

アドネットワークにおける広告効果指標の調査

柴山 りな¹ 草野 蘭之介² 菊池 浩明²

概要: インターネット広告は、効率的に広告を配信できるサービスとして大きく成長している。中でもディスプレイ広告はウェブサイトなどの掲載枠に画像や画像と文字の組み合わせで宣伝をする広告であり、認知拡大やコンバージョンの取得を目的として配信されている。インプレッション数やクリック数に応じて課金がされる仕組みである。ただ、広告枠に対する広告の掲載とそのクリック数などの効果指標の計測と報告はアドネットワーク業者が行っていて、その正確性を知ることは困難である。実際にあるまとめサイトが広告料が不正に横取りしていた事件も発生している。そこで本研究では、代表的なアドネットワークに広告を出稿して、流入したユーザの数、行動などを取得することにより、正しく広告が配布されているかを広告主の観点から調査することを目的とする。

Measurement of Advertising Effectiveness in Ad Networks

RINA SHIBAYAMA¹ RANNOSUKE KUSANO² HIROAKI KIKUCHI²

1. はじめに

電通によると [1], 日本におけるインターネット広告費は 6 年連続 2 桁成長で、テレビメディア広告費を超え、2019 年には 2 兆円を超えている。しかし、2018 年、曾於市ふるさと納税の広告を掲載しているあるまとめサイトで広告料を不正に横取りしていた事件を NHK が報道した [3]。これは、広告枠に対する広告の掲載とその効果指標の報告はアドネットワーク業者が行っており、広告主(曾於市)や広告配信業者が、広告がいつどこで掲載されているか知らないことが原因のひとつである。アドネットワークが出稿した広告のインプレッション数(表示回数)やクリック数の水増しを行うという不正が問題になってきている。

このオンライン広告の問題に対して、広告を掲示するサーバ側から検出する研究が行われている。例えば、金井らは、クリック座標に着目した提案を行っている [8]。Iqbal らも、不正クリックを検出する技術を提案している [4]。これらの研究の多くは、クリックポットなど外部の要素による広告不正を検出することを目的としている。しかしながら、これらの試みはいずれも広告を提供するアドネットワーク事業者の立場における対策であり、本研究で対象と

する広告主の観点での不正を確かめることはできない。

そこで、本研究では、広告主の観点から正しく広告が配布されているかを調査することを目的とする。具体的には、アドネットワーク業者が報告するクリック数とアクセスの数が一致するか、また広告クリックがユーザにとって誤操作によるものでないかを推定して明らかにする。

その為、代表的な複数のアドネットワークに実際に広告を出稿して、流入したユーザのアクセスを広告主のサーバで観測し、アドネットワークから報告されるインプレッション数やクリック数とサーバのアクセスログが整合しているかを検証する。加えて、ユーザのページ滞在時間やスクロール情報などを PHP, JavaScript で取得し、行動の推定を行う。

2. アドネットワークとその課題

2.1 インターネット広告の仕組み

インターネット広告には、広告を出稿する広告主、広告を閲覧するエンドユーザ、広告表示枠を提供するパブリッシャ、広告主とパブリッシャの間を取り持つアドネットワークが関わっている。

インターネット広告の仕組みを図 1 に示す。まず、(1) エンドユーザがパブリッシャ(Web サイト)を訪れた際、(2) その Web サイト上に仕込まれたアドタグ(HTML 内

