

検索意図の遷移検出に基づく 動的なクエリ推薦に向けた行動ログデータの分析

梅本 和俊¹ 中村 聡史^{1,2} 山本 岳洋^{1,2} 田中 克己¹

概要: 本論文では、検索クエリとそれに対する行動ログデータが与えられた際に、次の検索におけるユーザの検索意図の遷移を分類する手法を提案する。本手法では、検索・閲覧時に蓄積したユーザの行動ログデータから機械学習により分類器を構築することで、検索意図の遷移分類を実現する。分類器を用いた検索意図の遷移について実例をもとに考察を行うと共に、推定した検索意図の遷移に基づくクエリ推薦の可能性についても述べる。

キーワード: 検索意図, 行動ログ分析, 機械学習, クエリ推薦

Analysis of User's Behavior Log Data Towards Dynamic Query Suggestion Based on Search Intent Transition

KAZUTOSHI UMEMOTO¹ SATOSHI NAKAMURA^{1,2} TAKEHIRO YAMAMOTO^{1,2} KATSUMI TANAKA¹

Abstract: This paper proposes a method to classify transitions of users' search intents from current search to next one. This method uses a machine learning technique to construct a classifier from their behavior log data obtained from search and browsing processes. Based on the real search behavior log data, we discuss whether our method can estimate users' intent transitions accurately. We also describe transition-based query suggestions as a possible application of our method.

Keywords: Search intent, behavior log analysis, machine learning, query suggestion

1. はじめに

Web上の情報の爆発的な増加にともない、Web検索による情報収集が一般的なものとなっている。検索エンジンに入力される検索クエリのうち、約28%はそれ以前に入力されたクエリを修正したものが使われていること [1] から、ユーザはこうした検索タスクにおいては、複数のクエリを入力し反復的に検索を繰り返すことで自身の求める情報を探しているといえる。そのような複数クエリからなる検索タスクの例として「長期休暇時の旅行計画」や「就職活動の情報収集」といったタスクが挙げられる。「旅行計画」を例にとると、どこに旅行に行くか、旅行先で何をするか、

さらに、宿泊先の決定など、ユーザは様々な観点から情報検索を行う必要がある。そのためタスク中で入力される検索クエリ数も多くなる。他にも、検索対象に対するユーザの知識が浅い場合や、ユーザの情報要求が曖昧な場合（面白い情報が欲しいなど）では、1回の検索で適切な情報を得ることは難しい。

このような複数回の検索クエリ入力を含むタスクでは、ユーザの検索意図が検索の度に遷移することが考えられる。先述の旅行の例で考えると、まず行き先の候補地として近畿地方を設定して情報収集を始め、その過程でさらに候補を京都へと絞り込んだり、あるいは西日本へと広げたりするという行動は、「旅行の行き先を調べる」というユーザの検索意図が、詳細化あるいは一般化したとみなせる。また、旅行の行き先を調べ終えた後に、続いてその行き先で

¹ 京都大学大学院 情報学専攻 社会情報学専攻

² JST CREST

の観光スポットに関する情報検索を行うという行動は、旅行に関する検索の観点が「行き先」から「観光地」へと遷移したといえる。

今日の検索エンジンでは、ユーザがクエリを入力してはじめて検索が行われるが、ユーザの検索意図の遷移を検索エンジン側で検知することが可能となれば、次に入力されるクエリの候補を絞り込むことができるようになる。これによって、ユーザが次のクエリを入力する前に、意図を先読みしてクエリを推薦したり、自動検索を行うことで検索結果を動的に切り替えたり、関連情報をフィルタリングすることが可能になると考えられる。また、遷移の種類に応じて検索結果やクエリ推薦だけでなく、推薦する広告の内容を動的に変化させることも可能となる。こうした仕組みにより、各検索タスクで変化するユーザの検索意図に柔軟に対応した、より適切な情報提示手法の実現が期待できる。そのため、ユーザの検索意図の遷移をタスク実行時にオンラインで推定することは、今後の検索エンジンの発展、ならびに検索ユーザの支援に欠かせない重要要素の1つと考えられる。

本稿では、こうしたユーザの検索意図の遷移を、Web検索・閲覧時の行動ログから推定する手法を提案する。Buscherら[2]は、Webページ閲覧時におけるユーザの視線位置からの興味推定に取り組んでおり、その有効性を示している。また、マウス座標からの視線位置の推定に関する研究も行われており、検索結果ページにおいては比較的高い相関が見られるという報告がなされている[3]。これらの研究結果を踏まえると、本研究で取り組む検索意図の遷移に関しても、マウスや視線の位置が少なからず影響を与える可能性がある。例えば、検索結果ページ中では、検索意図に適合する結果ほど、ユーザの注目を引きやすいと考えられる。同様に閲覧中のWebページにおいても、どこからどこまでの領域を見ているか、ページ中の主要コンテンツと広告コンテンツのどちらに注目が集まっているかなどといった情報は、ユーザの検索意図の遷移を理解する手助けになるものと考えられる。

