

入札キーワードの重要度を考慮したスポンサードサーチの提案と評価

潘 献宇[†] 王 鑫[‡] 成 凱[‡]

[†]九州産業大学大学院情報科学研究科 〒813-8503 福岡県福岡市東区松香台 2-3-1

[‡]九州産業大学情報科学部 〒813-8503 福岡県福岡市東区松香台 2-3-1

E-mail: [‡] chengk@is.kyusan-u.ac.jp

あらまし スポンサードサーチとは検索サービス利用者をターゲットとして、検索条件と一致する広告をスポンサードクリックとして表示させ、クリックに基づき料金を徴収する検索連動広告の仕組みである。従来のスポンサードサーチでは、キーワードの重要度を区別しない単純なキーワード設定方法しかなく適切なキーワード設定が困難である。本研究では、スポンサードサーチのクリック率を高めるため、キーワードの重要度を考慮した柔軟なキーワード設定方式を提案し、それに適した広告検索方式を開発する。また、実験によって提案方式が広告検索の精度の向上やクリック率の改善に有効であることを検証した。

キーワード スポンサードサーチ, 検索連動広告, 辞書検索, Top-K 検索, ベンチマーク

Design and Analysis of a Sponsored Search with Weighted Keyword Auctions

Xianyu PAN[†] Xin WANG[‡] and Kai CHENG[‡]

[†] Graduate School of Information Science, Kyushu Sangyo University

[‡] Faculty of Information Science, Kyushu Sangyo University

2-3-1 Matsuda, Higashi-ku, Fukuoka, 813-8503 Japan

E-mail: [‡] chengk@is.kyusan-u.ac.jp.jp

Abstract In this paper, a new sponsored search scheme is proposed for weighted keyword auctions.

Keyword Sponsored Search, keyword targeted advertisement, Dictionary Lookup, Top-K, Benchmark

1. はじめに

インターネットの普及に伴い、ネットワークを経由した情報消費と情報発信が急速に広がり、インターネット利用者をターゲットとするオンライン広告が新しい広告媒体として確立されつつある[8]。特にポンスードサーチ(Sponsored Search)に代表される検索連動型広告は注目を集めており大半の市場シェアが占められており、ラジオ広告市場と並ぶ 1,524 億円市場への成長が見込まれる[9]。

スポンサードサーチでは、検索サービスプロバイダー(以降、プロバイダー)、広告依頼者(以降、広告主)、検索サービス利用者(以降、利用者)により次のようにサービスが成り立っている。まず、広告主が広告対象のコンテンツを用意し、関連するキーワードを入札し、掲載依頼を行う。プロバイダー側では依頼を受けて広告データベースに登録する。利用者が検索サービスを利用するとき、検索キーワードに関連する広告を

データベースから取り出し、掲載順位を決めてスポンサーリンクの形で検索結果ページに表示させる。利用者が掲載の広告コンテンツを消費した場合は広告料が発生し、広告主に課金する。一般的に利用者がスポンサーリンクをクリックした場合にのみ課金する Cost-Per-Click (CPC) 方式が採用されている[3][4][5]。

プロバイダーにとって、入札価格とクリック率とも高いほど利益が高くなる一方、広告主が満足できなければ長期的に顧客流失のリスクが高くなる。広告主にとって、できるだけ多くの潜在顧客を獲得しながらコストが一定の範囲にとどまることが望ましい。スポンサードサーチサービスが順調に成長していくために、プロバイダーと広告主の利害関係のバランスを保つ必要がある。そのため、適切な情報を適切なタイミングで適切な利用者に発信し、クリック率を高めるだけでなく確実な広告効果が得られると期待される。したがって広告内容を適切に表すキーワード設定により、広

告内容と検索サービスの利用者との適切なマッチングを実現することが必要不可欠である。

広告主がキーワードを設定する際に検索数の多い人気キーワードを設定すれば広告が表示される確率が高くなると想定される。しかし人気キーワードは必ずしも「広告内容にとって最も適切なキーワード」とは言い切れない。利用者の情報ニーズに合わない内容であれば、広告をクリックしてくれなかったり、クリックしても実際に購入や申し込みまで行いかなかったりすると、広告効果が薄く、無駄な広告料金を払うことになる。

