

複数の Web サイトにおけるユーザ滞在時間の解析

本間 涼介[†] 副島 啓一^{††} 吉田 光男[†] 梅村 恭司[†]

[†] 豊橋技術科学大学 〒441-8580 愛知県名豊橋市天伯町雲雀ヶ丘 1-1

^{††} 株式会社 Faber Company 〒107-0052 東京都港区赤坂 2-14-4 森崎ビル 5 階

E-mail: [†]r173360@edu.tut.ac.jp, ^{††}soejima@fabercompany.co.jp, [†]yoshida@cs.tut.ac.jp, [†]umemura@tut.jp,

あらまし Web ページをより良いものにするためにユーザの閲覧行動を利用する方法が考えられる。しかしながら、滞在時間そのものの特性は良く知られていない。本研究ではニュース記事ではない 13 の Web サイトについて、1 年分の閲覧ログデータを使用し、デスクトップとモバイルデバイスの両方を対象として解析を行なった。解析の結果、デスクトップとモバイルの両方で、滞在時間が長くなるほどセッション数は減少することが分かった。Web サイトがどちらかのデバイスのみに対してデザインされているとき、滞在時間の中央値はデスクトップとモバイルでは大きく違うことが分かった。また、アクセスする時間帯と月は滞在時間に大きく影響するが、曜日はあまり影響しないことが分かった。内部ユーザとクリックユーザはそれぞれ外部ユーザとクリックしないユーザより滞在時間が長くなる傾向が見られた。

キーワード 滞在時間, データマイニング

1 はじめに

Web ページをより良いものにするために、ユーザの閲覧行動を利用する方法が考えられる。滞在時間は Web ページにおけるユーザの行動指標の 1 つであり、ユーザがページをどれくらい長く見ているかを表す。滞在時間はユーザ評価が重要な分野、例えば検索エンジン [1, 2], コンテンツ推薦システム [3], 広告 [4-6] で使用されている。また、Kazuyuk ら [7] は Web ページに対するユーザの満足度の指標として滞在時間を使用している。しかし、この指標が重要であるにも関わらず、その特性はよく知られていない。Liu ら [8] は滞在時間の特性を解析しているが、2 週間分のデータのみを使用している。Vasiloudis ら [9] はメディア視聴における滞在時間（セッション時間）を予測しているが、ある 1 つのメディアサイトのみを扱っている。また、ニュースサイトにおける滞在時間の傾向はデスクトップとモバイルデバイスで大きく異なることが報告されている [3, 4, 10]。しかし、ニュースサイトではない Web サイトにおける研究はほとんど無い。

本研究では、複数の Web サイトにおいてデスクトップとモバイルデバイスの両方について 1 年分のログデータを解析する。本研究の目的は、滞在時間の予測に効果的な特徴量を発見するために、ニュースサイトではない Web サイトにおいて滞在時間の特性を明らかにすることである。この目的を達成するために以下の解析項目に取り組む。以下の解析項目は、使用した 4 種類のログデータから解析が可能な内、滞在時間予測の特徴量として利用出来そうなものを解析する目的で選んだ。

RQ1: 滞在時間はデバイスの種類により差が出るか？

RQ2: 滞在時間はアクセスする時間帯、曜日、月によって差が出るか？

RQ3: 滞在時間は内部ユーザと外部ユーザで差が出るか？

RQ4: 滞在時間とクリックに関係はあるか？

RQ5: 滞在時間とスクロール深度に関係はあるか？

解析の結果、デスクトップとモバイルデバイスの両方で、滞在時間が長くなるほどセッション数は減少することが分かった。Web サイトがどちらかのデバイスのみに対してデザインされているとき、滞在時間の中央値はデスクトップとモバイルデバイスでは大きく違うことが分かった。また、アクセスする時間帯と月は滞在時間に大きく影響するが、曜日はあまり影響しないことが分かった。さらに、時間帯による滞在時間の変化がセッション数と同じように変化する傾向が見られた。内部ユーザとクリックユーザはそれぞれ外部ユーザとクリックしないユーザより滞在時間が長くなる傾向が見られた。しかし、滞在時間とスクロールの関係を見出すことは出来なかった。