そこで我々は、行動ログデータの種類として、ユーザの入力したクエリや閲覧ページのURL、クリックした検索結果アイテムのランク(順位)などといった既存研究で用いられてきたデータ[4],[5]に加え、Webページ閲覧時のマウス座標や発生したクリック・スクロールイベント、さらにはユーザの視線位置といったデータに関しても、ユーザの検索行動から収集を行う。そして、得られた行動ログをもとに種々の素性を計算し、SVMを用いて分類器を構成することによって、ユーザの現在の検索クエリおよび、そのクエリにおける検索時の行動ログから、次の検索が実行される際の検索意図の遷移の種類を推定する。また、どういった素性が検索意図の遷移推定に有効に働くかなどについて議論を行う。

本稿の構成は下記の通りである。2章では、ユーザの行動ログの分析に関する研究や、視線位置検出に基づく情報探索支援に関する研究の概要を述べ、本研究の位置付けを行う。次に、3章では、検索意図の遷移の種類を述べ、本稿で扱う問題の定義を行う。4章では、まずWeb検索時におけるユーザの行動ログの収集方法を述べる。次に、得られたログデータに関して、その特性の分析を行う。5章では、得られた行動ログから素性を計算し、機械学習による分類器を構成する。そして、6章では、実際に構成した分類器に対して、その分類精度の評価を行い、最後に7章で、実験結果からの考察をもとに本研究の結論付けを行う。

2. 関連研究

ユーザの検索の効率的な支援には、検索時のユーザの意図をリアルタイムに推定し、その意図に応じた情報を提供することが重要である。こうした背景から近年では、ユーザの検索意図の推定に関する研究が盛んに行われている。本研究の位置付けとして、本章では関連する既存研究を2つの流れに沿って述べる。一方は、クエリログやクリックスルーログといった大規模なユーザの行動ログの分析および応用に関する研究であり、もう一方は、Web閲覧時における視線位置を利用した情報探索支援に関する研究である。

2.1 ユーザの行動ログ分析

Boldiら[6]や藤田ら[7]は、クエリログをもとに機械学習によって構成された分類器を用いて、連続する2つの検索クエリが与えられた際の、クエリ遷移の種類の分類を行っている。彼らの手法では、遷移前後の入力クエリを学習の素性とすることでクエリ遷移を推定している。しかし、実際のWeb検索において、次に入力されるクエリは一般的には予測不可能であるため、この手法では検索タスク中のクエリ遷移のオンライン推定を行うことができない。一方、本稿で提案する手法では、推定のための素性を、現在の検索クエリおよびその時の行動ログから計算する。そのため既存手法に比べて、より適用範囲が高いと考えられる。

Guoら[8]は、クエリログに加え、スクロールやクリックといったユーザのインタラクションデータも用いることにより、ユーザの検索タスクが情報収集指向であるのか、購買指向であるのかを分類する手法を提案している。本研究は、ユーザの行動ログデータをもとにユーザの検索行動を分類するという点では、彼らの研究と類似しているが、情報収集や購買に比べより一般的な検索意図の遷移という問題を扱う。

クエリログやクリックスルーデータを用いた分類や推定に関する研究は他にも多数行われており、例えば検索エンジンの切り替え[4]、検索時のユーザのストレス診断[5]などが挙げられる。こうした研究で用いられている素性や得られた知見を、本研究にも適用することで、より高精度な

遷移タイプの分類が可能になることが期待される。

2.2 視線検出技術を利用した情報探索支援

これまで、上述のとおり検索エンジン側で記録された大規模なクエリログなどを分析することで、Web 検索時におけるユーザの行動分析やそれに基づく支援手法に関する研究が行われてきた。しかし近年では、視線検出装置を用いて、ユーザの検索行動に関するより粒度の細かい分析などが盛んに行われるようになってきた citecutrell-chi07。さらに、検出機能を搭載した PC のプロトタイプの公開^{*1}や、市販のウェブカメラを利用した視線検出技術の開発^{*2}、さらには、マウス座標からの視線位置の推定に関する研究 [9] が、こうした分野の追い風となっており、今後ますます発展していくものと思われる。

Buscher ら [10] は、ユーザの視線位置を利用してクエリ拡張のためのキーワードを抽出する手法を提案した。彼らはこの手法の中で、リーディングやスキミングといったユーザの閲覧パターンを判定し、これらの特徴をキーワード抽出に利用している。また、我々もこれまでに、視線検出装置を用いた情報探索支援システムを提案してきた [11]。提案システムは、Web 検索時にユーザが注目している単語から検索意図を推定し、その意図に応じて検索結果の再ランキングや、Web ページ中の適合情報の強調を行う。これによって、既存の情報検索に比べ、より効率的な情報収集が可能になることが期待される。

本稿では、検索意図の遷移の分類を提案するが、その応用例の 1 つに遷移の種類に応じた動的なクエリ推薦がある。こうした推薦クエリの生成にあたって、既存研究で行われているような視線位置からの興味の推定が応用できると考えられる。

3. 検索意図の遷移タイプ

Boldi らの研究 [6] では、連続する 2 つの検索クエリの遷移の種類として (i) Generalization (G), (ii) Specialization (S), (iii) Parallel Move (P), (iv) Mission Change (M), および (v) Error Correction (E) の 5 種類を挙げている。本稿では、検索意図の遷移の種類も上記の 5 つからなると仮定し、以降ではこれらを遷移タイプと呼ぶ。