キーワード 🔍	完全一致の順位	実際のコスト (¥)	最高入札額 (¥) 📈	最低入札額	上位5位の最高入札額						
✓			入札価格の更新								
📄 マンション 京成線	3	93	93	9	515	514	93	93	92	その他の入札状況	

図 1 Yahoo! JAPAN におけるキーワード入札

しかし、既存の方式の問題点として、同じ広告また広告グループに対して複数のキーワードやフレーズを登録できるものの、キーワードやフレーズの間に OR 関係とみなし、重要度は区別されない (図 1)。また、必須と思われるキーワードと、不要と思われるキーワードの区別もできず、広告主にとって潜在顧客をターゲットする適切なキーワード設定は困難である。

本論文では、我々はスポンサードサーチにおけるキーワード設定方式を改善するために、キーワードの重要度を考慮したキーワード設計方式を提案する。提案のキーワード設定方式による新しいスポンサードサーチの実現と評価を行う。

2. スポンサードサーチにおけるキーワード登録

スポンサードサーチを利用する際に、広告主は検索サービスプロバイダーに掲載対象ページに関連するキーワードを登録する必要がある。キーワード登録と同時にクリックごとに支払える料金などの条件も設定しておき、オークションにより検索条件と一致する複数の広告から広告料金の高い広告を競争的に選出する。広告主がキーワードとともに掲載条件などを指定し広告掲載を依頼することは**キーワード入札**という。

2.1. キーワード入札

キーワード入札とは広告主が広告内容に応じてキーワードを指定するとともに、それぞれのキーワードに対して 1 クリック当たり支払える広告料金 (CPC: Cost-Per-Click) を設定することである。キーワード入札では、広告内容 (タイトル・説明文) とリンク先の URL を設定することで、関連キーワードが検索された

際に広告が表示されるようになる。同一の広告は複数のキーワードで入札可能である。キーワードごとに入札可能な最低入札価格があり、入札価格は、最低入札価格以上の価格に設定する必要がある。

キーワードは単語、フレーズ、単語リスト等の形、例えば、「エクステンション」、「女性向けウィッグ」、「ウィッグ 医療用」のいずれも可能である。広告表示の順番は、入札価格だけではなく、クリック率など広告の品質も考慮して判断される。

検索 q に関連して広告 $\{a_1, a_2, \dots, a_n\}$ が $1, 2, \dots, n$ の順で検索結果画面に表示されたとする。期待の収益 R は次のように求められる。

$$R = \sum_{i=1}^n ctr(q, a_i) \times cost(q', a_i, i) \quad (1)$$

ここでは、 $ctr(q, a_i)$ は広告 a_i の推定クリック率、 $cost(q', a_i, i)$ は入札キーワード q' により広告 a_i が i 番目に表示されるときにかかる広告費を表す。入札キーワード q' と検索キーワード q はキーワードマッチのアルゴリズムによって必ずしも同じとは限らない。キーワードマッチには部分一致マッチ (broad match)、完全一致マッチ (exact match)、フレーズマッチ (phrase match) 等がある [6]。

2.2. 重要度を考慮したキーワード登録

キーワード登録をより柔軟に行えるために、我々は、以下のような 3 種類キーワードを指定できるような方式を考案する。

表 1 提案キーワード登録方式

カテゴリ	キーワード	入札価格
必須	マンション	100
	アパート	105
	賃貸	120
不要	分譲	
重み付き	1. 京成線	130
	2. 不動産紹介	125
	3. 引っ越し	80