2 使用データ

2.1 使用したログデータ

本研究では、2017 年 7 月から 2018 年 8 月までの期間で 13 の日本語 Web サイトを使用した。このデータは MIERUCA¹ を使用して集められ、これはユーザの閲覧行動を可視化する機能を持った Web マーケティングツールである。MIERUCA は日本でよく知られている検索エンジン最適化 (SEO) ツールであり、株式会社 Faber Company² によって提供されている。使用したデータは 4 種類のログで構成されており、それぞれ referrer, click, scroll, read ログである。referrer ログは Web ページにアクセスするとすぐに記録され、デバイスの種類 (デスクトップ, タブレット, モバイル), ウィンドウサイズ (横幅と高さ), アクセス元 URL, Web ページの長さなどを含む。Web ページの長さはユーザが使用しているデバイスによって

1: <https://mieru-ca.com/heatmap/>

2: <https://www.fabercompany.co.jp/>

表 1 解析に使用した Web サイト

#	Data Period	Site Type	Mobile	Desktop			Mobile		
				Sessions	Time (s)	Clicks	Sessions	Time (s)	Clicks
A	2017/07/01 ~ 2018/08/21	Business		117,329 (74.1%)	7	0.70	40,947 (25.9%)	8	0.63
B	2017/07/02 ~ 2018/08/21	Blog	○	10,291 (35.2%)	21	0.12	18,914 (64.8%)	23	0.13
C	2018/04/11 ~ 2018/08/21	Blog	○	192,879 (30.5%)	47	0.18	439,695 (69.5%)	72	0.18
D	2017/08/02 ~ 2017/08/31	Affiliate	○	724 (41.7%)	12	0.06	1,014 (58.3%)	12	0.00
E	2017/07/01 ~ 2018/08/02	Business	○	520,792 (82.3%)	23	0.53	112,126 (17.7%)	14	0.27
F	2017/07/01 ~ 2018/08/15	Affiliate	○	26,884 (38.8%)	7	0.09	42,425 (61.2%)	9	0.07
G	2017/07/01 ~ 2018/08/21	Affiliate	○	171,565 (25.6%)	55	0.27	498,439 (74.4%)	83	0.24
H	2017/07/01 ~ 2018/08/21	BBS		286,139 (44.6%)	7	0.77	355,166 (55.4%)	7	0.77
I	2017/12/18 ~ 2018/08/27	Affiliate		102,872 (40.6%)	26	0.04	150,399 (59.4%)	20	0.02
J	2017/07/01 ~ 2018/08/21	Blog		100,158 (70.2%)	37	0.16	42,474 (29.8%)	9	0.05
K	2017/07/01 ~ 2018/08/21	Affiliate	△	51,340 (30.2%)	9	0.14	118,591 (69.8%)	35	0.18
L	2017/07/01 ~ 2018/08/21	Affiliate	○	194,838 (27.5%)	10	0.05	514,863 (72.5%)	25	0.08
M	2017/07/01 ~ 2018/04/27	Affiliate	○	18,073 (36.4%)	9	0.02	31,515 (63.6%)	10	0.02

異なる。click ログはユーザがマウスカーソルでクリックした場所を記録する。スマートフォンのようなモバイルデバイスでは、ユーザがタップした位置が記録される。scroll ログはセッション中にユーザがスクロールをストップした位置の中で、最も深い位置の（最下部に近い）ログのみを記録する。read ログはユーザがスクロールを4秒以上ストップした場合にその位置を記録する。さらに、これら全てのログデータがユーザがアクセスした URL、ログが記録された日時、セッションキーを持つ。セッションキーはユーザがある Web ページにアクセスしてからそのページを離れるまで同一である。

各セッションの滞在時間はセッション中で最初にログデータが記録された日時と最後に記録された日時の差として計算される。本研究では主に各 Web サイトでの滞在時間を解析する。Web サイトの滞在時間はそのサイトにアクセスした全てのユーザの滞在時間の中央値とする。また、セッション集合を分割する場合は、代表値としてそのセッション集合における滞在時間の中央値を使用する。ユーザが Web ページを閲覧せずにブラウザを立ち上げ続けている場合があるため、滞在時間の代表値として平均値を使用するのは適していない。MIERUCA はタイムアウト時間を設定している。タイムアウト時間より長い時間ログデータが記録されなかった場合、それ以降ユーザが4種類のログデータのどれかに該当する閲覧行動を起こした場合もログデータは記録されない。また、デバイスの種類による差を解析するために、デスクトップデバイス（タブレットとラップトップを含む）とモバイルデバイス（主にスマートフォン）を分けて滞在時間を計算した。