Generalization とは、ユーザが検索空間を広げる際に生じる遷移タイプである。例えば検索クエリが“京都 和食 おすすめ”から“京都 ごはん おすすめ”に変化した時は、「京都でおすすめのごはん」という条件が「京都でおすすめのごはん」へと緩められたと考えられる。このような意図の遷移を Generalization と呼ぶ。Specialization は Generalization と対をなす遷移タイプであり、ユーザが現在の検索範囲をさらに絞り込む際に生じる。検索クエリが

“京都 観光 お寺”から、より具体的な“京都 観光 清水寺”に変化した時などがこれに相当する。Parallel Move は、検索のコンテキストは変えないままで、別の観点から検索を行うといった際に生じる。この遷移タイプの具体例としては、旅行の計画のために宿泊先のホテルについて調べていたユーザが、続いて観光地を調べるという状況が挙げられる。Mission Change は、Parallel Move よりも広い概念であり、移動後の検索空間が現在とは大きく異なる際の遷移タイプと定められる。閲覧中のページで現在の検索対象とは関係のない広告の内容が気になり、それに関する情報を検索してしまうという行為は Mission Change の一例である。Error Correction は、スペルミスの修正や別表現への書き換えのように、検索クエリは変わるもののユーザの検索意図そのものは変わらない状況を指す。

上記の遷移タイプ集合を $T = \{G, S, P, M, E\}$ 、ユーザが入力するクエリ集合を $Q = \{q_i\}_{i=1}^{|Q|}$ 、クエリ q_i で検索する際の行動ログを $L(q_i)$ 、クエリが q_i から q_{i+1} へと変更される際の遷移タイプを $\text{Transit}(q_i, q_{i+1}) \in T$ とおくと、本研究で取り組む問題は、現在のクエリ q_{current} から次に入力されるクエリ q_{next} への遷移タイプ $t = \text{Transit}(q_{\text{current}}, q_{\text{next}})$ を、 q_{current} 自身および、そのクエリにおける検索時の行動ログ $L_{q_{\text{current}}}$ から推定することと定義できる。本稿では、この問題を機械学習による分類問題として扱い、Web 検索時におけるユーザの長期的な行動ログから、素性を計算し分類器を構成することで、その解決を試みる。

次章では、本問題を解くにあたって収集した Web 検索時におけるユーザの行動ログデータに関して、その収集方法および得られたデータの特性等を述べる。

4. 行動ログの収集と分析

4.1 行動ログ収集の流れ

検索意図の遷移の推定に向けて、まず Web 検索時のユーザの行動ログの収集を行う必要がある。1 章でも述べたように、本研究では通常のクエリログに加えて、ユーザのマウス操作や視線の動きに関する情報に関しても収集する。記録するログの一覧を表 1 に示す (各ログには、その発生時刻も付与される)。なお、ユーザの視線位置に関しては、Tobii 社^{*3}の視線計測装置 Tobii T60 およびソフトウェア開発キット Tobii SDK を用いて測定を行う。マウスおよび視線のモニタ上における座標の取得間隔に関しては、既存研究 [3] を参考にそれぞれ 100ms, 25ms とする。このようにして収集した検索結果、閲覧ページ、マウスおよび視線に関するログデータを、図 1 に示すように検索クエリごとのグループに分割を行う。そして連続する検索クエリ q_i, q_{i+1} に対する行動ログ L_{q_i} と $L_{q_{i+1}}$ に対して、その遷移タイプ $t = \text{Transit}(q_i, q_{i+1})$ を、図 2 に示すインタフェース

^{*1} <http://news.mynavi.jp/news/2011/03/02/018/index.html>

^{*2} <http://www.inference.phy.cam.ac.uk/opengazer/>

^{*3} <http://www.tobii.com>

表 1 Web 検索時に収集するログの種類一覧

カテゴリ	ログ項目
検索	検索クエリ, ヒットカウント, 表示範囲変化, ページ HTML, ページサイズ, 検索結果リスト, 検索結果表示位置, 検索結果クリック, クエリ推薦リスト, クエリ推薦表示位置, クエリ推薦クリック, 広告リスト, 広告表示位置, 広告クリック
閲覧	閲覧ページ, 閲覧 URL, ページ HTML, ページサイズ
マウス	検索閲覧切替, カーソル位置, クリック位置, スクロール位置, ホバー単語 (100ms ごとに座標取得)
視線	視線位置, 注目単語 (25ms ごとに座標取得)

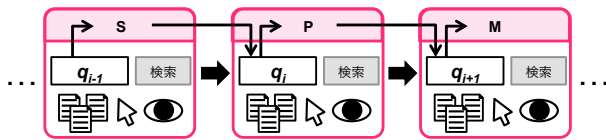


図 1 行動ログデータの分割およびラベリング



図 2 行動ログに対する正解ラベル付与プログラムのインターフェース

でユーザにラベリングしてもらい、その値を L_{q_i} の正解ラベルとすることで、分類器構成のためのデータ処理を行う。例えば、「京都 観光地」というクエリに続いて「京都 観光地 お寺」というクエリで検索が行われていると、前者の行動ログに対して付与される正解ラベルは Specialization となる。ただし、以降の検索クエリが存在しない行動ログに対しては、その正解ラベルを Mission Change として扱う。

