多腕バンディットによる検索結果の多様化に関する大規模クリックスルーログを用いた評価

下田　敬祐†　　江口　浩二†

† 神戸大学工学部情報知能工学科　〒 657–8501 兵庫県神戸市灘区六甲台町 1-1
E-mail: †kshimoda@cs25.scitec.kobe-u.ac.jp, ††eguchi@port.kobe-u.ac.jp

あらまし　検索エンジンによる検索結果の向上のため，類似性の高い内容を検索結果の中から自動的に排除することが望まれている．そのために，多腕バンディットアルゴリズムを用いた検索結果の多様化が最近注目されているが，実証評価が難しく，理論的な証明や仮想実験下での評価が多い．そこで，実際に運用されている検索エンジンの大規模なクリックスルーログデータを用いて，これらの評価及び考察を試みる．

キーワード　検索結果の多様化，多腕バンディット問題

Keisuke SHIMODA† and Koji EGUCHI†

† Kobe University
1-1 Rokkodaicho, Nada, Kobe, 657–8501 Japan
E-mail: †kshimoda@cs25.scitec.kobe-u.ac.jp, ††eguchi@port.kobe-u.ac.jp

1. はじめに

Web 検索は，インターネット基盤の不可欠な要素となっているため，機械学習の分野で重要視されている．特に，近年，ユーザ満足度向上のために注目されている技術として，検索結果の多様化があげられる．これは，検索結果の中から類似性の高い内容のものを自動的に排除し，より少ない手数でユーザが求める文書に辿り着くことを目的としている．このために，最近の検索エンジンでは，文書のドメイン情報を基にした多様化などの工夫が行われている．例えば，同じドメイン内の複数の記事が検索クエリにヒットした場合に全てを検索結果へ表示することなく，代表的なページのみを検索結果へと表示することで冗長性を排除しようとするものである．しかし，上記のような工夫では，内容が同じであってもドメインが異なる場合などに対応できない．具体的な例を挙げると，あるクエリに対し，同じニュース記事の内容を参照している異なるブログなどが多くヒットした場合，ドメインが異なるために検索結果にはそれらのページから複数が表示されてしまうことがあり得る．このために，ユーザのクリックスルーを基にした，多腕バンディット問題に対するアルゴリズムを検索技術へ応用する研究が最近注目されている [1].

ところが，これらの研究は，検索エンジンにおける多数のユーザのクリックスルー履歴を必要とするという性質上，実証評価が容易でなく，理論的な証明や仮想実験下での評価に留まっている．

そこで本研究では，これらの手法に対し，実際に運用されている検索エンジンの大規模なクリックスルーログデータを用いた評価を試みる．更に，今後実用化を見越した考察や実用上の課題，改良点などに関しても考察を加える．

2. 関連研究

2.1 検索結果の多様化

検索結果の多様化とは，ある検索クエリに対して表示される検索結果数件のうち，互いに類似性の高いものを排除し，そのクエリに対して多様な理由で潜在的に適合する文書を含む検索結果を生成することである [1] [2] [3] [7].

従来の検索結果の生成手法は，主にクエリ-文書間の類似度を判定し，高いスコアを持つものから順にランキング形式で表示する方法が主流である．ところが，このことはしばしば表示される検索結果の中に，内容が同じであるか，あるいは非常に類似性の高いものが含まれてしまうといった現象を引き起こす．このような検索結果はユーザにとって有用であるとは言えない．例えば，もし高級車の”jaguar”についての文書が，クエリ”jaguar”で検索しているユーザにとって適合でない場合，他の高級車について書かれている文書も適合である可能性が低くなる．つまり，クエリ-文書間の類似度スコアが常に他の文書と独立でありかつ一意に決まるという前提は，ユーザの情報ニーズの多様性とユーザ満足度(注1)の観点から見て適切であるとは

(注1)：ここでは主に検索結果の個々の文書がユーザに新たな情報をもたらすこ

言えない．上記の例（”jaguar”）で述べたような多様な潜在的意味を持つクエリの場合，特に顕著である．それらのような多様な潜在的意味を持ちうるクエリについて，それぞれの潜在的意味についての類似度スコアを定式化することは非常に容易でなく，現実的ではない．

また，Web 上の情報はニュースや流行などの影響を受け，常に変化し続ける．高級車の”jaguar”についての文書が多くのユーザにとってあまり適合とみなされず，クエリ”jaguer”に対して低い類似度スコアを付けられていたが，高級車に関するニュースなどの影響を受けてユーザの興味が高まり，ある時点で突然多くのユーザにとって求められる文書となる可能性もある．それらの変化の度に，クエリ-文書間の類似度スコアを更新し続けることは多くの労力を必要とし，かなり困難である．