- **必須キーワード (Must Keywords)**: 検索条件に最低一つ含まれる必要のあるキーワード。一つも含まない検索は広告を表示させない。
- **不要キーワード (Stop Keywords)**: 検索条件に含まれてはいけないキーワード。一つでも含まれた場合は、広告を表示させない。
- **重み付きキーワード (Weighted Keywords)**: このカテゴリのキーワードが検索条件に現れると、一定の重みを加算する。重みが大きいほど、適合するキーワード数が多いほど、広告を表示させる可能性が高くなる。

同じ広告に対して、上記の3種類のキーワードを設定できる。例えば、ある賃貸紹介サイトのキーワード登録例として、必須キーワード「マンション、賃貸」、不要キーワード「分譲」、重み付きキーワード「京成線、不動産紹介、引っ越し」を表1に示すように登録している。つまり、検索キーワードに、「マンション」か「賃貸」のいずれかが含まれないと表示しない。また、「分譲」で検索されるときにも広告を表示しない。必須キーワードとともに、重み付きキーワードも含まれる場合は、広告が表示されるチャンスが多くなる。

ちなみに、不要キーワードのみの登録や、同じキーワードを違う種類として登録できない。例えば、同じキーワードは同時に必須キーワードと不要キーワードとして登録してはならない。同様に、同じキーワードは必須キーワードでもあるし、重み付きキーワードでもあることができない。

キーワードごとに入札価格が設定される。例えば、必須キーワード「マンション」は100円、重み付きキーワード「不動産紹介」は125円である。広告が掲載される際に、クリックごとに実際に払うべき金額は検索条件によって平均入札価格を使い計算する。平均入札価格についてのちほど説明する。

2.3. 優先順位から重みへの変換

広告主の負担を軽減するために重み付けの方法を簡略化し、キーワードごとに重みを明示的に付けず、優先順位のみ指定すればよい。システムによって優先順位から重みの具体値に換算する。一つの広告に n 個のキーワードが $1, 2, \dots, n$ の順になっているとする。このとき、順位 i の重みは次のように求められる。

$$w_i = e^{1-i/\lambda} \quad (2)$$

e は自然対数の底、 λ はパラメータとする。重みはキーワードの数と関係のないようになっているので、パラメータ λ が決まれば、順位毎の重みが固定するので、あらかじめ計算しておけば、必要なとき参照すればよい。表2は $\lambda=10$ のとき各順位における重みを示している。

表2 掲載順位から重みへの変換表 ($\lambda=10$)

i	1	2	3	4	...	9	10	...
w_i	2.46	2.23	2.01	1.82	...	1.11	1.00	...

また、次章で述べるように掲載順位を決めるために必須キーワードを特殊な重み付きキーワードとして扱う。公式(2)では、必須キーワードは $i=0$ の重み ($e=2.72$) を持つとする。つまり、必須キーワードは重み付きキーワードより大きな重みをもつ特殊な重み付きキーワードである。

3. 平均入札価格による広告評価

提案のキーワード設定方式に合わせて、広告検索及び広告掲載の方式について述べる。利用者が検索条件を入力して検索サービスを利用するときに、検索結果とともに表示すべき広告及びその掲載順位を決める必要がある。本研究では、平均入札価格を使い、掲載順位を決める。

3.1. 総入札価格と平均入札価格

広告データベースに格納されている広告の集合 $A=\{a_1, a_2, \dots, a_n\}$ とする。それぞれの広告に必須キーワード、不要キーワード、重み付きキーワードが登録されている。複数のキーワードが含まれる検索条件 $Q=\{q_1, q_2, \dots, q_m\}$ で広告を検索し、掲載順位を決める。 $q_1, q_2, \dots, q_m$ のうち広告 a_i の必須キーワード、不要キーワード、重み付きキーワードはそれぞれ N_i, S_i, W_i とする。また、広告 a_i の必須キーワードを集合 M_i とし、 $w_{i,j}, cost_j$ はそれぞれ W_i の j 番目のキーワードと入札価格とする。このとき、広告 a_i の**総入札価格 (Aggregate Price)**は次のように求められる。