2.2 基本統計

基本的な Web サイトの情報は表 1 に示す。2017 年 7 月から 2018 年 8 月までのデータを使用した。各 Web サイトのログデータ取得期間には少しの異なりがある。MIERUCA のプライバシーポリシーの制限により、Web サイト名およびそのドメイン名を示さず、各 Web サイトを A から M までのアルファベットで示す。Web サイトのタイプは Affiliate, bulletin

board system (BBS), Blog, Business に分類した。Affiliate は主に特定の製品やサービスの購入の促進を行なう Web サイトである。BBS はユーザ同士でメッセージを送り合う Web サイトである。Blog はブログ形式で情報を提供する Web サイトである。Business は企業が特定の機能や情報を提供する Web サイトである。Mobile はモバイルデザインに対応しているかどうかを表し、具体的には html の meta タグで viewport がモバイル向けに設定されているかどうかである。○は viewport がモバイル向けに設定されている Web サイトである。△は viewport がモバイル向けに設定されていないものの、デスクトップのデザインとモバイルのデザインが明確に分かれている Web サイトである。

Sessions はログデータ取得期間の間のセッション数の合計であり、Time はその Web サイトにおける全てのセッションの滞在時間の中央値であり、Clicks はその Web サイトにおける全てのセッションのクリック数の平均値である。これらの値は全てデスクトップとモバイルデバイスを分けて計算している。デバイスの割合は各 Web サイトで異なっているが、全体的にはモバイルデバイスが長い滞在時間を持つ傾向にある。しかし、デスクトップデバイスはクリック数が多い傾向がある。

3 解析結果と考察

3.1 RQ1: 滞在時間はデバイスの種類により差が出るか？

Song ら [11] は検索エンジン（例えば Bing³）における滞在時間はデバイスによって大きく異なることを報告している。本研究ではニュースサイトではない Web サイトでは滞在時間がデバイスによって異なることを分析する。

セッション数と滞在時間の関係を図 1 に示す。デスクトップデバイスとモバイルデバイスの両方で、滞在時間が長くなるほどセッション数は減少した。Liu ら [8] はデスクトップデバイスではセッション数と滞在時間が Weibull 分布に従うことを示した。今後は本研究で用いたデータにおいても、滞在時間が

3: <https://www.bing.com/>

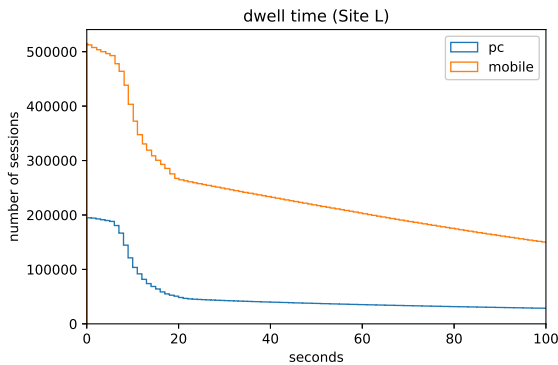


図 1 デバイス毎の滞在時間とセッション数の関係 (Web サイト L)

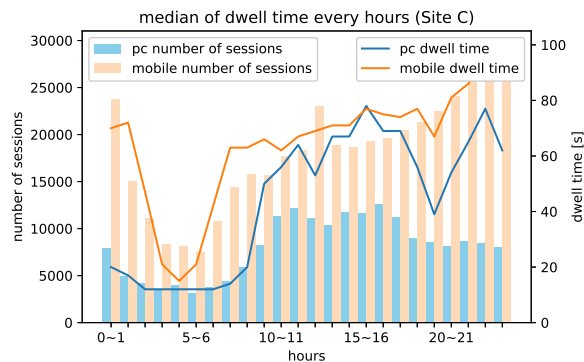
Weibull 分布に従うかどうかを検証したいと考えている。

Web サイトによって、デバイス毎の滞在時間は大きく異なっている。多くの場合、Web サイトはどちらかのデバイスのみに対してデザインされていた。表 1 において、例えば、Web サイト J はデスクトップの滞在時間は 37 秒であるが、モバイルの滞在時間は 9 秒であった。言い換えると、デスクトップユーザは Web サイト J により長く留まっている。この Web サイトは html の meta タグで viewport を設定していないため、モバイルユーザを考慮していないデザインであるように見える。viewport はページのウィンドウサイズを制御する設定値である。逆に、Web サイト C ではデスクトップの滞在時間が 47 秒である一方で、モバイルデバイスでは 72 秒である。この Web サイトは html の meta タグで viewport を設定しているため、モバイルユーザを考慮したデザインであるように見える。これらのことからモバイルデザインに対応しているかどうかによってモバイルの滞在時間が大きく変化することが分かった。

