

# 高速なウェブサイト最適化のための KPI 設計手法の提案

## Alternative KPI Design Method for Rapid Website Optimization

飯塚 修平      松尾 豊  
Shuheii Iitsuka      Yutaka Matsuo

東京大学工学系研究科技術経営戦略学専攻

Department of Technology Management for Innovation, The University of Tokyo

Website optimization is widely used to improve the performance of websites by comparing user behavior related to the different variations of the given website. Although the goal is to find the optimal variation which maximizes the performance within the given time, the required time varies according to the Key Performance Indicator (KPI) to optimize. Some KPIs take long time to reflect the effect of the change, which can be the bottleneck of the optimization speed. In this paper, we propose alternative KPI design method which makes the optimization process faster by using the log data. The experiment with an actual website showed that our proposed method finds the optimal variation in the 16.7% of the time which the baseline method consumes.

### 1. はじめに

近年、ウェブサイト最適化が様々なウェブサイトでもパフォーマンスを向上させるために用いられている。ウェブサイト最適化とは、ユーザをサンプルとした比較対照実験によってウェブサイトのパフォーマンスを向上させることである。ウェブサイトの要素の一部を変更することで異なるバリエーションを生成して訪問するユーザを振り分け、各バリエーションに振り分けられたユーザの行動の違いを比較して、好ましいユーザ行動を喚起するバリエーションを発見する。ウェブサイトの機能を実現するアルゴリズムのパラメータの変化や、ウェブページに表示されるラベルの文言やボタンの色の違いなど、小さな変化がウェブサイトのパフォーマンスに大きな影響を与えることが知られている。マイクロソフト社が運営する検索サービス Bing<sup>\*1</sup>では、検索結果ページのリンクの色の濃さをわずかに変更するだけで年間 10 億円の収入増を実現することに成功した [Kohavi 14].

ウェブサイト最適化は広く用いられている手法であるが、まだ未解決の問題がある。そのひとつが最適化する指標、すなわち **Key Performance Indicator (KPI)** の設計手法である。ウェブサイト最適化は、KPI の設計によって得られる結果が大きく異なる。たとえば、不適切な KPI を用いることによって短期的な収益は向上しても、ユーザへ提供するユーザ体験が劣化してしまう問題が知られている。また、同じ目的に沿った KPI であっても、値の取り方によって最適化に要する時間が異なる。ウェブサイト最適化において KPI の設計はパフォーマンスを大きく左右する重要事項である。

特にウェブサイト最適化の効率を左右する問題として、KPI によって最適化に要する時間が異なるということがある。たとえばユーザの再訪問率や生涯顧客価値などは、実際にユーザが訪問や購入などの行動を起こさない限りは算出することができない。また、Google AdSense<sup>\*2</sup>や MicroAd<sup>\*3</sup>のような外部の広告配信サービスをウェブサイトには貼り付けている場合、

連絡先: 飯塚修平, 東京大学工学系研究科技術経営戦略学専攻, 東京都文京区本郷 7-3-1 工学部 2 号館 92C1 号室, 03-5841-7718, iitsuka@weblab.t.u-tokyo.ac.jp

\*1 Bing <http://www.bing.com/>

\*2 Google AdSense <http://www.google.com/adsense/>

\*3 MicroAd <http://www.microad.co.jp/>

連動するサービスの情報が更新されるまでは最適化の結果を得ることができない。このような性質がウェブサイト最適化のスピードのボトルネックになってしまうことがある。

そこで本研究ではこの問題に着目し、高速にウェブサイト最適化を行うことができる KPI の設計手法を提案する。提案手法では最適化に時間を要する目的の KPI (**目的 KPI** とよぶ) に対して説明力が高く、かつ速報性が高い KPI (**代替 KPI** とよぶ) を設計することで、高速なウェブサイト最適化を実現する。実際のウェブサイトに対して代替 KPI に着目してウェブサイト最適化を行った結果、目的 KPI のみに着目して最適化を行った場合に比べて 16.7% の時間で最適解を発見できることがわかった。提案手法はウェブサイトの最適化だけではなく、学校教育における教育効果の測定や政策の効果測定など、施策効果の測定に時間がかかる領域全般に適用可能な手法である。

### 2. 関連研究

ウェブサイト最適化は様々なウェブサイトでも成果を収めており、より効率的な実験を行うための研究も盛んに行われている。Bing や Amazon<sup>\*4</sup> では、ウェブページのレイアウトやスタイルに小さな変更を加えてウェブサイト最適化を行うことで、収益を増加させることに成功した [Kohavi 14]. Google<sup>\*5</sup> では効率的な実験を実現するために、検索結果にまつわるパラメータをグループ化することで、大量のパラメータを同時に最適化する仕組みを開発している [Tang 10].