¹ 明治大学 先端数理科学研究科 先端メディアサイエンス専攻

² 明治大学 総合数理学部 先端メディアサイエンス学科

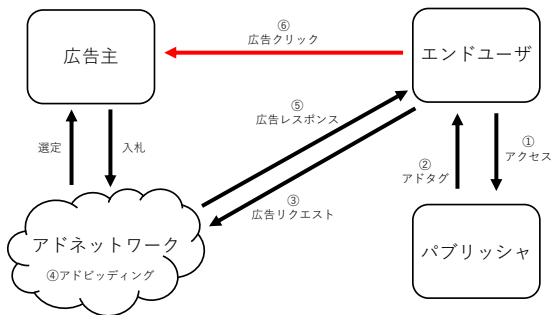


図 1 インターネット広告の仕組み

の<script>タグなど)がブラウザ上に読み込まれる。(3)その後、アドタグについて実行される JavaScript からアドネットワークに対して広告のリクエストが送られる。(4)この際、アドネットワーク上では、アドビiddingと呼ばれる広告枠の競売が行われる。この競売では、入札者の入札額アドタグについて送信されたパブリッシャの情報(広告枠のサイズなど)や、エンドユーザの情報(Cookie, OSバージョンなど)が基準となり、落札され、広告が決定する。(5)最終的に、その広告枠を買った広告主のディスプレイ広告が、閲覧しているページに表示される。(6)エンドユーザは、ディスプレイ広告をクリックすることで広告主の Web サイトに飛ぶ。同時に、インプレッション数や広告クリック回数が評価される。

広告効果指標を次のように呼ぶ。

インプレッション ユーザが Web サイトを訪問し広告枠に広告が掲載され、ユーザが広告を見ること。

リーチ 広告を 1 回以上見た人の数。同一人物が複数回閲覧しても 1 回とカウントする。

クリック 表示された広告がクリックされること。

コンバージョン クリックされた先の Web ページで商品購入など広告主の目標が達成されること。

CTR(Click Through Rate, クリック率) 表示された広告に対して、クリックされる割合

Adstage 社によると、2020 年の CTR は、Facebook 広告で 1.11%、Instagram 広告で 0.22%、Twitter 広告で 0.86%、Google ディスプレイ広告で 0.46% などである [2]。

また、ここでは代表的な広告の種類を次のように定義する。

リスティング広告 検索エンジンやオンラインショッピングサイトなどにおいて、検索したときに、検索結果に表示される広告。

ディスプレイ広告 Web ページ内やアプリ内の決められた広告枠に表示される広告。バナー広告、レスポンス広告、動画広告、テキスト広告などの種類がある。

バナー広告 ディスプレイ広告のひとつで、バナーと呼ばれる画像を表示する広告。

レスポンス広告 ディスプレイ広告のひとつで、画像と文章を組み合わせた広告。広告枠に合わせて、サイズやレイアウトが変化する。

動画広告 ディスプレイ広告のひとつで、動画による広告。

2.2 アドネットワークに関する不正

広告主と広告媒体 (Web サイトやアプリなど)をつなぐアドネットワークというプラットフォームにより、簡単に大量の広告媒体に出稿できるようになった。一方、数ある広告枠の中から広告の出稿先を選定できないために、知らないうちに不正なサイトに出してしまう可能性がある。不正不正とは、実際には発生していない広告効果を不正な手段で発生したように見せかける攻撃である。主な攻撃者として以下の 3 つが考えられる。

2.2.1 ユーザやボットによる不正

ユーザが広告主に不利益を与える目的で広告をクリックするといった問題ある。例えばリスティング広告のクリック単価は高い場合 1000 円を超えるため、たとえ手動によるクリックでも被害は大きくなりうる。ボットによる不正インプレッション、不正クリックはより深刻である。ボットに感染したユーザの端末は、画面上に大量の広告を表示させたり、リックしたりして不正に広告収入を搾取する。

2.2.2 パブリッシャの不正

アドネットワークは、広告の掲載枠を審査をしたうえで、広告を配信する枠 (Web ページやアプリなど)を選定する。しかし、その審査基準はさまざまである。例えば、ユーザを騙して広告クリックを誘導するパブリッシャも存在する。クリック数が増加するとアドネットワークの利益にも繋がるため、このようなパブリッシャを黙認するアドネットワークも存在する。また、第三者のパブリッシャサイトから強制的にユーザーを移動する飛ばし裏広告という不正も存在する。

2.2.3 効果指標の評価の不正

アドネットワークでは、時間、曜日、地域指定など細かな配信条件を広告主が設定でき、インプレッション、クリック、コンバージョンなど効果指標もアドネットワークが計測してくれる。ただし、報告される効果指標が正確である保証はない。例えば、広告が準備されただけで実際にはユーザの目に入っていないでもインプレッションとしてカウントする不正などが考えられる。