4.2 収集したログの分析

前節で述べた収集方法に従って、Web 検索時のユーザの行動ログの収集およびラベリングを行った。今回は主に、著者の 1 人である梅本の行動ログを対象として、分類器の構成を試みる。2012 年 4 月 4 日, 25 日, 5 月 4 日, 6 日, 7 日の計 5 日間にわたって、検索タスクを明示的に与えることはせずに、検索時の行動ログデータの収集を行った。得られた総クエリ数は 183 であった。実際のログに含まれていた複数クエリからなる検索タスク、およびそのタスクで用いられたクエリ遷移の例を表 2 に示す。さらに、収集した行動ログに対して付与された遷移タイプの内訳を、Boldi

表 3 ログデータ中の遷移タイプの内訳。左は本研究で収集した行動ログ、右は商用検索エンジンのイギリス国内およびアメリカ国内におけるクエリログ（それぞれ UK, US）を用いた既存研究 [6] における結果を示している。

遷移	内訳 (M を除いた内訳)	遷移	UK 内訳	US 内訳
G	13.1% (15.1%)	G	4.4%	9.5%
S	36.6% (42.1%)	S	37.5%	30.1%
P	29.5% (34.0%)	P	47.7%	55.5%
E	7.7% (8.8%)	E	10.4%	5.0%
M	13.1% (—)	M	—	—
総数	0.2k	総数	6.5M	10.5M

ら [6] が行った、商用検索エンジンのイギリス国内およびアメリカ国内におけるクエリログ（それぞれのデータセットを UK, US とする）を用いた既存研究の結果とともに表 3 に示す。

今回収集した行動ログデータでは、遷移タイプが Specialization または Parallel Movement である行動ログが全体の 66.1% を占めていた。更に、遷移タイプが Mission Change である行動ログを除いた際（既存研究におけるログデータをセッション分割したものに相当）の上タイプの占める割合は、76.1% に上昇し、この結果は既存研究の 85.2% および 85.6% に近い値となった。また、Generalization や Error Correction を遷移タイプとする行動ログの割合が低いという結果も、既存研究のものの一貫性があるといえる。両者の違いとしては、既存研究では S よりも P の割合が高いが、本研究で用いるログデータに関してはそれが逆転していることが挙げられる。これは、今回の行動ログの収集にあたって、表 2 に示したように、ある事柄に関する情報収集という検索タスクを行うことが多かったことに起因すると考えられる。しかしながら、両者の差はそこまで大きくないといえる。以上の簡易的な比較より、今回収集した行動ログデータは総数は少ないものの、十分に実際の検索データの分布と一貫性があると我々は判断した。そこで、以降ではこの行動ログデータを用いて、検索意図の遷移分類を実現する分類器の構成を行う。

5. 遷移タイプの分類

5.1 素性候補の検討

本稿では、機械学習による分類問題において広く用いら

表 2 収集したログデータに含まれていた検索タスクおよびそのときのクエリ遷移の例。
遷移間の矢印上の記号は、その遷移に対して付与された遷移タイプを示す。

検索タスク	クエリ遷移例
就職活動の情報収集	“NTT ドコモ グループディスカッション” \xrightarrow{E} “NTT ドコモ グループ 面接” \xrightarrow{P} “NTT データ 二次面接”
博士課程への進学	“ドクター 進学” \xrightarrow{S} “ドクター 進学 就職” \xrightarrow{P} “ドクター 進学 決めて” \xrightarrow{E} “ドクター 進学 決め手”
ヨーロッパ旅行	“イタリア カプリ島” \xrightarrow{S} “イタリア カプリ島 青の洞窟” \xrightarrow{S} “イタリア カプリ島 観光スポット”
PCトラブル	“パソコン 排熱” \xrightarrow{S} “パソコン 温度 高い” \xrightarrow{P} “パソコン 温度 speecy” \xrightarrow{E} “パソコン 温度 speccy”
プログラミング言語	“型付け” \xrightarrow{S} “動的型付け 強い 弱い” \xrightarrow{P} “python オブジェクト指向” \xrightarrow{S} “python オブジェクト指向 継承”

れている SVM を、分類器の構成手段に採用する。分類器の構成にあたって、まず得られた行動ログからどのような特徴量を分類のための素性として用いるかを考察する必要がある。そこで、本節では分類器構築のための素性の検討を行う。我々は得られた行動ログから計算できる素性群として、大きく3つのカテゴリを用いる。以下では、各カテゴリに関して、それに属する具体的な素性候補を述べる。

5.1.1 クエリ・クリック素性群

クエリ・クリック素性群は、「クエリの単語数」や「ヒットカウント」、「クリックされた検索結果数」、「直前の行動ログの遷移タイプ」といったような、ユーザが入力したクエリおよびそのクエリに関する検索全体から得られる特徴量のカテゴリである。