ウェブサイトをパラメータの組み合わせとして表現することで、ウェブサイト最適化は組合せ最適化問題として捉えることができる [飯塚 14]. また、ウェブサイトをエージェント、ユーザの行動を報酬とみなすことで、ウェブサイト最適化を強化学習の枠組みで捉えることも可能である。特に環境の状態は一定とみなして多腕バンディット問題として定式化することで、**バンディットアルゴリズム**を用いることが可能になる [White 12].

KPI の設計手法についても研究が行われており、Borodovsky らは、ボタンのクリックや購入の有無などの指標をモデル化する際はポアソン分布よりも負の二項分布を用いたほうが当てはまりがよいことを示している [Borodovsky 11]. また、ウエ

\*4 Amazon <http://www.amazon.com/>

\*5 Google <http://www.google.com/>

ウェブサイト最適化を始める前に過去のログデータを活用して各 KPI の性質を分析することで、KPI の変化の感度を向上させる手法が提案されている [Deng 13].

このようにウェブサイト最適化は様々なウェブサイトで用いられており、広く研究されている。KPI の設計手法についても研究が進められているものの、クリック率や購入率など即座に変化が観測できるものを対象としたものが多い。一方、本研究ではユーザのリピート率や外部サービスから与えられる情報など、観測に時間を要する KPI を最適化する手法を提案する。

### 3. 提案手法

この章では、まずウェブサイト最適化で用いられる KPI を、KPI を算出する母集団の違いおよび観測までの期間が長期化する原因によって整理する。その上でウェブサイト最適化問題を定式化し、高速なウェブサイト最適化を実現する代替 KPI の設計手法を述べる。

#### 3.1 KPI と長期化の要因の整理

ウェブサイトにおけるユーザ行動にはばらつきがあり、同じバリエーションを表示してもすべてのユーザが同じ行動を示すとは限らない。そのため、ウェブサイト最適化における KPI はある確率分布にしたがって与えられるユーザ行動から算出される統計量となる。このとき KPI を分類するひとつの方法として、統計量を算出する母集団によって分類する方法が考えられる。たとえば、ユーザの週間リピート率はバリエーション  $x$  を表示したユーザ集合  $U_x$  のうち、一週間以内にウェブサイトを再訪問したユーザ数の比で表すことができる。一方、検索エンジンにおける表示順位はバリエーション  $x$  を表示するウェブページ集合  $P_x$  に含まれるウェブページの表示順位の平均によって表すことができる。ここでは、バリエーション  $x$  を閲覧したユーザ集合  $U_x$  を母集団とする KPI を **User-based KPI**、バリエーション  $x$  を表示したウェブページ集合  $P_x$  を母集団とする KPI を **Page-based KPI** と呼ぶことにする。他にもバリエーション  $x$  を表示した時点のタイムスタンプや乱数による母集団の形成も考えられるが、ウェブサイト最適化ではユーザとウェブページについて統計量を算出したものを用いることが多いため、今回はこの2つに限定して考える。User-based KPI ではユーザ ID やユーザに付与されたクッキーなどによってユーザをグループ分けして母集団が形成される。一方、Page-based KPI はページの URL やコンテンツ ID などによってウェブページをグループ分けして母集団が形成される。

ウェブサイト最適化において KPI の観測が長期化する要因には様々なものが考えられるが、ここでは特にステークホルダーの行動頻度による長期化要因に着目する。ここに含まれる要因のひとつに、ウェブサイトを利用するユーザの行動頻度によるものがある。たとえばウェブサイトのリピートユーザ数を KPI とする場合には、一週間や一ヶ月といった期間をデータ収集のために設ける必要がある。これはユーザがウェブブラウザを通してインターネットに触れる頻度以上にウェブサイトの利用頻度を上げることは不可能であり、ユーザの再訪問を観測するためにはある程度の期間が必要であるためである。このように、ユーザの行動頻度によって KPI の観測が長期化することを **ユーザ行動頻度要因** による長期化と呼ぶことにする。一方、連携する外部サービスの行動頻度によって長期化する場合も考えられる。たとえば、広告配信プラットフォームから配信された広告のクリック率や検索エンジンにおける表示順位などは、その外部サービスによって情報更新の頻度が決まってい