$$r_i = \sum_j w_{i,j} \times cost_j \quad (3)$$

また、**平均入札価格 (Average Price)**は次のように求められる。

$$\bar{r}_i = \sum_j w_{i,j} \times cost_j / \sum_j w_{i,j} \quad (4)$$

例えば、広告 a に含まれる検索条件 q の重み付きキーワードは順位 1,3,4 の三つあり入札価格はそれぞれ 100,110,80 とする。このとき、広告 a の総入札価格と平均入札価格は以下のように計算する。

$$\begin{aligned} r &= 2.46 * 100 + 2.01 * 110 + 1.82 * 80 = 612.7 \\ 2.46 + 2.01 + 1.82 &= 6.29 \\ \bar{r} &= 612.7 / 6.29 = 97 \end{aligned}$$

提案方式における広告検索と掲載は次のように行う。

1. 検索条件と一致する広告のうち、**総入札価格**の高い K 個の広告を検索する。結果は集合 A とする。
2. **平均入札価格**の高い順で、広告の掲載順位を決める。
3. **平均入札価格**をベースにクリック単価を決めて課金する。

ただし、必須キーワードしか登録していない広告は、これまで通りに広告検索、掲載順位を行えるとよい。つまり、入札しているキーワードのうち、入札価格の高いものを一つ選んで評価し、掲載順位を決める。

3.2. 広告の Top-K 検索

大量の広告から総入札価格の高い K 個を求める問題は Top-K 検索アルゴリズムを用いて解決できる。Top-K 検索とは m 個の順序付きリストから総合的重みが最も大きな K 個を探し出す問題である。

広告検索の場合は、m の検索キーワードに対してそれぞれの $p_{i,j} = w_{i,j} \times \text{cost } t_j$ を求め並べ替えて、図 2 に示すような m 個の順序付きリストが得られる。

順位	List 1		List 2		List 3	
	広告	p1	広告	p2	広告	p2
0	a_4	230	a_0	250	a_3	130
1	a_2	182	a_7	245	a_1	100
2	a_3	165	a_4	182	a_7	92
3	a_5	82	a_8	149		
4			a_2	100		
5			a_5	82		

図 2 順序付きリストの Top-K 検索

Top-K 検索を効率的に行うために、FA, TA, BPA, BPA-2 等の有名なアルゴリズムが知られている [7]。BPA アルゴリズムは以下のように、総入札価格の高い K 個の広告を検索する。

1. m 個リストに対し先頭から同時調べていく。これまで触れたことのない新たな広告 a があるリストに現れた場合は、広告 a の存在をほかのリストにも確認し、存在する場合は、その順位と重みを取得したうえ a の重みの合計を求める。
2. ここまで調べた広告のうち、総合重みの最も大きな k 個を集合 Y として保持する。さらにここで調べた順位とその重みの合計も記録しておく。
3. リスト i のこれまで調べたことのある順位が 1 位から連続している最大の順位をベスト順位 bp_i とし、その広告の重みを $w_i(bp_i)$ とする。さらに $\lambda = \sum_{i=1}^m w_i(bp_i)$ 。Y の要素のうち総重みが λ より低いものがなければ、アルゴリズムが終了。そうでなければ、1 へ戻って繰り返す。

ベスト位置以下に広告が残っても、総重みが λ 以下なので Top-K の対象外である。逆に集合 Y にある広告の総重みが λ より大きいので Top-K の対象と判断できる。

BPA-2 アルゴリズムでは、1 へ戻ってアルゴリズムを繰り返すとき、表の一行一行で調べるのではなく、

ベスト順位の次の行から調べていく。原理が BPA と同じであるが、調べたデータを重複に調べることがない。

4. 実験評価

重み付きキーワードにより、検索の精度が高まり、クリック率の改善につながることが期待される。本章では、この改善を実験より評価する結果を報告する。

実験では、従来のキーワード登録方式と重み付きキーワード登録方式によって、同じ検索を行う際に、クリック率の変化を評価する。

4.1. テストベッドの構築

実験のベースとして、N 個の広告、M 個のキーワードを考える。広告ごとにそれぞれ必須キーワード、不要キーワード、重み付きキーワードをランダムに生成しておく。広告ごとの重み付きキーワードに、入札価格も設定しておく。