2.3 先行研究

このうち、ボットによる不正クリックを検出する研究が多くなされている。HTTP リクエスト情報から不正なアクセスを判別する方法 [5] や、CTR に着目した検出方法 [6] が提案されている。また、JavaScript の実行でクリックボットを検出し、広告主サイトでのユーザの行動を監視することでより高度なクリックボットや人間によるクリックを



図 2 実験サイト^a

表 1 取得した情報の一覧

	データの種類	データの例
アクセ スログ	IP アドレス	111.239.175.143
	アクセス時刻	2020/11/29 00:12:03
	ユーザー エージェント	Mozilla/5.0(Linux;Android7.1.2;SH-M05)AppleWebKit/537.36(KHTML,likeGecko)Chrome/81.0.4044.96MobileSafari/537.36
	リクエストメ ソッド	GET
	リファラー	https://www.pixiv.net/
行動	滞在時間	00:00:01
	スクロール率	0.3

検出する、2つの組み合わせで不正広告を検出する方法も提案されている [7]。Iqbal らは、ユーザのインターネットアクティビティを用い、オペレーションレベルでの不正クリック検出技術を提案している [4]。

金井らは、広告によって生じる通信や広告が表示されるクライアントアプリケーションからではなく、広告ネットワーク上で観測された広告閲覧やクリックのログを用いて広告の実態調査を行った [8]。結果、クラウド事業者が管理する IP アドレスから高い割合で不正な広告アクセスが発生している事例が報告された。

しかしながら、これらの試みはユーザやクリックボットなどのアドネットワーク外部の要素への対策であり、本研究で対象とする広告主の観点からのアドネットワークやパブリッシャによる不正を確かめることはできない。

3. 実験

3.1 目的

インターネット広告による集客の効果とアドネットワークによって報告される効果指標の正しさを調査することを目的とする。

3.2 方法

本研究では、代表的なアドネットワークとして、Google

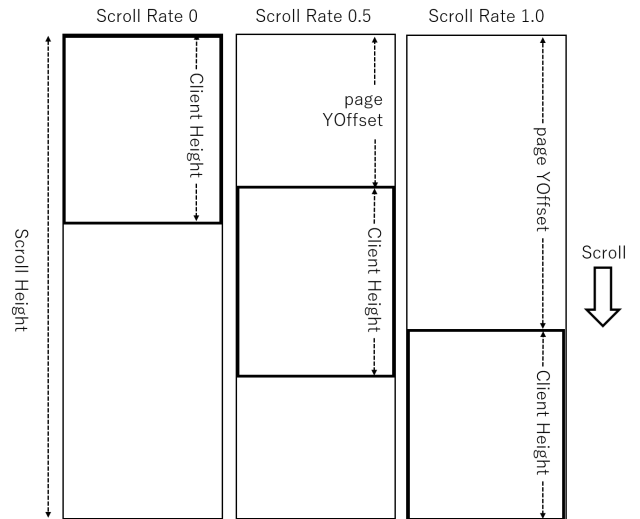


図 3 スクロール率の算出例

広告^bと Facebook 広告^cに出稿する。Google 広告は、広告枠のあるサイトが日本だけで 1 万以上ある大規模なアドネットワークである。一方 Facebook 広告の配信先は主に 4 つあり、Facebook, Instagram, Messenger, Audicence Network(Facebook 社と提携しているスマホアプリやスマホサイト)である。Facebook 広告の広告媒体は、自社で運営している媒体が多いことが特徴である。Google 広告と Facebook 広告に出稿したのは、こういった特徴の違うアドネットワークを比較するためである。

表 2 のように、レスポンスディスプレイ広告を出稿し、アクセス数や行動などを観測する。ただしレスポンス広告であるため、これらの図は一例であり、この通りにユーザに表示されるとは限らない。広告の配信設定は、Google 広告と Facebook 広告でできるだけ同じになるように設定した。配信地域は日本、オーディエンスにはエンターテインメントに興味があるユーザを選んだ。また、サーチキーワード(例えば「映画 おすすめ」というワードで検索した人に広告を表示する)やリターゲティング(一度サイトを訪れたことがある人へ配信する)などは選択しなかった。