そこで，より直接的に，ユーザのクリックスルーを基にして多様性のある検索結果を生成することを考える．問題の定式化のためにより詳しく言い換えると，”すべての時刻で任意の新しいユーザに対し表示する検索結果の中から少なくともどれかひとつを適合とみなしクリックされる確率が最も高くなるような検索結果を求める問題”となる．以上で述べた問題に対処するため，本論文では多腕バンディット問題を解くためのアルゴリズムを検索結果の生成に用いる [1] [7] ．

2.2 多腕バンディット（MAB）アルゴリズム

ここで，多腕バンディット問題（multi-armed bandit:以降 MAB と表記する）及び解くためのアルゴリズムについて紹介する．

MAB とは，単腕バンディットと呼ばれるカジノのスロットマシンをモデル化したものである．標準的な MAB アルゴリズムの目的は，複数個のスロットマシンを仮定したとき，有限回のプレイ回数で得られる報酬を最大化するためにどのスロットマシンを選択すべきかの意思決定を行うことである．具体的には，どのスロットマシンが当たりやすいか調べるための”探査 (exploration)”と，それまでに最も当たりやすいと予想されるスロットマシンを選択する”知識利用 (exploitation)”との間にトレードオフがあり，どのようにバランスをとりながら意思決定を行うかが鍵となる．

細かな問題設定の違いにより既に幾つかの解法が知られているが，最も自然な設定下で最大のパフォーマンスを発揮することが知られている UCB1 [4]，その改良バージョン UCB1+ [1] と，各スロットがそれぞれある一定の当選確率分布を持つという前提を置かないモデルでの最適解である EXP3 [4] を今回紹介する．UCB1 は各時刻 t で各スロットマシンに対し探査項と知識利用項からなる UCB1 値を以下の式で計算し，最大となるスロットマシンを選択する決定性アルゴリズムである．UCB1 値は，時刻 t，各スロット i に対し，スロット i のそれまでの平均利得 $\bar{x_i}$，スロット i をそれまでに選択した回数 n_i を用いて以下の式で計算できる．

$$\mathrm{UCB1}_i = \bar{x_i} + \sqrt{\frac{2\ln t}{n_i}}$$

とを意味する新規性を指す.

UCB1+はその探査項の時刻 t による部分を定数に置き換えたよりシンプルなアルゴリズムである．以下に示す.

$$\mathrm{UCB1+}_i = \bar{x_i} + \sqrt{\frac{1}{n_i}}$$

各スロットがある一定の確率分布に従って報酬を得られるという仮定に基づき，各時刻で最適であると予想されるスロットを決定論的に選択する UCB1 アルゴリズムに対し，EXP3 はそれらの仮定を置かない adversal な問題について考えだされたアルゴリズムで，非決定性アルゴリズムとして知られている．EXP3 はそれまでの探査から各時刻 t 毎に報酬の不偏推定量を求め各スロットに重み付けし，その確率でスロットを選択する．以下に数式を示す．時刻 t で，K 個のスロットの中からスロット i を以下の $p_i(t)$ の確率で選択する．

$$p_i(t) = (1-\gamma)\frac{w_i(t)}{\Sigma_{j=1}^{K} w_j(t)} + \frac{\gamma}{K}$$

その際得られた利得 $x_i(t)$ によって各スロットの重み w_i を更新する．

$$\hat{x_j}(t) = \begin{cases} x_j(t)/p_j(t) \text{ if } j = i_t \\ 0 \text{ otherwise} \end{cases} \quad \text{for } j = 1, \ldots, K$$

これらのアルゴリズムにおける，スロットを文書，報酬をクリック動作とみなせば，より報酬が大きくなるスロットを選択するアルゴリズムでよりクリックされやすい文書を選択することができる．

2.3 MAB アルゴリズムによる検索結果の多様化の仕組み

ここで，MAB アルゴリズムを用いた検索結果の生成方法及び多様化の仕組みについて説明する [1] [7].

検索結果に表示できる件数が k 件の場合，2.2 節で説明した MAB アルゴリズムを k 個用意し，各ランクにそれぞれ割り当てる．新規ユーザによる検索が行われた際，各 MAB アルゴリズムがそれぞれ最もクリックされやすい文書の選択を行い，k 件の検索結果を生成する．ユーザは各ランクに提示された文書を上位ランクから順にチェックし，適合であるものをクリックする．クリックされた文書を提示した MAB アルゴリズムには利得を与え，クリックされなかった場合には利得を与えない．各 MAB アルゴリズムがそれぞれ得た利得を基に学習し，次の検索結果の生成に活かす.