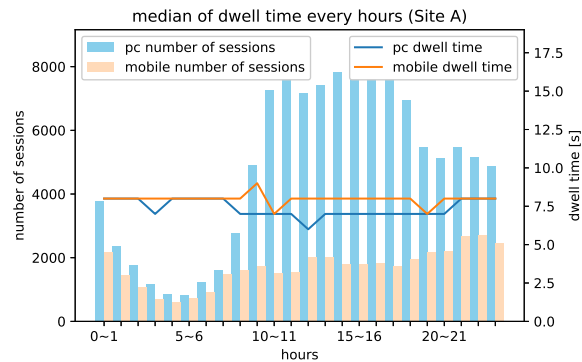
デバイス毎の滞在時間は大きく異なっているが、セッション数が急激に降下する秒数はほぼ同じである。どちらのデバイスを使用して Web ページを閲覧したとしてもデザインは異なる場合があるがコンテンツは同一である。セッション数が急激に降下する秒数がほぼ同じであることから、ユーザがその Web ページを見続けるかどうかを判断するためにかかる時間は同じ程度だと考えられる。

3.2 RQ2: 滞在時間はアクセスする時間帯、曜日、月によって差が出るか？

滞在時間がアクセスする時間によって変化するかを解析する。アクセスする時間によりセッション数が変化することは人間の生活リズムの関係から理解出来るが、滞在時間が変化するかどうかは知られていない。本実験では 1 年分のログデータを使用したため、アクセス時間として、時間帯、曜日、そして月に着目した。各項目では Web サイト毎にセッションを時間帯、曜日、月毎に分割し、さらに各セッション集合で滞在時間の中央値を計算する。それらの中央値を比較した場合に 20 秒以上の変化があった場合はその Web サイトにおいて滞在時間がその時間軸で「変化した」と定義する。例えば、ある Web サイトで 1 時の滞在時間と 8 時の滞在時間で 20 秒以上の差があった



(a) 時間帯別で滞在時間が変化する例 (Web サイト C)



(b) 時間帯別で滞在時間が変化しない例 (Web サイト A)

図 2 時間帯別の滞在時間 (変化する例と変化しない例)

場合、そのサイトは時間帯により滞在時間が変化したとみなす。

3.2.1 時間帯別の滞在時間

時間帯別で滞在時間が変化する例と変化しない例を図 2 に示す。これらの図では滞在時間に加え、セッション数が示されている。図 2a は時間帯で滞在時間が変化する例であり、図 2b は変化しない例である。

セッション数は時間帯で変化する傾向があり、例えば早朝ではセッション数が少ないが夜は多い。約半数の Web サイトでは滞在時間は時間帯で変化した。Blog に分類される Web サイトは滞在時間が大きく変化する傾向があり、Business では大きな変化は見られなかった。

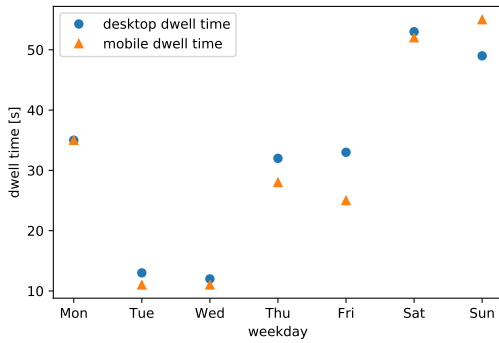
3.2.2 曜日別の滞在時間

曜日別で滞在時間が変化する例と変化しない例を図 3 に示す。これらの図では滞在時間が示されている。図 3a は曜日で滞在時間が変化する例であり、図 3b は変化しない例である。

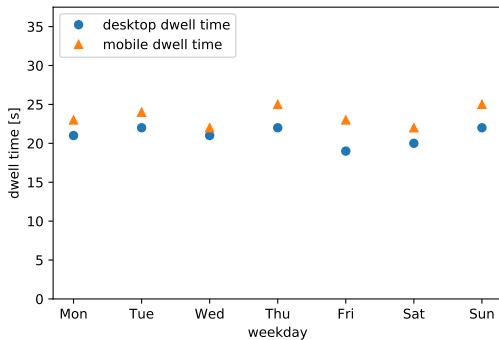
曜日で滞在時間が大きく変化したのは Web サイト I だけであった。この Web サイトは中古車の価格情報を提供している。ユーザが中古車販売店を訪問するなどの行動を準備しているときにユーザの閲覧行動がそれに影響を受け、したがって滞在時間が増加したと考えられる。ほとんどの Web サイトでは滞在時間は曜日によって変化しないという結果になった。