広告をクリックしたユーザは、図 2 の実験用コンテンツページへ遷移する。Google 広告と Facebook 広告で同じページを使用した。興味のあるユーザは数十秒ページに滞在して下までスクロールし、興味がない場合は、数秒でページから離脱すると考えられる。遷移先コンテンツページのアクセスログ、ページ内での行動を PHP と JavaScript で取得する。取得した情報の詳細を表 1 に示す。

アクセスログとして、IP アドレスとアクセス時刻、リクエストメソッド、ユーザーエージェント、リファラー(遷移元 URL)を取得し、アクセス数を推定する。またアクセスに使用したデバイスやアクセス時刻を比較する。ただしリ

^a <https://windy.mind.meiji.ac.jp/~Rina/ad/m.php>

^b Google 広告, https://ads.google.com/intl/ja_jp/home/

^c Facebook 広告, <https://www.facebook.com/business/ads>

表 2 各アドネットワークの広告出稿情報

アドネットワーク	Google 広告	Facebook 広告
配信期間	2020/11/29~12/5 の 7 日間	2021/04/29~05/02 の 4 日間
配信地域	日本	日本
興味と関心	メディア, エンターテインメント 映画ファン	ブラック, エンターテインメント, テレビジョン
配信形式	レスポンス広告	レスポンス広告
広告配信例		

ファラーは、広告掲載サイトの設定によっては、取得することができない。

3.3 行動推定

アクセスしたユーザはコンテンツに興味を持っているのか、クリックがユーザにとって誤操作ではないのかを確認するため、各アクセスの滞在時間とスクロール率を JavaScript で取得した。ページ滞在時間とスクロール率は、0 秒、1 秒、3 秒、5 秒、その後 5 秒ごとに取得した。

スクロール率は、コンテンツ全体の高さにおける、ユーザがスクロールした領域の割合と定める。JavaScript でスクロール量 (`window.pageYOffset`)、コンテンツページ全体の高さ (`document.documentElement.scrollHeight`)、ユーザの画面の高さ (`document.documentElement.clientHeight`) の 3 つの値を取得した (単位: ピクセル)。スクロール率は以下の式で計算される。

$$\text{ScrollRate} = \frac{\text{pageYOffset}}{\text{ScrollHeight} - \text{ClientHeight}}$$

スクロール量については、ページ離脱時の値ではなく最大値を採用した。概要を図 3 に示す。例えば、本実験では `ScrollHeight` が 3700 であり、`ClientHeight` はデバイスによって異なるが 500 などである。このとき、300 ピクセルスクロールしたとすると、最大スクロール率は

$$\text{ScrollRate} = \frac{400}{3700 - 500} = 0.125$$

となる。

3.4 推定結果

Google 広告と Facebook 広告の出稿結果、計測したアクセスログ数とその分析結果を表 3 に示す。Google 広告では 72,756 インプレッションと 303 クリックが、Facebook 広告では 4440 インプレッションと 364 クリックが報告された。クリック単価はどちらも指定しなかったところ、Google 広告が 26 円、Facebook 広告が 8 円であった。

表 3 Google, Facebook 広告の掲載結果

	Google	Facebook
インプレッション数 (広告表示回数)	72,756	4,440
リーチ数	-	3,385
広告クリック数	303	364
クリック率 [%]	0.42	8.31
アクセスログ数	414	369
Google からのアクセス数	12	0
Facebook からのアクセス数	0	0
ヘッダーリクエスト	6	0
重複アクセス数	96	8
推定アクセス数	300	361
一致度 [%]	99.0	99.2