このとき，各ランクの MAB がそれぞれ独立して文書選択を行い提示することで，自動的に内容が似通っているものが排除されていく事になる．$k > 2$ 件目以降の MAB アルゴリズムの振る舞いに注目すると，仮に上位の MAB アルゴリズムで提示されている文書と類似性の高い文書を提示した場合，その文書がもしもユーザにとって適合であっても，上位のものがクリックされて検索が終了し，利得を得られないからである．

このようにして，任意の新しいユーザに対し多様な潜在的内容を含む，つまり表示する検索結果の中から少なくともどれかひとつを適合とみなしクリックされる確率が最も高くなるような検索結果を生成する事ができる．

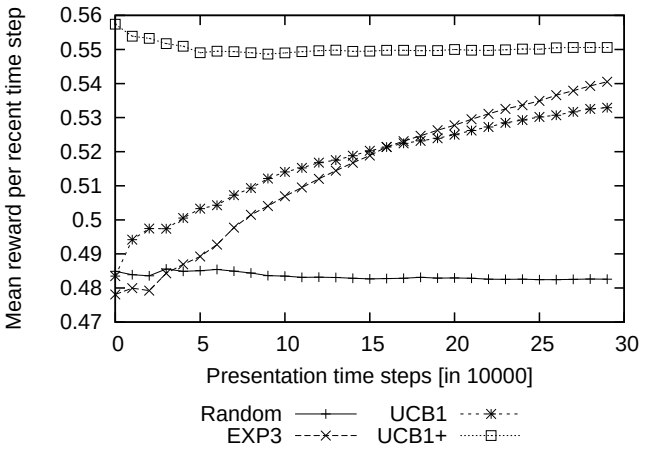

図 1　Simulation-based experimental results with UCB1,EXP3,and UCB1+

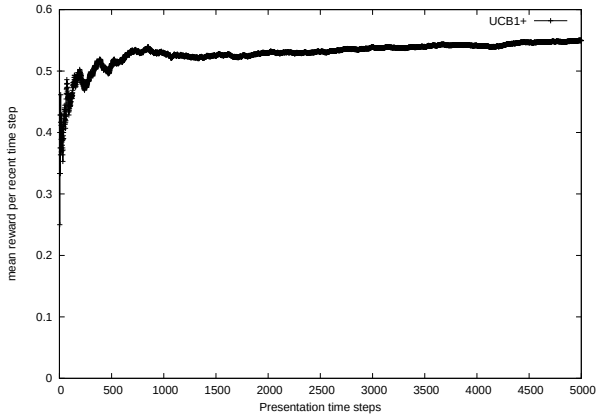

図 2　Simulation-based experimental results with UCB1+ (for the first 5,000 steps)

3. シミュレーションによる評価実験

本章では，MAB アルゴリズムを用いた検索結果の生成の挙動を確認するためのシミュレーションによる評価実験の結果を示す.

まずユーザ 20 人，文書 50 件があると仮定し，ユーザはそれぞれ文書 50 件の中から少なくともいずれかの文書を適合であると判断するものとする．各ユーザがどの文書を適合であると判断するかは，より現実に即した設定にするため中華レストラン過程（$\theta = 3$）を用いてある程度人気が偏るようなモデルを生成する．次にユーザを 1 人ランダムに選び，ユーザが検索を行ったものとして MAB アルゴリズムを用いて生成した長さ 5 件の検索結果を提示する．ユーザは提示された検索結果を上位から順に，適合であると判断した文書があるかどうかチェックを行う．もし適合であると判断した文書が検索結果内にあればクリックしたものとし，その結果を基に MAB アルゴリズムを更新する．これを 30 万回繰り返し，クリック率の推移を見る．これを MAB アルゴリズムに UCB1，UCB1+，EXP3 を用いた場合それぞれに対して行い，結果を比較する．また，ベースラインとしてランダムに検索結果を生成した場合のクリック率とも比較する．結果を図 1 に示す.

MAB アルゴリズムを用いたものはいずれもランダムに検索結果を生成した場合と比較してクリック率の上昇が見られた．UCB1，EXP3 を用いた場合はそれほど性能に差は見られなかった．また，UCB1+を用いたものが最も性能がよく，最も早く最適な検索結果へと収束していた．UCB1+についての詳細な挙動を確認するため、最初の 5000 回のクリック率の推移を示したものを図 2 に示す。

図 2 よりわかるように，およそ 1,000 回目ほどで最適な検索結果へと収束していた．これは文書数が 50 件と小規模であるため，あまり探索を行わないアルゴリズムが有利であると考えられる．また，今回の設定ではユーザと文書のモデルを生成した時点で各文書それぞれについて適合と判断するユーザの割合が決定するため，UCB1 や UCB1+にとって想定外である adversalial な状況が起こらないため，このような結果になったと考えられる．実際の Web 検索シーンでは，流行やニュース，文書の更新などの影響を受け，文書それぞれについて適合と判断するユーザの割合が逐次的に変化するため，異なる結果となることが予想される.

4. クリックスルーログを用いた評価実験

4.1 準　　備

これまでに紹介した UCB1，UCB1+，EXP3 を用いた検索結果の生成方法に対し，大規模なクリックスルーログデータを用いて実証評価を行うために必要な準備について説明する.

