the American Society for Information Science and Technology, 54(7), 638-649.

Jansen, B. J., Liu, Z. and Simon, Z. (2013). The effect of ad rank on the performance of keyword advertising campaigns. Journal of the American Society for Information Science and Technology, 64(10), 2115-2132.

Jones R. (2011). Keyword intelligence: keyword research for search, social, and beyond. John Wiley and Sons.

Joshi, A. and Motwani, R. (2006). Keyword generation for search engine advertising. In Workshops of 6th IEEE International Conference on Data Mining (ICDM’ 06), 490-496.

Kelly, D., Gyllstrom, K. and Bailey, E. W. (2009). A comparison of query and term suggestion features for interactive searching. In Proceedings of the 32nd International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR’ 06), 371-378.

Kschischang, F. R., Frey, B. J. and Loeliger, H. (2001). Factor graphs and the sum-product algorithm. IEEE Transactions on Information Theory, 47(2), 498-519.
Lathia, N., Hailes, S., Capra, L. and Amatriain, X. (2010). Temporal diversity in recommender systems. In Proceedings of the 33rd International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR' 10), 210-217.
Liu, H., He, J., Gu, Y., Xiong, H. and Du, X. (2012). Detecting and tracking topics and events from web search logs. ACM Transactions on Information Systems, 30(4), 21.
Loeliger, H. A. (2004). An introduction to factor graphs. Signal Processing Magazine, IEEE, 21(1), 28-41.
Ma, Z., Pant, G. and Sheng, O. R. L. (2011). Mining competitor relationships from online news: A network-based approach. Electronic Commerce Research and Applications, 10(4), 418-427.
Powers, D. M. (2011). Evaluation: from precision, recall and F-Measure to ROC, informedness, markedness and correlation, Journal of Machine Learning Technologies, 2(1), 37-63.
Sarmento, L., Trezentos, P., Gonçalves, J. P. and Oliveira, E. (2009). Inferring local synonyms for improving keyword suggestion in an on-line advertisement system. In Proceedings of the 3rd International Workshop on Data Mining and Audience Intelligence for Advertising, 37-45.
Sayedi, A., Jerath, K. and Srinivasan, K. (2014). Competitive poaching in sponsored search advertising and its strategic impact on traditional advertising. Marketing Science, 33(4), 586-608.
Schwaighofer, A., Candela, J.Q., Borchert, T., Graepel, T. and Herbrich, R. (2009). Scalable clustering and keyword suggestion for online advertisement. In Proceedings of the 3rd International Workshop on Data Mining and Audience Intelligence for Advertising, 27-36.
Szpektor, I., Gionis, A. and Maarek, Y. (2011). Improving recommendation for long-tail queries via templates. In Proceedings of the 20th International Conference on World Wide Web (WWW' 11), 47-56.
Wei, Y., Wei, Q. and Zhang, J. (2013). From query log to competitive advertising: A business intelligence method for elaborating consideration set of keywords. In Proceedings of the 20th International Conference on Management Science and Engineering (ICMSE' 13), 179-185.
Wu, H., Qiu, G., He, X., Shi, Y., Qu, M., Shen, J. and Chen, C. (2009). Advertising keyword generation using active learning. In Proceedings of the 18th International Conference on World Wide Web (WWW' 09), 1095-1096.
Zhang, Y., Zhang, W., Gao, B., Yuan, X. and Liu, T. (2014). Bid keyword suggestion in sponsored search based on competitiveness and relevance. Information Processing and Management, 50(4), 508-523.