3.2.3 月別の滞在時間

月別で滞在時間が変化する例と変化しない例を図 4 に示す。これらの図では滞在時間に加え、セッション数が示されている。



(a) 曜日別で滞在時間が変化する例 (Web サイト I)



(b) 曜日別で滞在時間が変化しない例 (Web サイト B)

図 3 曜日別の滞在時間 (変化する例と変化しない例)

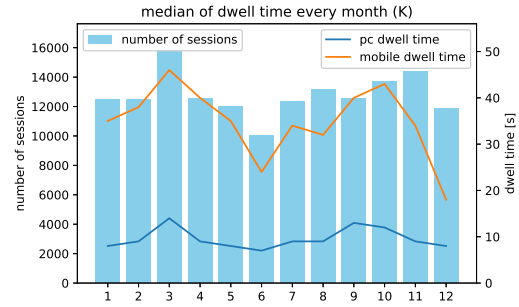
Web サイトによってはログデータが1年分に満たない場合があり、その場合の滞在時間とセッション数は0である。図 4a は月別で滞在時間が変化する例であり、図 4b は月別で滞在時間が変化しない例である。

月別の滞在時間が変化する Web サイトは多数あったが、Web サイト毎になぜ結果の違いが生じるかを明らかに出来なかった。その理由として、月が変わること以外の要因が入り込んでいることが挙げられる。Web サイトが更新されるか人気のコンテンツが追加されると滞在時間は変化すると思われる。この要因を除いて月別の解析を行なうために、今後の課題として Web サイトのコンテンツを考慮して解析する必要がある。

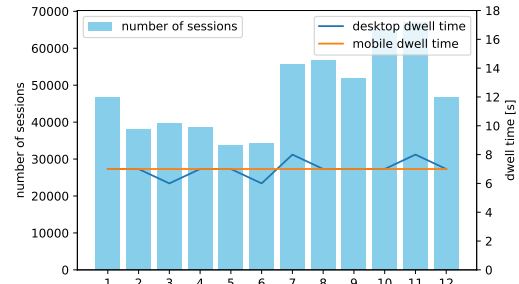
3.3 アクセス時間と滞在時間のまとめ

各 Web サイトのアクセス時間と滞在時間変化のまとめを表 2 に示す。この表では Y が該当する時間軸変化で 20 秒以上の変化があったことを表す。例えば、ある Web サイトで 1 時の滞在時間と 8 時の滞在時間で 20 秒以上の差があった場合、そのサイトの Hour カラムに Y が追加される。Web サイト D のデータはひと月分しかないため、月別の変化は存在しない。

BBS と Business タイプの Web サイトは滞在時間が時間帯と曜日による変化が無い傾向があるが、セッション数は時間帯と曜日によって変化している。滞在時間が変化しない理由はユーザが Web サイトを訪れる目的が変わらないからだと考えられる。



(a) 月別で滞在時間が変化する例 (Web サイト K)



(b) 月別で滞在時間が変化しない例 (Web サイト H)

図 4 月別の滞在時間 (変化する例と変化しない例)

表 2 アクセス時間と滞在時間変化のまとめ

#	Hour	Day	Month
A			
B	Y		Y
C	Y		
D			-
E			Y
F			
G	Y		Y
H			
I	Y	Y	Y
J	Y		
K	Y		Y
L	Y		Y
M	Y		

3.4 RQ3: 滞在時間は内部ユーザと外部ユーザで差が出るか?

ユーザが内部ユーザか外部ユーザかで滞在時間に差が出るかを解析する。本研究では、同一ドメインからページ遷移してきたユーザを内部ユーザ (Inside User) と定義し、そうではないユーザを外部ユーザ (Outside User) と定義した。各 Web サイトにおいてデスクトップとモバイル、さらに内部ユーザと外部ユーザにセッション集合を分割し、それぞれ滞在時間の中央値を計算する。

内部ユーザと外部ユーザの滞在時間を表 3 に示す。内部ユーザは外部ユーザに比べて滞在時間が長い傾向がある。同じ Web サイトで複数の Web ページを見るユーザは、他のユーザよりもその Web サイトにより興味を持っていると思われる。興味