4.1.1 クリックスルーログデータ

まず，本研究で使用した，検索エンジンのクリックスルーログデータについて説明する．今回はロシアで主に運用されている，Yandex という検索エンジンの約 2 年間の間に取られたログデータを用いた(注2)．データセット内には，セッション，クエリ，検索結果，クリックスルーが含まれている．また，クエリや検索結果内の URL はすべて匿名化されており，無意味な数字による ID に置き換えられている．またそれぞれのサイズは次のとおりである.

- クエリ数：30717251
- URL 数：117093258
- セッション数：43977859
- レコード総数：340796067

データセットはセッションごとに区切られ，ひとつの検索クエリ ID に対して長さ 10 件の URL-ID が順に提示される．その後，クリックスルーがあればクリックされた URL-ID が全て記録される．上記の検索動作とクリックスルー動作は 1 対 1 で対応しているとは限らず，同一セッション内で連続して複数回検索が行われる場合や，1 つの検索に対し複数個の URL-ID がクリックされる場合も存在する点が，予備実験の設定と大きく違う点である.

4.1.2 実験に用いたデータについて

今回の実験では，Yandex のクリックスルーログデータから

(注2)：http://imat-relpred.yandex.ru/en/datasets

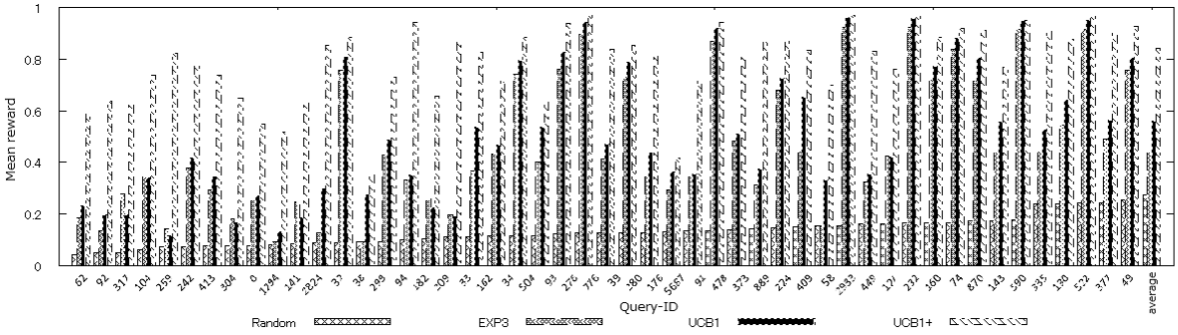

図 4　Clickthrough rate of each query by EXP3,UCB1,and UCB1+

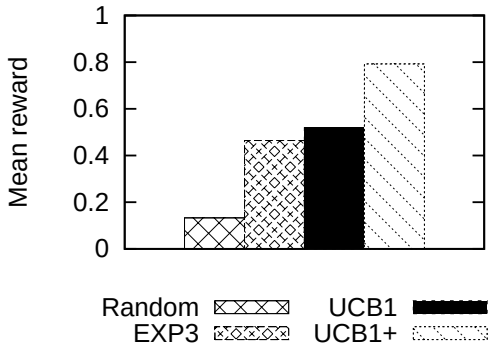

図 3　Average clickthrough rate of each approach

最もよくクリックスルーの起こっているクエリ ID 上位 50 件を抽出した．最も多かったクエリ ID のクリックスルー回数は 86637 回，50 番目に多かったクエリ ID のクリックスルー回数は 13485 回，平均すると 1 つのクエリ ID 毎に 31824.92 回のクリックスルーが起こっていた．また，それぞれのクエリ ID についてクリックスルーの起こった URL を順に抽出し，ユーザの検索動作によって適合であると見なされたデータ（以降，適合 URL-ID と呼ぶことにする）として扱った．次に，それぞれのクエリ ID 毎に Yandex の検索結果に一回以上表示された URL-ID の集合を抽出し，MAB による探索の候補とした．クエリ ID によって探索の候補となる URL-ID 集合の大きさは異なり，最大で 242 件，最小のものは 32 件の URL-ID を含み，平均で 86 件であった．

4.2　実験方法及び結果

シミュレーション実験と同様に，まず探索の候補となる URL-ID 集合から MAB を用いて検索結果を生成し，適合 URL-ID があるかどうかのチェックを行う．その結果を基に MAB アルゴリズムを更新し，次の検索結果の生成を行う．

生成する検索結果の長さは Yandex のログデータに合わせ，10 件とした．上記の動作をクエリ ID 上位 50 件毎に行い，ベースラインとしてランダムに検索結果を生成した場合との比較を行った．

最終的なクリック率をクエリ毎に計算し，すべてのクエリにわたって平均をとった結果を各種法毎に示したものを図 3 に示す．クエリ毎の最終的なクリック率は図 4 に示す．横軸の数字は各クエリの ID である．