現在の検索対象の広さや粒度は、ユーザの検索の絞込みの度合いを間接的に表していると考えられる。例えば、5単語で構成されるクエリによる検索結果は、一般的に2単語からなるクエリよりも、ユーザの検索結果に対する絞込みの意図が強く現れているとみなせる。同様に、ヒットカウントの少ない検索結果が生じる要因の1つとして、ユーザが特定の限られた情報を検索しようとしてクエリを入力することが考えられる。また、クエリログ分析の研究 [6] では、Specialization と Generalization は交互に発生しやすい、という報告がなされている。これらを総合すると、クエリの語数が少ない、あるいは、ヒットカウントの多い検索からの遷移タイプとしては、Specialization が生じる可能性が高くなることが期待される。対照的に、長いクエリや少ない検索結果は、検索結果の絞込みを示唆するため、その後の検索への遷移タイプは Generalization となる確率が高くなると考えられる。同様に、直前の検索における遷移タイプも、後続の検索時の遷移タイプの推定にあたって、少なからず影響を与えるであろう。

クリックスルーに関する既存研究 [12] で述べられているように、検索結果のクリック数は、その検索クエリに対するユーザの満足度の尺度として利用することができる。もし、どの検索結果もクリックされずに次の検索に移動した際には、ユーザはより良い情報を求めてクエリ修正をするのではないかと考えられる。そのため、全く別の検索へと遷移する Mission Change は比較的発生しにくいものと予想される。

5.1.2 検索素性群

このカテゴリには、あるクエリに関する各検索結果ページ（以降 SERP と呼ぶ）、およびそのページ上でのユーザの行動から得られる特徴量が属する。具体的には、「何件目までの SERP を表示したか（検索結果表示件数）」、「SERP 閲覧時間」、さらには「広告、クエリ推薦、検索結果に対する注目度合い」などが考えられる。

SERP を何ページ見たかや何秒見たかといった指標は、その検索タスクの難易度や性質を表す一要因といえる。例えば、Navigational な検索タスク [13] においては、ある1件の目的ページを見つければ完了するため、ユーザの検索結果に対する従事度合いはそれほど高くないことが予想される。さらに、この後で検索が行われる場合は、既に目的ページを訪れるというユーザの情報要求は達成されているため、次の遷移タイプが Mission Change となる確率が通常時よりも高くなるはずである。

同様に、SERP 中のどの領域によく目が向けられているかといったことも検索タスクの種類によって変わってくる。実際に既存研究 [8] において、商品購入のための検索に代表される Transactional タスク [13] では、他のタスクに比べて広告のクリックスルーレートが上昇するという実験結果が報告されている。このような事例では、ある商品の購入のために訪れた通販サイトで全く別の商品に興味を惹かれ、その後、その商品に関する検索を行うことも十分起こりうると予想される。そのため注目領域の同定も、遷移タイプの推定に役立つ可能性があるといえる。

5.1.3 閲覧素性群

これは、ユーザが閲覧した（SERP を除く）Web ページから得られる特徴量のカテゴリである。これに属するものとして、「Web ページ閲覧時間」や「視線の移動量（軌跡の長さ）」、「ページ中において注目していた領域」などが考えられる。

訪れた Web ページに対して、その中の情報をよく読んでいる場合は、そうでない場合に比べてそのページの適合性が高いと考えられる。そのため、どの程度の時間を Web ページ閲覧に費やしたかは、現在の検索に対する満足度の近似といえる。ページ中のどこからどこまでを読んだか、といった指標に関しても同様の議論ができる。現在の検索

に対して一定の満足度が得られると、別の検索タスクへの遷移が予想されるので、こうした指標は Parallel Movement や Mission Change の推定に有用となる可能性がある。

また、検索結果の時と同様、Web ページ中のどの部分に注目しているかといった情報も重要と考えられる。Web ページは 2 もしくは 3 ペインで構成されていることが多く、そうしたページでは、ページ中の左右には外部サイトの広告あるいはサイトナビゲーション用のメニュー項目が表示されている。Web ページ中の広告に対する注目は、別ドメインへの興味を表しているとも考えられるため、次の遷移タイプが Mission Change となる確率が高くなるかもしれない。

5.2 分類器の構成

前節での考察をもとに、本研究では分類器構築にあたって、表 4 に示す素性群を採用する。本研究が対象とする問題の出力集合は、3 章でも述べたように、遷移タイプ集合 $T = \{G, S, P, M, E\}$ であるため、本問題は多クラス分類問題である。今回は、本問題を解くための分類器構成にあたって、既存の SVM ライブラリである LIBSVM^{*4} を使用する。LIBSVM では、遷移タイプの組み合わせ (G と S など) ごとの分類器を組み合わせることで上記の多クラス分類器を構成する。また、分類に用いるカーネルとしては RBF カーネルを選択し、カーネルのパラメータは LIBSVM のデフォルト値を採用する。以降では、本章で述べた方法により構成した分類器に対して、その分類精度に関する評価および考察を行う。