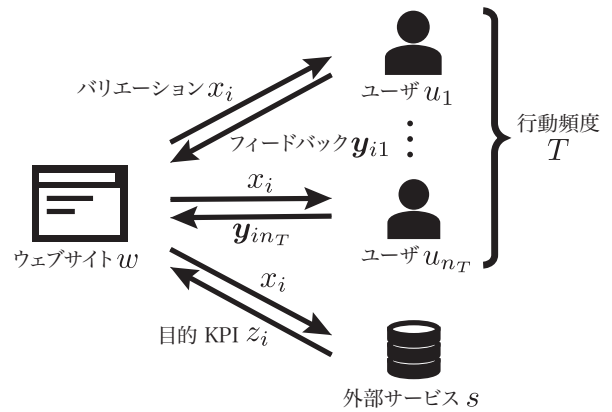


図 1: 外部サービス行動頻度要因によって長期化した KPI に対するウェブサイト最適化の概要

るため、それ以上の頻度で KPI の変化を観測することができない。このように、ウェブサイトと連携する外部サービスの行動頻度によって KPI の観測が長期化することを **外部サービス行動頻度要因** による長期化と呼ぶことにする。ユーザと連携する外部サービス以外にも、対象のウェブサイトのステークホルダーに関する KPI を最適化する場合には、その行動頻度が最適化のスピードのボトルネックになる可能性がある。

#### 3.2 長期化する KPI のウェブサイト最適化の定式化

ここでは外部サービス行動頻度要因によって観測が長期化した Page-based KPI を最適化する場合を取り上げて、ウェブサイト最適化全体のプロセスを図 1 に示す。ウェブサイト  $w$  はあらかじめ設定された振り分け方法に基づいて、ユーザ  $u_j \in U$  に対してバリエーション  $x_i$  を表示する。このときウェブサイト  $w$  はユーザの行動を観測し、ユーザ  $u_j$  から即座にフィードバック  $y_{ij}$  を受け取る。このフィードバック  $y_{ij}$  はウェブページの閲覧に関してその場で観測できる  $m$  種類の値で構成されるベクトルとする。外部サービス  $s$  は行動頻度  $T$  が経った時点で、ウェブサイト  $w$  にバリエーション  $x_i$  に対する目的 KPI の値  $z_i$  を公開する。ここで、 $n_T$  は行動頻度  $T$  の中でバリエーション  $x_i$  を表示したユーザ  $u$  の数である。

例としてブログサイトのインターネット広告の配置を変更し、広告のクリック率を最大化する場合を取り上げる。あるブログサイト  $w$  には 5 種類の広告配置案  $x_1, \dots, x_5$  があり、最もユーザの広告クリックを喚起する案を知りたい。一方、広告配信プラットフォーム  $s$  は、各ブログ記事ページの広告クリック率  $z_i$  を  $T = 1$  日の更新頻度で公開する。ブログサイト  $w$  はブログ記事の ID を 5 で割った時の余り  $k$  に対応する広告配置案  $x_k$  をユーザに表示する。ブログサイト  $w$  はユーザ  $u$  の滞在時間  $y_1$ 、ユーザの直帰の有無  $y_2$ 、記事のスクロール回数  $y_3$  の 3 種類の値を計測しており、ユーザが記事を閲覧し終えてウェブページから離脱すると、フィードバック  $y = (y_1, y_2, y_3)$  が記録される。

ウェブサイトが更新頻度  $T$  で外部サービス  $s$  から公開される目的 KPI  $z$  に基づいて最適化を行う場合、バリエーションを作成し、データを収集して最適なバリエーションを探索するサイクルを期間  $T$  で 1 度しか回すことができないが、フィードバック  $y$  から目的 KPI を説明する代替 KPI  $z$  を設計することができれば、期間  $T$  に複数回最適化のサイクルを回すことができる。