図 3 から，シミュレーション実験とほぼ同様に，EXP3，UCB1 はほぼ同じクリック率となり，UCB1+が最も優れているという結果が得られた事がわかる．また，図 4 より，どのクエリに対しても UCB1+が優れている結果が見られるが，クエリによって各手法の学習効果に差が見られることがわかる．これは，実験に用いたクエリの潜在的意味の多様性の違いや，繰り返し数の差などによって学習に影響を与えたためであると考えられる．

5.　時間変化を考慮した評価実験

前章までの結果を受け，改良案についてここで考察する．

5.1　UCB1+における考察

上記の実験結果では UCB1+が最もクリック率が高いという結果が得られたが，UCB1 の拡張版である UCB1+はもともと ”各 URL のクリックされる確率が時刻によらず常に一定の確率値を持つ ”という前提のもとに，その確率値を推定し，意思決定を行なっている．このことは言葉を変えると，”あるクエリについて，そのクエリを任意の潜在的意味で用いているユーザの割合は，時刻を問わず常に一定である ”という仮定を置いている事を示している．しかし，この仮定は WEB 検索の状況に即しているとは言えない．

2.1 節の”jaguar”の例を用いるならば，ユーザ全体のうち，例えば高級車についての”jaguar”を求めて検索クエリに”jaguar”を用いるユーザの割合が常に一定であるという前提は，現実のユーザ行動の状況を反映しているとは言えない．高級車に関するニュースなどの影響を受けてその割合が変化することはあり得るし，これまで用いられていなかった意味（例えば，jaguar という名前のついた有名人が最近になって急激に話題に上がった場合など）での”jaguar”で検索が急激に増えることも想定できる．

この場合，過去全てにわたっての利得を用いている UCB1，UCB1+はそういった場面について適切な結果を生成できない場合がある．ユーザの興味が変化し，それに伴い任意の URL のクリック率が増加した場合でも，十分に探査を行った後の UCB1 では URL に対してのクリック率の変化を想定できないからである．

そこで，今回，UCB1 における知識利用項（各 URL 毎の平均クリック率）に，1 よりわずかに小さな値 α を時刻 t 毎の計算の度にかけ合わせることで，”直近に得られた利得を尊重し，

過去に得られた利得に対しては徐々にその重みを減少させる”改良版を考案した．以下に数式を示す．

$$\mathrm{UCB1}_i^+ = \bar{x_i}\alpha^t + \sqrt{\frac{1}{n_i}}$$

この変更によって，”各 URL のクリックされる確率（つまり，ユーザの興味の分布）は一定ではなく常に変化する可能性があり，また直近の結果に影響される”という，より WEB 検索の実情に近い想定に対応できるようになる．

このような改良を今回の実験で最も優れた結果を残した UCB1+に対して行い，改良前と同様の実験を実施し，UCB1+との結果を比較した．次章にて詳細を示す。

5.2 追加実験及び考察

α の値は，シミュレーションによる評価実験で UCB1+が凡そ 1000 回以内で最適な検索結果へと収束していたことを参考に，α^{1000} が凡そ 1 %となる $\alpha = 0.997$ を中心に，さまざまな値に対して実験を行った．クエリ毎の最終的なクリック率を並べたものを文末の図 5 に示す．図 5 は改良版が最も優れた結果を残した $\alpha = 0.99987$ の際の結果である．全クエリにわたって平均をとった結果は一番右端に示している．

図 5 から，クエリによって結果が異なることが読み取れる．また，UCB1+ではあまりクリック率の上昇が見られなかったクエリに対しても，改良版は高い学習効果が見られたことが読み取れる．全クエリにわたってクリック率の平均を取ると，改良によってもたらされた差は大きくはない．しかし，クエリ毎のクリック率に関して Wilcoxon 符号付順位検定を行った結果，$P < 0.05$ となり，有意に差が認められたと言える．

クエリによって結果が異なることから，今回用いたクリックスルーログ内に URL に対するクリック率の時間的変化が見られるクエリに対しては改良版がより良い結果を残し，そうでない場合は UCB1+がより優れているのではないか，と推察できる．

上記の推察の確認のため，UCB1+と改良版とで大きく結果の異なるクエリに対してどのような学習を行い，結果に影響したのかを推察するため，学習途中のクリック率についても記録をとった．改良版が優れていたクエリ（ID= 304, 182, 2824 , 5687) と，UCB1+が優れていたクエリ (ID=176, 94, 299, 317) を選び，繰り返し途中のクリック率の推移をとって観察した．代表的な推移を示したクエリ ID= 304，クエリ ID= 176 の結果を図 6，図 7 に示す．