6. 評価と考察

我々は前章で述べた方法をもとに遷移タイプを推定する分類器を構成した。本章では構成した分類器に関して、実際の行動ログに対する遷移タイプの分類精度の評価および分類に有効に働く素性の検証を行う。

6.1 分類精度

まずはじめに、収集した行動ログデータ (183 クエリ) に対して leave-one-out cross validation (データのうち、1 つをテストデータとし、残りを訓練データとすることを各組み合わせに対して行う手法) を用いて、その分類精度を評価した。精度評価のベースラインとしては、各遷移タイプを等確率で毎回ランダムに決定する手法 (Baseline_{rand})、遷移確率の最も高いタイプを常に出力する手法 (Baseline_{dom}) の 2 つを考える。分類器の精度の評価尺度としては、Accuracy を用いる。ここで、Accuracy はテストデータのうち、その遷移タイプを正しく分類することのできた割合と定義される。Baseline_{rand} の分類精度は、その性質上 0.2 である。

表 4 分類器の構成に用いる素性一覧

クエリ・クリック素性群	
[f ₁]: 前ラベルが G か, [f ₂]: 前ラベルが S か, [f ₃]: 前ラベルが P か, [f ₄]: 前ラベルが M か, [f ₅]: 前ラベルが E か, [f ₆]: クエリ単語数, [f ₇]: クエリ文字数, [f ₈]: ヒットカウント, [f ₉]: クエリ推薦数, [f ₁₀]: 広告数, [f ₁₁]: $f_{12} + f_{13} + f_{14}$, [f ₁₂]: 検索結果クリック数, [f ₁₃]: クエリ推薦クリック数, [f ₁₄]: 広告クリック数, [f ₁₅]: 検索閲覧切替数, [f ₁₆]: f_{15}/f_{11} , [f ₁₇]: $f_{12} > 0$, [f ₁₈]: $2^{-(\text{初クリック検索結果ランク})}$, [f ₁₉]: Avg. $2^{-(\text{クリック検索結果ランク})}$, [f ₂₀]: $f_{13} > 0$, [f ₂₁]: $2^{-(\text{クリッククエリ推薦ランク})}$, [f ₂₂]: $f_{14} > 0$, [f ₂₃]: $2^{-(\text{初クリック広告ランク})}$, [f ₂₄]: Avg. $2^{-(\text{クリック広告ランク})}$,	
検索素性群	
[f ₂₅]: 検索結果表示件数, [f ₂₆]: SERP 遷移数, [f ₂₇]: SERP 総閲覧時間, [f ₂₈]: Avg. SERP 閲覧時間, [f ₂₉ , f ₃₀ , f ₃₁]: Avg., Max., Min. SERP マウス移動量, [f ₃₂ , f ₃₃ , f ₃₄]: Avg., Max., Min. SERP 視線移動量 [f ₃₅ , f ₃₆ , f ₃₇]: Avg., Max., Min. 検索結果注目率, [f ₃₈ , f ₃₉ , f ₄₀]: Avg., Max., Min. クエリ推薦注目率, [f ₄₁ , f ₄₂ , f ₄₃]: Avg., Max., Min. 広告注目率	
閲覧素性群	
[f ₄₄]: Web ページ遷移数, [f ₄₅]: Web ページ総閲覧時間, [f ₄₆]: Avg. Web ページ閲覧時間, [f ₄₇ , f ₄₈ , f ₄₉]: Avg., Max., Min. Web ページマウス移動量, [f ₅₀ , f ₅₁ , f ₅₂]: Avg., Max., Min. Web ページ視線移動量 [f ₅₃ , f ₅₄ , f ₅₅]: Avg., Max., Min. Web ページ左部注目率, [f ₅₆ , f ₅₇ , f ₅₈]: Avg., Max., Min. Web ページ右部注目率, [f ₅₉ , f ₆₀ , f ₆₁]: Avg., Max., Min. Web ページ上部注目率, [f ₆₂ , f ₆₃ , f ₆₄]: Avg., Max., Min. Web ページ下部注目率, [f ₆₅ , f ₆₆ , f ₆₇]: Avg., Max., Min. Web ページ中央部注目率	

表 5 提案手法およびベースラインの分類精度比較

提案手法	Baseline _{rand}	Baseline _{dom}
0.410	0.200	0.366

また、Baseline_{dom} の分類精度については、今回収集した行動ログデータの遷移タイプ分布から 0.366 とした。各手法の分類精度をまとめたものを表 5 に示す。

提案手法の分類精度は 0.410 であり、ベースライン 2 つのそれを上回る結果となった。しかしながら、Baseline_{dom} との差は大きいものではなかった。その原因の一つとして、今回収集したログデータの数の少なさが挙げられる。収集するログデータ数の増加によって、分類精度に向上が見られるかどうかを検証するために、我々は異なる訓練データ数に対する分類器の精度比較を行うことにした。以下でその手順および結果を述べる。

まず収集した 183 クエリの全行動ログデータから事前にテストデータとしてその 10% (18 個) を抽出した。抽出においては、元々の各遷移タイプの分布割合に沿って、サンプリングを行った。次に残りの 165 クエリに関する行動ログデータから、30, 60, 90, 120, 150 個をランダムサンプリングした際の、それぞれにおける分類精度を比較した。ただし、どのデータをサンプリングするかによって、分類精度が大きく異なる可能性がある。そのため、各サンプルサイズに対して分類精度の計算を 5 回行った。その結果を