の度合いが滞在時間の長さに影響し、内部ユーザは外部ユーザより滞在時間が長くなったと考えられる。BBS タイプの Web サイトのみが外部ユーザより内部ユーザの人数が多いが、このタイプのサイトではユーザは他のユーザ達と交流をするために同じ Web サイトに留まり続けるからと考えられる。

3.5 RQ4: 滞在時間とクリックに関係はあるか?

ユーザが Web ページ上でクリックをしたかどうかによって滞在時間が変化するかを解析する。クリックを一度でもしたユーザをクリックユーザ (Click User) と定義し、一度もクリックをしなかったユーザをノンクリックユーザ (Non-Click User) と定義した。クリックをするユーザの方がそのページにより興味を持っているため、滞在時間が長くなるのではないかと考えた。

クリックをしたユーザとしなかったユーザの滞在時間を表 4 に示す。クリックユーザはノンクリックユーザよりも滞在時間が長い傾向がある。ユーザが何か特殊な行動を取るとき、今回の場合はクリックをするという行動をするとき、ユーザはそのページに興味を持っていると考えられる。興味度合いが滞在時間の長さに影響し、クリックユーザの滞在時間がより長くなったと考えられる。クリックユーザの方がノンクリックユーザよりも数が少ない傾向があり、その傾向はデスクトップとモバイルデバイスの両方で同様である。一方で 3 つの Web サイト (Web サイト A, D, H) では、クリックユーザの方がノンクリックユーザよりも滞在時間が短い。Web サイト A と H はリンクが多い Web サイトである。日常的にその Web サイトに訪れるユーザはリンク先を知っており、短い時間でクリックによりページ遷移をする。Web サイト D はログデータの数が少なく、クリックユーザの数はたった 17 人である。正確な解析を行なうためにはより多くのクリックユーザが必要である。

クリックユーザに対して、各 Web サイトにおけるクリック前とクリック後の滞在時間を表 5 に示す。クリック後の滞在時間はクリック前よりも短い傾向がある。ユーザがページのリンクをクリックし別のページに移動したとき、クリック後の滞在時間は 0 秒になる。そのようなユーザがクリック後の滞在時間の中央値を減少させている。

クリックユーザの方が滞在時間が長いことと、クリック前の滞在時間の方が長いことにより、滞在時間が長いことでクリックが発生していることが考えられる。もしクリックにより滞在時間が長くなっていればクリック後の滞在時間の方が長くなるからである。この仮説をより強く裏付けるために、今後の課題として滞在時間が長いをクリックをしていないユーザがどの程度存在するかを調べる必要がある。

3.6 RQ5: 滞在時間とスクロール深度に関係はあるか?

Web ページ上のスクロールの深さで滞在時間が変化するかを解析する。本研究では Web ページの高さが 0%以上 25%以下の位置までスクロールしたユーザを Top ユーザ、25%より大きく 75%以下の位置までスクロールしたユーザを Middle ユーザ、76%より大きく 100%以下の位置までスクロールしたユーザを Bottom ユーザとして定義した。Top は Web ページ先頭

のヘッダー部分が含まれるように、Middle はページのコンテンツが含まれるように、Bottom はページのフッター部分が含まれるように定義をした。ユーザはセッション中で最も深いスクロール位置から Top, Middle, Bottom ユーザのどれかに分類される。

ページの下部に近づくほどスクロールする時間が長くなるため、全ての Web サイトにおいて Top ユーザの滞在時間が最も短く、次いで Middle ユーザがそれより長く、Bottom ユーザが最も長いと考えた。

Top, Middle, Bottom ユーザの滞在時間を表 6 に示す。Top ユーザの滞在時間が最も短い傾向がある。約半分の Web サイトで Bottom ユーザの滞在時間が最も長かった。したがって、幾らかのユーザはページの下部までスクロールをするが、滞在時間は短い。ユーザがページの下部まで閲覧したとしても、ページ全体を読んだとは限らないからである。