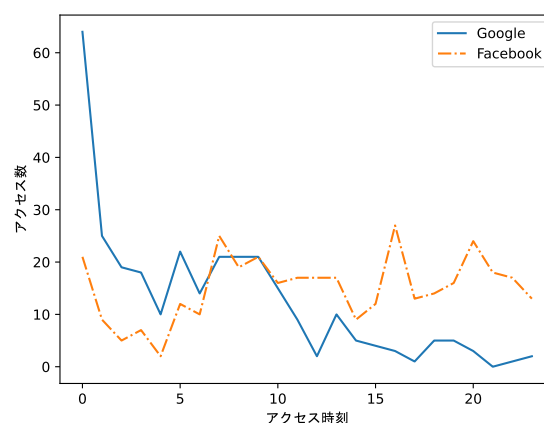


図 4 アクセス時間帯の分布

リファラーより、Google 広告が掲載された場所は YouTube(22 件)からのアクセスが多かった。ただ、300 中 241 アクセスでリファラーは記録されなかった。他に取得できたリファラーには <https://www.pixiv.net/>、<http://www.smartnews.com/> などがあつた。Facebook 広告では配信先は主に 4 つあるが、すべて Facebook に出稿された。記録されたリファラーは <http://m.facebook.com/>

や android-app://m.facebook.com/ などであった。

アクセス数の時刻分布を図4に示す。Googleは0時台が最も多く、12時から23時の活動時間帯が少なかった。一方、Facebook広告は、アクセスが集中した時間は見られなかった。Facebook広告を出稿したのは祝日3日を含む4日間であったのに対し、Google広告の7日間は土日を含むものの祝日はなかった。

アクセスに使用されたデバイスをユーザエージェントから判別した結果を図5に示す。どちらのアドネットワークでもiPhoneとAndroidが9割以上を占めた。

流入したユーザの滞在時間とスクロール率の分布をJavaScriptで取得した結果をそれぞれ図6、7に示す。滞在時間は、Google広告では5秒以内の比較的短い層と2分以上の長い層に別れたが、Facebook広告では20から90秒が一番多かった。Google広告では195アクセスのうち154で全くスクロールしないまま、ページを離れていることが確認できた。一方、Facebook広告では、全く対照的で354のうち180のアクセスでページ最下部までスクロールされた。ただし、Google広告では300のうち105のアクセスで、Facebook広告では、361のうち7のアクセスでデータが取得できなかった。