*4 <http://www.csie.ntu.edu.tw/~cjlin/libsvm/>

表 6 異なるサンプルサイズに対する分類精度比較

反復回数	サンプルサイズ				
	30	60	90	120	150
1 回目	0.33	0.39	0.39	0.44	0.44
2 回目	0.33	0.33	0.33	0.33	0.44
3 回目	0.33	0.33	0.44	0.33	0.44
4 回目	0.28	0.33	0.33	0.39	0.38
5 回目	0.33	0.33	0.44	0.44	0.44
Avg.	0.32	0.34	0.39	0.39	0.43
SD	0.02	0.02	0.06	0.06	0.02

表 7 素性を除去した際の分類精度比較.

ここで全素性をういた場合の精度は, 0.410 である.

除去前の精度に対する減少率の高い (分類寄与度の高い) 素性		
f_{17} : 0.393	f_{54} : 0.399	$f_{3,6,46}$: 0.404
除去前の精度に対する減少率の低い (分類寄与度の低い) 素性		
f_1 : 0.432	$f_{9,66}$: 0.421	$f_{5,10,29,30,31,50,65}$: 0.415

表 6 に示す.

この結果から, サンプルデータによって多少のばらつきはあるものの, 一般的にはサンプルサイズが大きい時ほど, 分類精度が高いことがわかる. この結果は, 収集する行動ログデータのサイズが大きくなるほど, 分類精度の上昇が期待できることを示唆している. そのため, 行動ログをさらに蓄積し続けることで, 提案手法とベースラインの分類精度の差はさらに広げられると考えられる.

6.2 有用な素性の検証

次に, 分類器の構成に用いた素性の中で, 遷移タイプの推定に大きく寄与しているものを明らかにする. そのため, 我々は表 4 に示した各素性に対して, その素性を除いて構成した分類器の精度の比較を行った. ここでも分類精度の評価においては, leave-one-out cross validation を用いた. その結果について, 精度の減少率が大きかった素性および, 最も小さかった素性を表 7 に示す. なお, この表に示していない素性に関しては, 分類精度がその素性を除く除かないにかかわらず変化しなかった.

表 7 から, 検索結果を 1 つでもクリックしたか, Web ページの左側に注目していた割合, 前回の遷移ラベルが P であるか, 検索クエリを構成する単語数, Web ページを閲覧していた時間, という 5 つの素性に関しては, その素性を除いた際の分類精度が, 全ての素性をういた際の精度に比べて, 減少していることがわかる. この結果は, 上の 5 つの素性は他の素性に比べて, 遷移タイプの分類に対する寄与度が高いということを示唆している.

Web ページの左側には他サイトへの広告リンクが表示されていることがよくある. 前章の素性検討の項目では, こうした領域への注目が Mission Change を引き起こすと述

べていた. 今回の結果からは, 後続する遷移タイプが何かまではわからないものの, この素性が遷移タイプの分類に大きく貢献するということが示されたといえる. 上述の他の素性に関しても, 同様の議論を行うことができる. 例えば, 検索クエリの単語数が多いか少ないかといった情報が加わることで, 現在の検索対象の範囲が大まかに推定できるため, 次に行われる検索への遷移タイプの予想がしやすくなると考えられる.

このような分類寄与度の高い素性の存在とは対照的に, いくつかの素性に関しては, それを取り除くと分類精度が元の値よりも上昇するという結果が得られた. こうした素性の事例として, 前回の遷移タイプが G か, SERP 上に表示される推薦クエリの数, Web ページ中の中央部を注目していた割合, SERP 上におけるマウスの移動量などが挙げられる. 今回の結果からは, これらの素性は分類器の構成にあたって, 分類精度の妨げになってしまうということがいえる. 既存研究 [6] において, G と S が交互に起きるという報告があったことも手伝い, 我々は直前の遷移タイプが G であることは, 次の遷移タイプ (特に S) の推定に有効に働くことを期待していた. しかしながら, 今回収集した行動ログから構成した分類器においてはその傾向は見られなかった. 既存研究の結果との相違が生じた要因として, 収集した行動ログの遷移タイプの分布が考えられる. 表 3 を見ると, 既存研究と比べた際に遷移タイプ G の生じる割合が比較的高いことがわかる. これは今回, 収集した行動ログ中で行われていた検索タスクが複数クエリからなる事例が多く, そのようなタスクにおいて G と S が通常のクエリログにおける分布よりも高い確率で生じていたからであると考えられる. そのため, 直前の遷移タイプが G である行動ログにおいて, 次の検索への遷移タイプのエントロピーが高くなり, そのことが分類精度に負の影響を与えた可能性がある. この素性が本当に有効かどうかを検証するためにも, 今後も行動ログを収集し, 再度分析を行う必要があると我々は考える.