図 6，図 7 の結果から，最初の収束段階である程度差が見られることがわかる．また，手法毎に両クエリに対する推移を比較すると，改良版はどちらのクエリを用いた場合も $t = 2,000$ 付近で一度クリック率が低下し，その後クリック率の回復が見られたのに対し，UCB1+ではそのような低下後の回復の様子は観察できなかった．また各クエリに対し，クリック率の低かった方のアルゴリズムでは，試行を重ねるほどクリック率のわずかな低下が見られる．この事から，初期の試行では最適であった検索結果に収束してした後，URL に対するクリック率の時間的変化が起こり，最適な検索結果でなくなったとしても以降の更新が行われていないことが予想される．そこで，クエリ ID= 304，クエリ ID= 176 に対しての収束の様子をより細かく観察するため，上記の実験の最初の 5000 回の試行について同様にクリック率の推移を調べた．結果を図 8，図 9 に示す．

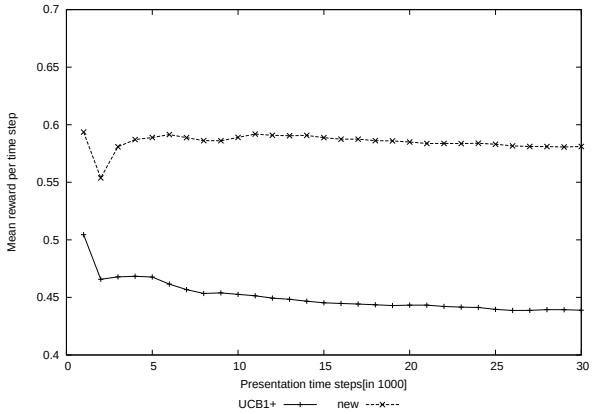

図 6　Transition of clickthrough rate [Query-ID = 304]

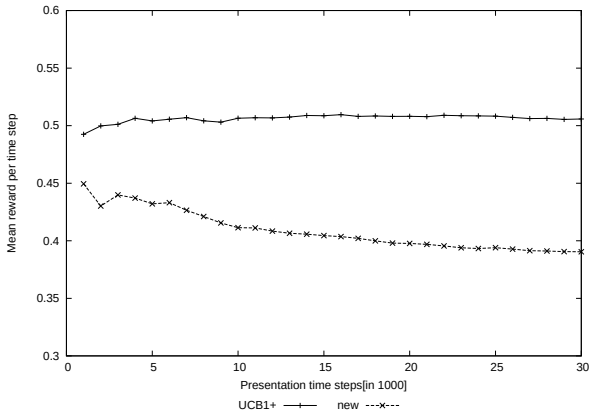

図 7　Transition of clickthrough rate [Query-ID = 176]

図 8，図 9 より，全体的なクリック率に差はあるものの，どちらもその推移の様子については同じような傾向が見られる．また，t が小さい場合の学習結果の差が，全体的なクリック率の差となっている事が読み取れる．この事は，t が小さい場合の各アルゴリズムの意思決定は主に探査項による影響が大きいことや，今回用いた α の値が 5000 回程度ではあまり影響を及ぼさない ($\alpha^{5000} = (0.9987)^{5000} = 0.522023719$ となり，$t = 5000$ で $t = 1$ の利得が約 52 %残っている) ことから説明できる．また，実際に URL 毎のクリック率の時間的変化について調べるため，クエリ ID= 304，クエリ ID= 176 を用いて行われた検索によってどの URL がクリックされたかを時間毎にプロットした．以下の図 10，図 11 に示す．縦軸に URL-ID を取り，横軸は時刻を示している．

クエリ ID= 304，クエリ ID= 176 共に，どの時刻にも普遍的にクリックの起こっている URL が存在することがわかる．またどちらのクエリに関しても，t が非常に小さい段階（$t < 1000$）で，クリック率の時間的変化が起こっていることが観察できる．特徴的なのは，クエリ ID= 176 の図 11 では，$t = 6000$ 付近で大きなクリック率の時間的変化の起こった URL が 2 件確認できるのに対し，クエリ ID= 304 の図 10 ではそのような変