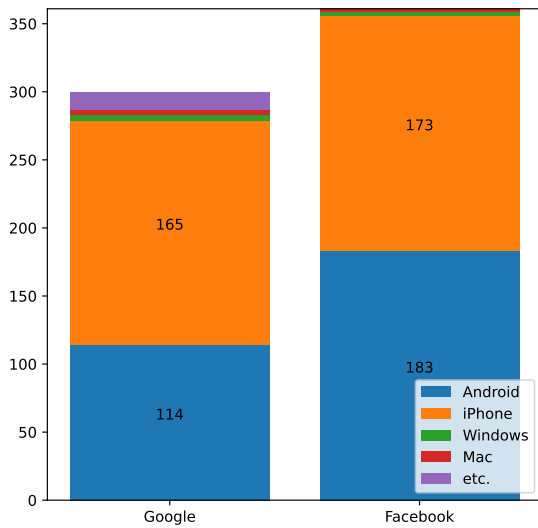


図5 デバイスごとのアクセス数

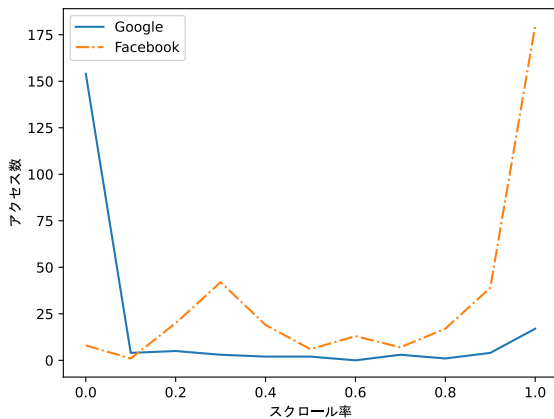


図6 スクロール率のアクセス分布

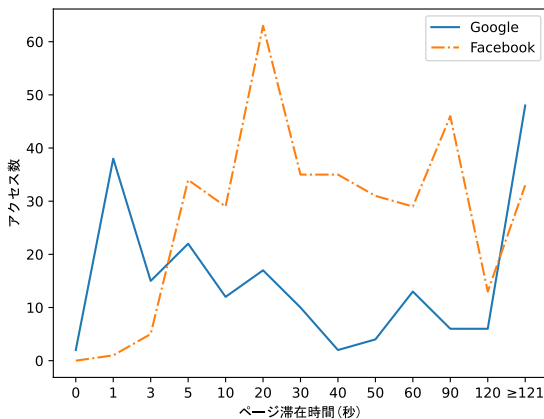


図7 ページ滞在時間のアクセス分布

4. 分析と考察

4.1 広告クリック数とアクセス数の比較

Google広告とFacebook広告のアドネットワークにより計測されたクリック数と我々のサイトでの計測数を比較する。表3より、Google広告でのクリック数が303だったのに対し、観測されたアクセスログ数は414であった。414からヘッダーリクエスト6件、Botからのアクセス12件、同一IPアドレスからの5分以内のアクセス96件を除外すると300となる。この結果は、Google広告のクリック数303と99.0%一致する。Facebook広告も同様に369のアクセスログから重複アクセス8件を除くと、361となり、Facebook広告のリンククリック数と99.2%一致する。

本研究では同一アクセスと判断する条件を5分以内としたが、Google広告で1分以内とすると311、30分以内とすると295、120分以内とすると292となり、いずれにせよ96%以上一致する。Facebook広告では5分以内と120分以内で重複アクセスは同数であった。

また、CTRは、Google広告で0.42%、Facebook広告で8.31%と、結果が大きく異なった。

4.2 時刻、デバイス

図4より、Google広告でのアクセスの時刻は0時台が最も多く、0時から11時にかけてが比較的多い時間帯であった。しかし、総務省通信白書[9]によると日本のインターネットの利用時間帯は朝6時台から23時台までが10%を超えて多く利用される時間帯である。特に12時と20か

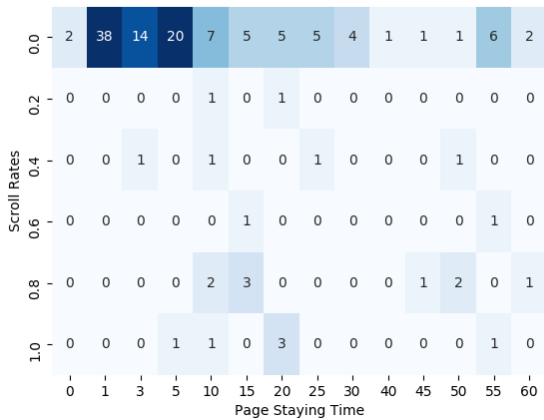


図 8 Google 広告のページ滞在時間 [秒](縦) とスクロール率 (横) ごとのアクセス分布

ら 22 時は 20% を超える最多時間帯である。したがって両者は整合しない。よって、広告は必ずしも時間的に均等にユーザに表示されていないと考えられる。

Facebook 広告では、日本人がインターネットを利用するとされる 6 時から 20 時の時間帯に多くアクセスがあった。また図 5 より、アクセスに使用されたデバイスは、スマートフォンが 9 割以上を占めた。

4.3 流入したユーザの振る舞い

取得できたアクセスのページ滞在時間とスクロール率ごとのアクセス分布をそれぞれ図 8, 9 に示す。図の左上 (ページ滞在時間 10 秒以内) は、広告をクリックしたものの求めていたコンテンツと違ったため、すぐに離脱したと考えられる。また、ユーザの意図しない誤操作による広告クリックの可能性もある。逆に、右下は興味を持って最後までコンテンツを読んでいると考えられる。このどちらかや真ん中のようにユーザは振るまうと予想していた。Facebook 広告では予想通りになったが、Google 広告ではスクロールせずに 50 秒以上長く滞在するアクセスも見受けられた。これは興味のない広告を開いたままにしていることなどが考えられる。