### 3.3 代替 KPI の設計手法

代替 KPI にはユーザの閲覧と同時に得られるフィードバックに基いて算出できる速報性と、目的 KPI の値を高い精度で予測できる説明力が求められる。そこで、ウェブサイト蓄積された過去データを解析し、フィードバックに含まれる素性およびその組み合わせから目的 KPI を精度高く予測する KPI を設計することができれば、それを代替 KPI として用いることができると考えられる。素性ベクトルから出力を予測する問題の解法には、重回帰分析やロジスティック回帰など様々な解法があるが、ここでは簡単に、素性のなかで目的 KPI との相関が高い KPI を代替 KPI として使う手法を説明する。ウェブサイト最適化ではバリエーションの優劣のみに興味があるため、ここではスピアマンの順位相関係数  $\rho = 1 - \frac{6 \sum_{i=1}^N (x_i - y_i)^2}{N^3 - N}$  を相関の強さを示す指標として用いる。ここで、 $x_i, y_i$  は  $i$  番目のペアのそれぞれの値、 $N$  はペアの総数を表す。

代替 KPI は以前に行ったウェブサイト最適化の際に収集された過去データ  $D$  から設計される。以前のウェブサイト最適化の際に観測されたフィードバック集合を  $Y = \{y_1, \dots, y_{n_T}\}$  とすると、過去データ  $D$  はフィードバック集合  $Y$  とそれに対応する目的 KPI  $z$  のペアの集合  $D = \{(Y_1, z_1), \dots, (Y_N, z_N)\}$  で表される。ここで  $N$  は過去に収集されたフィードバック集合と目的 KPI のペアの総数である。フィードバック集合  $Y$  について各素性の平均値からなるベクトル  $\bar{y}$  を考えると、フィードバックの平均値と目的 KPI のペア集合  $\bar{D} = \{(\bar{y}_1, z_1), \dots, (\bar{y}_N, z_N)\}$  (過去平均データとよぶ) を生成することができる。このペア集合  $\bar{D}$  について各素性の相関分析を行い、相関が強い素性を代替 KPI として利用する手法を提案する。

## 4. 評価実験

代替 KPI が高速なウェブサイト最適化を実現することを確かめるため、実際のウェブサイトを用いた検証実験を行う。今回は、ウェブページの広告配置を変更することによって広告クリック率を最適化する例を取り上げる。これは外部サービス行動頻度要因によって長期化した Page-based KPI を最適化する場合に相当する。本実験では、まずウェブサイト最適化以前に収集されたアクセスログデータと外部サービスから収集されたデータから過去平均データを算出する。その後、目的 KPI とフィードバックに含まれる各素性とのスピアマンの順位相関係数を算出し、最も相関が大きいものを代替 KPI として採用する。ある行動頻度で結果が与えられる目的 KPI に対して最適化を行う場合と、即座に結果が与えられる代替 KPI に対して最適化を行う場合とで最適なバリエーションに探索するのに要する時間を比較する。

今回は対象のウェブサイトとして「あのひと検索 SPYSEE\*6」(サイト A とよぶ)を用いる。サイト A は人物に特化した検索エンジンであり、人名を検索して該当した人物のプロフィールおよびウェブから観測される人間関係を表したネットワークを閲覧することができる。それぞれの人物のネットワークを表示するネットワークページには、ある広告配信プラットフォーム(外部サービス B とよぶ)が配信する広告が配置されている。ウェブサイトの構造がシンプルに、ネットワークページの広告配置について定常的にウェブサイト最適化が行われているため、サイト A を実験対象に選定した。

実験には、サイト A のネットワークページでウェブサイト最適化を行っていた 2013 年 5 月 14 日から 2013 年 5 月 17 日

表 1: 各素性と目的 KPI の順位相関係数

素性	相関係数 $\rho$	有意水準 $p$
ページ滞在時間 $y_1$	-0.273	0.417
直帰率 $y_2$	-0.045	0.894
離脱率 $y_3$	0.536	0.089

までのデータを用いる。この期間、サイト A は広告のクリック率を目的 KPI  $z$  としてウェブサイト最適化を行っていた。バリエーションは 3 種類  $x_1, x_2, x_3$  であり、ページに表示される人物の ID にしたがって表示するバリエーションが決定される。観測されるフィードバック  $y$  に含まれる素性はページ滞在時間  $y_1$ 、直帰率  $y_2$ 、離脱率  $y_3$  である。外部サービス B から情報を得られる頻度  $T$  は 1 日 1 回であるものとし、その間にユーザにウェブページが表示される回数は  $n_T = 2000$  とする。シミュレーションでは 2013 年 5 月 14 日に得られたフィードバックおよび外部サービス B からの過去平均データ  $\bar{D}$  をもとに各素性と目的 KPI の相関  $\rho$  を計算し、相関が大きい物を代替 KPI  $\zeta$  として採用する。2013 年 5 月 15 日から 2013 年 5 月 17 日のサイト A のアクセスログからユーザのフィードバックを、外部サービス B から得られたクリック率から広告のクリックを再現し、ウェブサイト最適化のシミュレーションを行う。