また, Web ページの左側への注目率が遷移タイプの分類に貢献したのに対して, 中央部への注目に関する素性の追加は, 分類精度の悪化という結果をもたらすといえる. 上述の遷移タイプ G に関する考察と同様, Web ページの中央部がよく注目されているという情報からは, 後続する検索への遷移タイプの同定が困難であることを示している. そのため, この素性は遷移タイプの分類という問題に関しては, あまり適切なものではなかったと考えられる.

7. まとめと今後の課題

本稿では, Web 検索時におけるユーザの検索意図の遷移を推定する手法を提案した. 提案手法は, ユーザの長期的な行動ログデータから機械学習によって分類器を構成することで, 遷移推定を行う. 我々は, 実際に収集した行動口

グデータから分類器を構成し、その分類精度の評価および遷移推定に有効な素性の検証を行った。実験結果から、5種類の遷移タイプに対して約40%の精度で分類が可能であることが明らかになった。しかしながら、実験で用いた行動ログの数は、一般的な行動分析で用いられているものと比較すると少ない。

今後は、行動ログの収集数を増やすことで、遷移タイプの分類精度がどこまで向上するか検証を行う。また、今回の手法で機械学習に用いた素性は、クエリの単語数やページ閲覧時間、ページ中の注目位置など、計算は容易なもの比較的シンプルなものが多かった。しかしながら、本研究で行動ログとして記録している視線情報からは、注目単語とクエリの関係や、ページを読む速度、視線の動きのパターンなど、よりリッチな特徴量も計算することができる。そこで、こうした特徴量も素性に加えることで分類精度の向上に努めていきたい。

謝辞 本研究の一部は、文部科学省科学研究費補助金若手研究(A)「インタラクティブな再ランキング・再サーチを可能とする次世代検索に関する研究」(研究代表者:中村聡史, 課題番号23680006), 挑戦的萌芽研究「モバイル協調検索に関する研究」(研究代表者:中村聡史, 課題番号22650018)によるものです。ここに記して謝意を表すものとします。

参考文献

- [1] Huang, J. and Efthimiadis, E. N.: Analyzing and evaluating query reformulation strategies in web search logs, in *Proceedings of the 18th ACM conference on Information and knowledge management*, CIKM '09, pp. 77-86, ACM (2009).
- [2] Buscher, G., Dengel, A. and Elst, van L.: Query expansion using gaze-based feedback on the subdocument level, in *Proceedings of the 31st annual international ACM SIGIR conference on Research and development in information retrieval*, SIGIR '08, pp. 387-394, ACM (2008).
- [3] Huang, J., White, R. W. and Dumais, S.: No clicks, no problem: using cursor movements to understand and improve search, in *Proceedings of the 2011 annual conference on Human factors in computing systems*, CHI '11, pp. 1225-1234, ACM (2011).
- [4] White, R. W. and Dumais, S. T.: Characterizing and predicting search engine switching behavior, in *Proceedings of the 18th ACM conference on Information and knowledge management*, CIKM '09, pp. 87-96, ACM (2009).
- [5] Feild, H. A., Allan, J. and Jones, R.: Predicting searcher frustration, in *Proceedings of the 33rd international ACM SIGIR conference on Research and development in information retrieval*, SIGIR '10, pp. 34-41, ACM (2010).
- [6] Boldi, P., Bonchi, F., Castillo, C. and Vigna, S.: From "Dango" to "Japanese Cakes": Query Reformulation Models and Patterns, in *Proceedings of the 2009 IEEE/WIC/ACM International Conference on Web Intelligence*, WI '09, pp. 183-190, IEEE Computer Society (2009).
- [7] 藤田遼治, 太田学, 徳永徹郎: ユーザのクエリ変更意図に基づく先読み検索, 第4回データ工学と情報マネジメントに関するフォーラム, DEIM '12 (2012).
- [8] Guo, Q. and Agichtein, E.: Ready to buy or just browsing?: detecting web searcher goals from interaction data, in *Proceedings of the 33rd international ACM SIGIR conference on Research and development in information retrieval*, SIGIR '10, pp. 130-137, ACM (2010).
- [9] Guo, Q. and Agichtein, E.: Towards predicting web searcher gaze position from mouse movements, in *Proceedings of the 28th of the international conference extended abstracts on Human factors in computing systems*, CHI EA '10, pp. 3601-3606, ACM (2010).
- [10] Buscher, G., Elst, van L. and Dengel, A.: Segment-level display time as implicit feedback: a comparison to eye tracking, in *Proceedings of the 32nd international ACM SIGIR conference on Research and development in information retrieval*, SIGIR '09, pp. 67-74, ACM (2009).
- [11] Umemoto, K., Yamamoto, T., Nakamura, S. and Tanaka, K.: Search Intent Estimation from User's Eye Movements for Supporting Information Seeking, in *Proceedings of the 11th International Working Conference on Advanced Visual Interfaces*, AVI '12, ACM (2012).
- [12] Joachims, T.: Optimizing search engines using click-through data, in *Proceedings of the eighth ACM SIGKDD international conference on Knowledge discovery and data mining*, KDD '02, pp. 133-142, ACM (2002).
- [13] Broder, A.: A taxonomy of web search, *SIGIR Forum*, Vol. 36, No. 2, pp. 3-10 (2002).