ユーザの行動を JavaScript で取得を試みたが、Google 広告では約 35% の頻度で取得に失敗した。原因としては、ユーザが JavaScript をブロックする設定にしていることが考えられる。Facebook では、取得できなかったアクセスは 2% にとどまった。

5. おわりに

本論文では、Google ディスプレイ広告と Facebook 広告における広告効果指標の分析を行った。その結果、クリック数とアクセス数は Google 広告、Facebook 広告ともに 99% 以上一致した ($n = 303, n = 364$)。一方、Google 広告ではアクセスしたユーザの 21% が 1 秒以内にページを閉

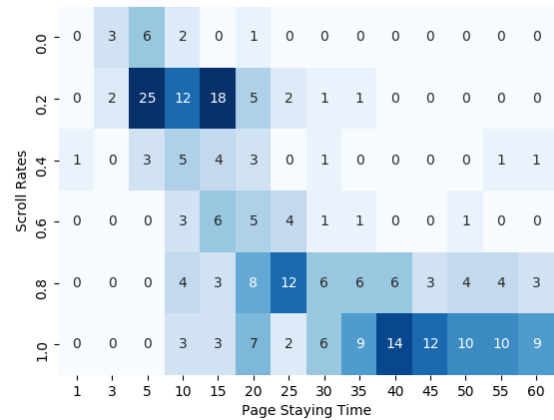


図 9 Facebook 広告のページ滞在時間 [秒](縦) とスクロール率 (横) ごとのアクセス分布

じ、79.0% がスクロールを全くせずにページを閉じていたことが明らかになった。広告がクリックされた時間帯は、0 時から 12 時が多く日本のインターネットの利用時間とずれていた。Facebook 広告では、ページの 20% 以上をスクロールし、5 秒以上滞在するアクセスが多かった。また広告クリックの時間帯も日本の利用時間と一致していた。

ただし、広告効果指標からクリックやインプレッションの水増しなどの広告不正について調査するには限界がある。今後は、離脱の理由についてユーザにアンケートを実施するなど、広告効果指標以外の側面から調査を行う必要がある。

参考文献

- [1] 2019 年日本の広告費, (https://www.dentsu.co.jp/knowledge/ad_cost/2019/, 2020/12/16 参照。)
- [2] ADSTAGE, (<https://www.adstage.io/>, 2020 年 12 月参照。)
- [3] NHK, 「もうけは誰の手に? 闇に消えるネット広告費」, (https://www3.nhk.or.jp/news/special/net-koukoku/article/article_05.html, 2020 年 6 月参照。)
- [4] Md Shahrear Iqbal, Mohammad Zulkernine, Fehmi Jaafar, Yuan Gu, "Protecting Internet users from becoming victimized attackers of click-fraud", WILEY, Journal of Software Evolution and Process, 2018.
- [5] Metwally A, Agrawal D, Abbadi AE. Using association rules for fraud detection in web advertising networks. Proceedings of the 31st International Conference on Very Large Databases, VLDB Endowment, Trondheim, Norway, 2005.
- [6] Immorlica N, Jain K, Mahdian M, Talwar K. Click fraud resistant methods for learning click-through rates. Proceedings of the Internet and Network Economics: Springer, Hong Kong, China, 2005.
- [7] Xu H, Liu D, Koehl A, Wang H, Stavrou A. Click fraud detection on the advertiser side. Proceedings of the 19th European Symposium on Research in Computer Security: Springer, Wroclaw, Poland, 2014.
- [8] 金井 文宏 ほか, 広告ネットワーク上で観測されたユーザアクティビティの分析による広告不正の実態調査, 情報

処理学会 研究報告セキュリティ心理学とトラスト (SPT),
pp. 1-6, No. 17, Vol. 2018-SPT-27, 2018.

- [9] 総務省, 「平成 29 年版 情報通信白書 主なメディアの利用時間帯」, (<https://www.soumu.go.jp/johotsusintokei/whitepaper/ja/h29/html/nc262520.html>, 2020 年 12 月参照.)