3.7 今後の課題

本研究では、滞在時間の予測に効果的な特徴量を発見するという目的を持って解析を行なった。まず RQ1 ではデバイス毎の滞在時間を解析し、デバイスによって滞在時間が大きく違うことを示した。このことから、デバイスによって予測モデルを変える必要があることが示唆される。また、RQ2 では時間軸変化が滞在時間に影響を与えるかを解析した。時間帯別と月別の滞在時間では滞在時間に変化があったことから、時間帯と月を予測モデルの特徴量として使用すると効果的な可能性がある。RQ3 では内部ユーザが外部ユーザよりも滞在時間が長いことを示した。理由として、ページ遷移をしたユーザの方がコンテンツにより興味を持っているからだと考えられる。本研究ではセッション毎の滞在時間のみ測定したが、ユーザ毎に滞在時間を予測する場合もページ遷移の回数が特徴量として効果的なのではないかと考えられる。RQ4 ではクリックが滞在時間に与える影響を解析した。クリックをするユーザの方が滞在時間が長いことから、滞在時間を予測する時の特徴量としてクリックの有無が効果的な可能性がある。RQ5 ではスクロールと滞在時間の関係を見出すことは出来なかった。Yi らは [3] コンテンツの長さが滞在時間と比例関係になることが示している。コンテンツの長さを考慮することでより正確な結果を得ることが出来ると考えられる。また、RQ4 では滞在時間が長くなることでクリックが発生することを示した。広告の分野などではユーザが広告を見る回数を増やすために、クリック数を多くすることを目的としている場合がある。滞在時間が長くなることでクリックが発生することから、クリック数を増やしたいという目的を滞在時間を長くするという目的として捉えることが出来る可能性がある。

本研究ではニュースサイトではない 13 サイトを扱ったが、データ提供元の性質上からマーケティングを行う Web サイトが多かった。今後の課題として、取り扱う Web サイトの数と種類を増やすことでより汎用的な解析を行ないたい。

4 おわりに

複数の Web サイトにおいてデスクトップとモバイルデバイスの両方について 1 年分のログデータを解析した。本研究の目的は、滞在時間の予測に効果的な特徴量を発見するために、ニュースサイトではない Web サイトにおいて滞在時間の特性を明らかにすることである。本研究ではデバイスの種類、アクセス時間、Web サイト上の行動、スクロール深度に着目し解析をした。

解析の結果、デスクトップとモバイルデバイスの両方で滞在時間が長くなるほどセッション数は減少することが分かった。さらに、Web サイトがどちらかのデバイスのみに対してデザインされているとき、滞在時間の中央値はデスクトップとモバイルデバイスでは大きく違うことが分かった。また、アクセスする時間帯と月は滞在時間に大きく影響するが、曜日はあまり影響しないことが分かった。さらに、内部ユーザとクリックユーザはそれぞれ外部ユーザとクリックしないユーザより滞在時間が長くなる傾向が見られた。この結果はユーザのコンテンツに対する興味の大きさが滞在時間の長さに影響したと考えられる。滞在時間とスクロールの関係を見出すことは出来なかった。これはユーザが Web ページの最下部まで閲覧したとしてもページ全体を読んだとは限らないからである。

上記解析において滞在時間に大きく影響を与えることが分かった要因は、滞在時間を予測するときの特徴量として利用出来る可能性がある。今回の結果を利用して、今後は滞在時間の予測を行ないたいと考えている。

文 献

- [1] Eugene Agichtein, Eric Brill, and Susan Dumais. Improving web search ranking by incorporating user behavior information. In *Proceedings of the 29th annual international ACM SIGIR conference on Research and development in information retrieval*, pp. 19–26, 2006.
- [2] Masahiro Morita and Yoichi Shinoda. Information Filtering Based on User Behavior Analysis and Best Match Text Retrieval. In *Proceedings of the 17th annual international ACM SIGIR conference on Research and development in information retrieval*, pp. 272–281, 1994.
- [3] Xing Yi, Liangjie Hong, Erheng Zhong, Nanthan Nan Liu, and Suju Rajan. Beyond clicks: dwell time for personalization. In *Proceedings of the 8th ACM Conference on Recommender systems*, pp. 113–120, 2014.
- [4] Mounia Lalmas, Janette Lehmann, Guy Shaked, Fabrizio Silvestri, and Gabriele Tolomei. Promoting Positive Post-Click Experience for In-Stream Yahoo Gemini Users. In *Proceedings of the 21th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pp. 1929–1938, 2015.
- [5] Ke Zhou, Miriam Redi, Andrew Haines, and Mounia Lalmas. Predicting Pre-click Quality for Native Advertisements. In *Proceedings of the 25th International Conference on World Wide Web*, pp. 299–310, 2016.
- [6] D Sculley, Robert Malkin, Sugato Basu, and Roberto J Bayardo. *Predicting Bounce Rates in Sponsored Search Advertisements*. 2009.
- [7] Shima Kazuyuki, Makoto Sakai, Ken-ichi Matsumoto, and