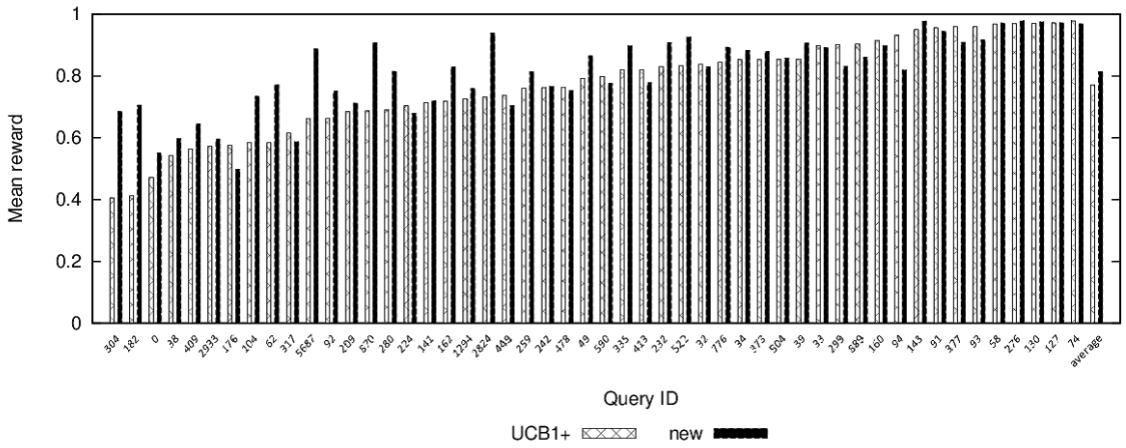

図 5　Clickthrough rate for each query with UCB1+ and new approach

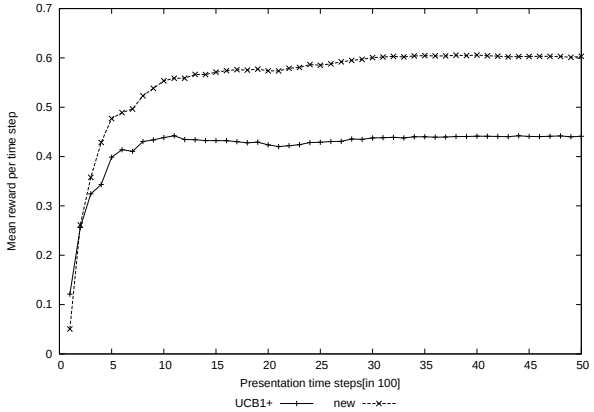

図 8　Transition of clickthrough rate for the first 5,000 steps[Query-ID = 304]

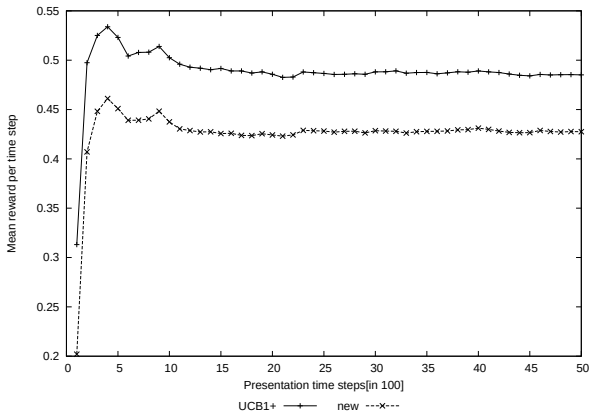

図 9　Transition of clickthrough rate of the first 5,000 steps[Query-ID = 176]

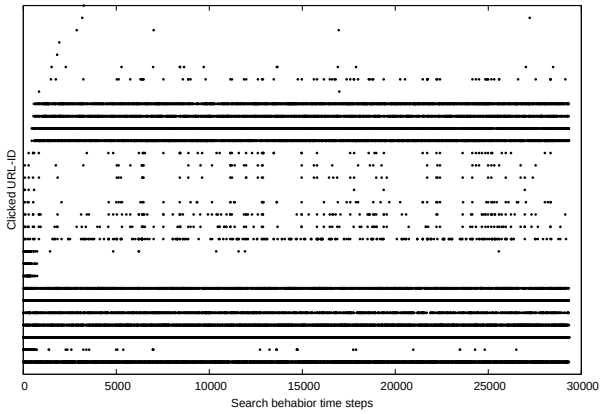

図 10　Clickthrough operation for each URL[Query-ID = 304]

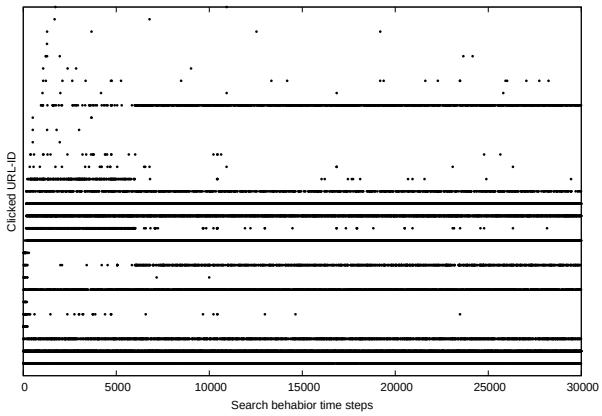

図 11　Clickthrough operation for each URL[Query-ID = 176]

化は見られない．また，クエリ ID= 304 の図 10 では非常によくクリックされている（線の濃い）URL 約 10 件と，しばしばクリックされている（線の薄い）URL が 7，8 件確認できるが，クエリ ID= 176 の図 11 では，$t = 6000$ 以降はほぼ固定の URL がよくクリックされていることがわかる．言いかえると，クエリ ID= 304 は潜在的意味の多様性が大きく，またクリック率の時間的変化があまり起こらなかったが，クエリ ID= 176 は潜在的意味の多様性が小さく，またクリック率の時間的変化が起こりやすかった，と言える．