ウェブサイト最適化手法にはバンディットアルゴリズムのひとつである  $\epsilon$ -greedy アルゴリズムを用いる。 $\epsilon$ -greedy アルゴリズムにはパラメータ  $0 < \epsilon < 1$  が設定されており、ユーザの訪問があるごとに、確率  $\epsilon$  で探索、確率  $1 - \epsilon$  で活用の行動を取る。活用の行動をとった場合は最も期待値が高いバリエーションを表示し、逆に探索の行動をとった場合はそれ以外からひとつバリエーションを無作為に選択して表示する。このようにして探索と活用の間を繰り返すことで、最適なバリエーションを表示する状態へ収束する。本研究ではバリエーションが組み合わせの構造であるかどうかは問わないこと、またバンディットアルゴリズムの中でシンプルな解法であることから、 $\epsilon$ -greedy アルゴリズムを採用した。

表 1 に、フィードバック  $y$  に含まれる各素性とクリック率  $z$  の順位相関係数  $\rho$  およびその有意水準  $p$  を示す。離脱率については目的 KPI  $z$  にたいして中程度の相関が見られることがわかった。その他の素性については目的 KPI と相関が見られるものはなかった。

図 2 に、 $\epsilon$ -greedy アルゴリズムを用いてウェブサイト最適化を行った結果を示す。今回は評価指標として、各ユーザ数が得られた時点で表示されるバリエーションの平均クリック率を用いた。各 KPI に対して 100 回シミュレーションを行った。 $\epsilon = 0.3$  と設定した。なお、負の相関が見られた素性に関しては、代替 KPI を最小化するように最適化するものとする。目的 KPI に対して最適化した場合、振り分けられたユーザ数が  $n_T$  に到達するまでは低いクリック率のバリエーションを表示する状態が続くが、ユーザ数が  $n_T$  に到達した時点で大きく平均クリック率を上げることができている。一方、相関が比較的大きかった離脱率とページ滞在時間を代替 KPI として採用した場合には、ユーザ数が  $n_T$  に到達する以前から代替 KPI の速報性によって高いクリック率に到達することができている。相関が見られなかった離脱率を代替 KPI とした場合は、ユーザ数が増えても低いクリック率でとどまっている。ページ滞在時間または離脱率を KPI として用いた場合は、クリック率を KPI として用いた場合に比べて約 16.7%の時間で最適なバリエーションに到達することができた。以上より、過去平均デー

\*6 あのひと検索 SPYSEE <http://spysee.jp/>

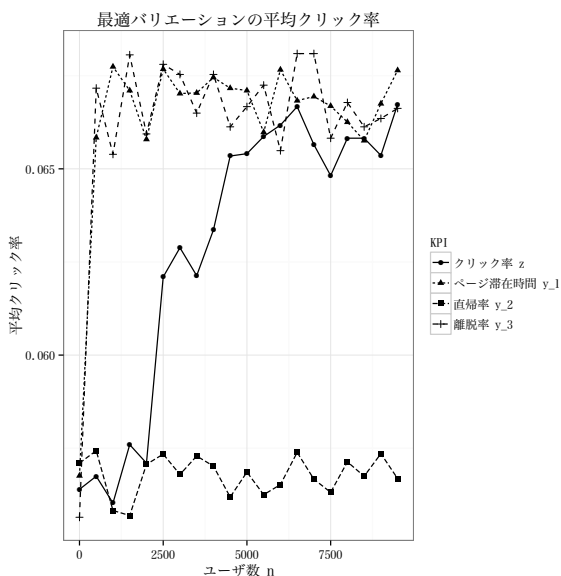


図 2: 各 KPI に着目した場合の最適バリエーションのクリック率の推移

タ  $\bar{D}$  から目的 KPI  $z$  と相関が高い素性が観測されれば、それを代替 KPI  $\zeta$  として用いることで、より高速なウェブサイト最適化を実現できることがわかった。