クリック率は、掲載順位 (Rank)、キーワードの特性 (Length)、販売業者 (Retailer)、商品ブランド (Brand) などの要素によって左右される [4][5]。本実験では、広告の掲載順位ごとにクリック率は一定と想定する。図 3 はアキュラカスト (AccuraCast) 社に提供された 2009 年の Google アドワーズの広告掲載順位ごとのクリック率の調査結果をしめしている [10]。本実験は図 3 のような分布でクリック率を自動生成する。

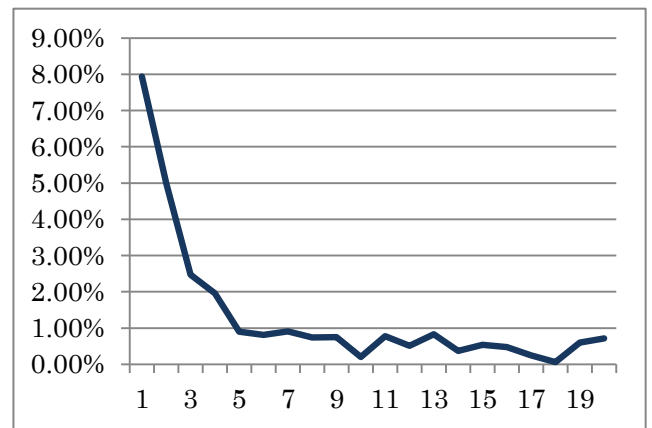


図 3 掲載順位とクリック率の関係

4.2. 実験結果

キーワード数 $M=1,000$ 、広告数 $N=10,000$ の場合で評価実験を行った。検索条件は次のようにランダム生成している。検索キーワード 1 個の条件を 1,000 セット、キーワード数が 2~5 の条件それぞれ 10,000 セットを生成し、検索キーワード数が 1 個の場合と合わせて 41,000 セットの検索キーワードで実験を行った。

実験結果は図 4 に示している．図 4 の横軸はクリック率の改善範囲，縦軸はその範囲の広告数を示している．例えば，横軸 4%のところ，クリック率の改善は 2~4%の広告数は 10,000 個のうち 3,093 個である．0% だけ，クリック率が変わらない広告数である．この実験結果からみて，広告数がある一定規模の場合では，提案手法は効率よく検索できることが分かった．