クエリ ID= 304 に対しては改良版のアルゴリズムが優れており，クエリ ID= 176 に対しては UCB1+が優れていたことと合わせて考察すると，改良版のアルゴリズムでは知識利用項を減じているため，探査戦略を取る割合が増加し，クエリ ID= 304 のような潜在的意味の多様性を持つ語に対して有効性が見られたと考えられる．また改良版のアルゴリズムでは，$t < 1000$ 以下でのクリック率の時間的変化に対しては有効に対処できたが，

$t = 6000$ 付近での変化にはうまく対処できていない．この事の詳しい原因については，α をさまざまな値に変えながら詳細な動作確認などを行うことで，今後調査を行う予定である．

また，今回の実験は，Yandex の検索ログデータの中からよく検索されているクエリの上位 50 件を用いている．どのような検索クエリが頻繁に用いられているのか，またそれらのような頻繁に用いられているクエリに特徴的な傾向はないか調べるため，Google Insights for Search(注3)を用いて検索クエリのランキングについて調べたところ，特定の WEB サービスを示す固有名詞 (”facebook”，”twitter”など) や，潜在的意味の多様性のない一般的な内容を指す語（”動画”，”天気”など）が多く含まれていることがわかった．

前者のような特定の WEB サービスを示す固有名詞は，ユーザがそのサービスを用いる際に検索される場合が多く，潜在的意味の多様性が少なかったり，その時間的変化があまり起こらず，今回の改良版の有効性が見られにくいのではないかと推測される．後者のような一般的な内容を指す語についても，ユーザが求める結果を 1 回の検索で表示しようとしていない際に用いられることが多いのではないかと推察できる．そのために，このようなクエリを用いた検索によって表示された URL に対するクリック率の時間的変化が起こりにくいのではないかと推測される．

今回の実験では全てのクエリに対して一様な α の値を用いたが，前述のようなクリック率の時間的変化の少ないと予想されるクエリに対してクエリ毎に適切な α の値を選択することによって，より有効な検索結果の生成を行えないかどうかを今後の課題としたい．

6. ま と め

Web 検索におけるユーザ満足度の向上のための MAB アルゴリズムを用いた多様な検索結果の生成を実装し，実際の検索エンジンでのクリックスルーログデータを用いた検証実験でその有用性を確認した．また，仮想実験による既存研究では想定できていなかった実際の Web 検索におけるユーザの興味の分布の変化に対応するための改良を考案し，既存手法との比較実験を行った結果，その有効性について確認することができた．

今後の課題として，今回新たに提案した手法では，どのクエリに対しても一様な α の値を用いて実験を行ったが，クエリによってユーザの興味の分布の時間的変化の起こりやすさは異なるはずである．クエリによって最適な α の値を複数吟味し，また MAB の学習によって自動的に更新し，どのようなクエリに対してもユーザの興味の分布を捉え，更なるクリック率の改善を図ることができるかどうか今後調査を進める予定である．

また更なる課題として，より大規模なデータセットに対して効率的な探索が可能であるとされている GridBandit，Zooming Algorithm を実装し，UCB1，UCB1+，EXP3 を用いて検索結果を生成する場合と比較して評価を行う予定である．

(注3)：http://www.google.com/insights/search/

謝　　辞

本研究の一部は，科学研究費補助金基盤研究 (B)（20300038，23300039）による

文　　献

[1] F. Radlinski, R. Kleinberg, and T. Joachims, “Learning diverse rankings with multi-armed bandits,” Proceedings of the 25th International Conference on Machine Learning, pp.784–791, 2008.

[2] C.L.A. Clarke, M. Kolla, G.V. Cormack, O. Vechtomova, A. Ashkan, S. Büttcher, I. MacKinnon, “Novelty and diversity in information retrieval evaluation,” Proceedings of the 31st Annual International ACM SIGIR Conference on Research and Development in Information Retrieval, pp.659–666, 2008.

[3] R. Agrawal, S. Gollapudi, A. Halverson, and S. Ieong, “Diversifying search results,” Proceedings of the 2nd ACM International Conference on Web Search and Data Mining, pp.5–14, 2009.

[4] P. Auer, N. Cesa-Bianchi, and P. Fischer, “Finite-time analysis of the multiarmed bandit problem,” Machine Learning, vol.47, No.2-3, pp.235–256, 2002.

[5] R. Kleinberg, “Nearly tight bounds for the continuum-armed bandit problem,” Advances in Neural Information Processing Systems 17, pp.697–704, MIT Press, 2005.

[6] R. Kleinberg, A. Slivkins, and E. Upfal, “Multi-armed bandits in metric spaces,” Proceedings of the 40th Annual ACM Symposium on Theory of Computing, pp.681–690, 2008.

[7] A. Slivkins, F. Radlinski, and S. Gollapudi, “Learning optimally diverse rankings over large document collections,” Proceedings of the 27th International Conference on Machine Learning, pp.983–990, 2010.