## 5. 考察

高速な最適化ができるということは、少ないサンプル数でも最適解に到達することができるということである。したがって、対象サイトの規模が小さくユーザー数が少ない場合でも提案手法を用いてウェブサイト最適化を行うことができると考えられる。しかし、代替 KPI を設計するためには以前のウェブサイト最適化から得られた過去データが必要であるため、対象のウェブサイトは定常的にウェブサイト最適化を行っている必要がある。

ウェブページから得られるフィードバックが少なく、代替 KPI を設計するための素性が少ない場合には、回帰分析を用いて素性を組み合わせ、新たな素性を生み出すことが改善案として考えられる。また、javascript で記述されたプログラムを実行すれば、アクセスログからは得られないユーザ行動を観測して素性を増やすことも可能である。また、ウェブサイト最適化の回数を重ねるにつれてウェブサイトの状況が変化するため、代替 KPI と目的 KPI との相関関係も変化する可能性がある。定期的に相関を計測して設計を更新することが、長期にわたって代替 KPI によるウェブサイト最適化を行う場合には有効であると考えられる。

今回の実験結果では目的 KPI である広告のクリック率に対して離脱率が相関、ページ滞在時間が逆相関していた。しかし、代替 KPI が原因で目的 KPI が結果であるという因果関係が必ずしも成立するわけではない。今回の実験ではむしろ「広告がよくクリックされるバリエーションでは、ユーザが広告先のページに遷移するために離脱率が上昇する」もしくは「広告がクリックされやすい配置では、ウェブページが表示されてすぐに広告をクリックして遷移してしまうため、ページ滞在時間が下がる」という逆の因果関係が成立している可能性がある。

このことを見誤ると、遷移後すぐにユーザが離脱してしまうような魅力の無いページが最適という結果を導き出してしまう危険性がある。

提案手法によって観測に長期間かかる KPI についても最適化することができるようになると、ユーザの満足度やブランドへの印象など、長時間かけて醸成する必要がある指標についても最適化することができるようになる可能性がある。また、ウェブ以外の分野においても学校教育における新しい教育カリキュラムや、行政における新たな政策の実施など、今までは効果の測定に数年かかった分野についても提案手法を用いることによって最適化を実現できる可能性がある。

## 6. まとめ

本研究では、ウェブサイト最適化において観測までの期間が長期化する KPI の問題を取り上げた。まず、目的とする KPI の長期化の要因を整理し、ウェブサイト最適化のプロセスを定式化した。その上で、ユーザから即座に得られるフィードバックに含まれる素性のなかから、目的とする KPI と相関が高いものを代替 KPI として最適化に利用する手法を提案した。実際のウェブサイトに対してシミュレーション実験を行った結果、代替 KPI を用いることで目的 KPI に対して最適化を行う場合に比べて高速にウェブサイト最適化を行うことができることがわかった。本手法はウェブの分野だけではなく教育や政治など、施策の成果を観測するのに時間がかかる分野で汎用的に適用することができる手法である。

## 参考文献

- [Borodovsky 11] Borodovsky, S. and Rosset, S.: A/B testing at SweetIM: The importance of proper statistical analysis, in *Data Mining Workshops (ICDMW), 2011 IEEE 11th International Conference*, pp. 733–740 (2011)
- [Deng 13] Deng, A., Xu, Y., Kohavi, R., and Walker, T.: Improving the sensitivity of online controlled experiments by utilizing pre-experiment data, in *Proceedings of the sixth ACM international conference on Web search and data mining*, pp. 123–132 (2013)
- [Kohavi 14] Kohavi, R., Deng, A., Longbotham, R., and Xu, Y.: Seven rules of thumb for web site experimenters, in *Proceedings of the 20th ACM SIGKDD international conference on Knowledge discovery and data mining*, pp. 1857–1866 (2014)
- [Tang 10] Tang, D., Agarwal, A., O'Brien, D., and Meyer, M.: Overlapping experiment infrastructure: More, better, faster experimentation, in *Proceedings of the 16th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pp. 17–26 (2010)
- [White 12] White, J.: *Bandit Algorithms for Website Optimization*, O'Reilly (2012)
- [飯塚 14] 飯塚 修平, 松尾 豊: ウェブページ最適化問題の定式化と最適化手法の提案, *人工知能学会論文誌*, Vol. 29, No. 5, pp. 460–468 (2014)