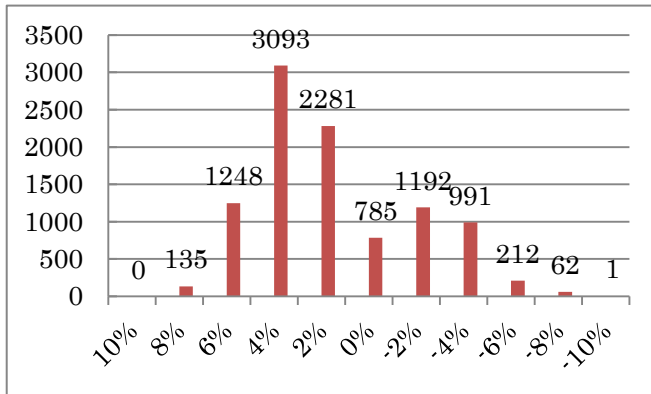


図 4 クリック率改善の実験結果

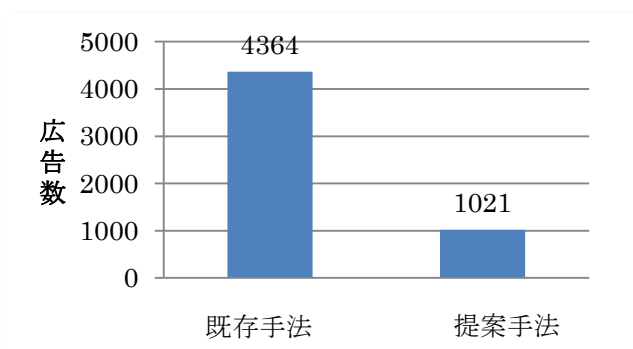


図 5 一度も表示されない広告数

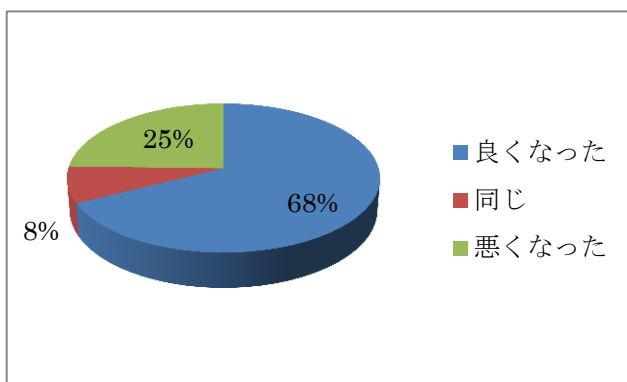


図 6 クリック率改善の割合

図 5 は提案手法により、一度も表示されない広告の数の減少結果を示している。さらに、図 6 はクリック率が改善した割合とそうではない割合を示している。結果として，クリック率は改善できた割合は 68%と多

く，提案方式の有効性を確認できたといえる．

終わりに

本論文では我々はスポンサードサーチにおける新しいキーワード設定方式及びそれに関する広告評価方式を提案した．重み付きキーワードの優先順位から重みへ変換の方法や平均入札価格による広告評価，順序リストの Top-K 検索による効率的な実装技術を検討した．実験によって提案方式のクリック率が改善できたと確認した．

今後の課題として，掲載順位に加え，キーワードの特性，販売業者，商品ブランドなどの要素を考慮した場合への対応やコンバージョン率を評価する仕組みの検討等が挙げられる．また，大規模の広告データベースにおいて，平均入札価格の計算や TOP-K 検索の効率的な方法を検討する必要がある．

参考文献

- [1] B. J. Jansen and M. Resnick, Examining Searcher Perceptions of and Interactions with Sponsored Results, Workshop on Sponsored Search Auctions at ACM Conference on Electronic Commerce (EC'05) 2005
- [2] [1] B. J. Jansen and M. Resnick, Factors Relating to the Decision to Click on a Sponsored Link, Decision Support Systems 44 (2007) 46–59
- [3] Daniel C. Fain, Jan O. Pedersen, Sponsored Search: A Brief History, Bulletin of the American Society for Information Science and Technology, Vol. 32, Issue 2, pp. 12–13, December/January 2006
- [4] A. Ghose, S. Yang, An Empirical Analysis of Sponsored Search Performance in Search Engine Advertising, WSDM '08, pp.241-250, 2008
- [5] D. Hillard, S. Schroedl, E. Manavoglu, H. Raghavan, C. Leggetter, Improving Ad Relevance in Sponsored Search, WSDM '10, pp.361-369, 2010
- [6] A. C. König, K. Church, and M. Markov. A Data Structure for Sponsored Search. In Proceedings of the 25th IEEE International Conference on Data Engineering (ICDE 2009), pp.90-101, 2009
- [7] R. Akbarinia, E. Pacitti and P. Valduriez. Best Position Algorithms for Top-K Queries, In Proceedings of VLDB 2007, pp. 495-506, 2007
- [8] 戸田 淳，わが国におけるインターネット広告の歴史的変遷とその本質，情報文化学会誌 17(1), 49-54, 2010-09-07
- [9] インターネット検索エンジンの現状と市場規模等に関する調査研究，総務省情報通信政策研究所調査研究報告書，2009-08-20
- [10] アク्यूラカスト (AccuraCast) , Google AdWords Click Through Rates Per Position, 2009-10-09 <http://knowledge accuracast.com/articles/adwords-clickthrough.php>