- Nakamichi Noboru. Detecting Low Usability Web Pages using Quantitative Data of Users' Behavior. In *proceedings of the 28th International Conference on Software Engineering*, pp. 569–576, 2006.
- [8] Chao Liu, Ryen W. White, and Susan Dumais. Understanding web browsing behaviors through Weibull analysis of dwell time. In *Proceeding of the 33rd international ACM SIGIR conference on Research and development in information retrieval*, pp. 379–386, 2010.
- [9] Theodore Vasiloudis, Hossein Vahabi, Ross Kravitz, and Valery Rashkov. Predicting Session Length in Media Streaming. In *Proceedings of the 40th International ACM SIGIR Conference on Research and Development in Information Retrieval*, pp. 977–980, 2017.
- [10] Dmitry Lagun and Mounia Lalmas. Understanding User Attention and Engagement in Online News Reading. In *Proceedings of the Ninth ACM International Conference on Web Search and Data Mining*, pp. 113–122, 2016.
- [11] Yang Song, Hao Ma, Hongning Wang, and Kuansan Wang. Exploring and Exploiting User Search Behavior on Mobile and Tablet Devices to Improve Search Relevance. In *Proceedings of the 22nd International Conference on World Wide Web*, pp. 1201–1212, 2013.

表 3 内部ユーザと外部ユーザの滞在時間

#	Desktop				Mobile			
	Inside User		Outside User		Inside User		Outside User	
	Sessions	Time (s)	Sessions	Time (s)	Sessions	Time (s)	Sessions	Time (s)
A	14,283	7	103,046	7	8,159	7	32,788	8
B	694	121.5	9,597	20	400	86	18,514	23
C	11,935	65	180,944	46	7,656	101	432,039	71
D	8	86.5	716	12	1	14	1,013	12
E	113,371	40	407,421	20	11,668	13	100,458	14
F	2,148	40	24,736	7	3,062	39	39,363	9
G	30,314	79	141,251	50	123,728	56	374,711	90
H	164,223	7	121,916	7	202,771	7	152,395	6
I	1,757	18	101,115	28	1,376	17	149,023	20
J	8,887	30	91,271	38	1,210	24	41,264	8
K	6,496	47	44,844	8	15,374	51	103,217	32
L	20,895	82	173,943	9	28,830	73	486,033	21
M	193	88	17,880	9	412	69.5	31,103	10.0

表 4 クリックをしたユーザとしなかったユーザの滞在時間

#	Desktop				Mobile			
	Click User		Non-Click User		Click User		Non-Click User	
	Sessions	Time (s)	Sessions	Time (s)	Sessions	Time (s)	Sessions	Time (s)
A	68,292	6	49,037	10	22,164	8	18,783	8
B	999	136	9,292	19	1,905	83	17,009	21
C	27,922	123	164,957	34	63,439	122	376,256	62
D	17	5	707	12	0	-	1,014	12
E	132,655	43	388,137	18	11,974	74	100,152	14
F	2,021	56	24,863	7	2,755	56	39,670	9
G	34,747	94	136,818	44	91,994	135	406,445	71
H	216,383	7	69,756	10	267,790	6	87,376	10
I	2,965	62	99,907	17	2,478	63	147,921	14
J	12,661	55	87,497	34	1,849	47	40,625	8
K	5,639	70	45,701	8	17,930	95	100,661	20
L	8,304	46	186,534	10	35,839	131	479,024	19
M	199	93	17,874	9	446	190	31,069	10

表 5 クリック前とクリック後の滞在時間

#	Desktop		Mobile	
	Until (s)	Since (s)	Until (s)	Since (s)
A	5	0	7	0
B	49	17	21	35
C	23	38	20	58
D	4	1	-	-
E	20	1	32	9
F	46	1	48	1
G	73	0	95	0
H	6	0	6	0
I	24	2	18.5	12
J	40	1	38	12
K	52	1	75	1
L	21	1	55	6
M	52	1	128	10.5

表 6 滞在時間とスクロール深度の関係

#	Desktop			Mobile		
	Top	Middle	Bottom	Top	Middle	Bottom
A	6	25	44	7	30	47
B	24	26	20	26	76	15
C	12	42	89	16	116	163
D	6.5	18	12	304.5	19	11
E	11	48	86	14	15	14
F	7	8	7	8	9	9
G	13	56	87	11	78	131
H	5	11	22	6	16	21
I	67	24	9	67	9	20
J	27	43	35	8	9	7
K	7	7	10	9	40	54
L	11	9	10	18	61	16
M	8	10	9	9	23	9