

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†

草野蘭之介†

明治大学 総合数理学部 先端メディアサイエンス学科 菊池研究室†

1 はじめに

電通によると [1], 日本の広告は 8 年連続でプラス成長をしている. 特に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, NHK が, あるまとめサイトに曾於市ふるさと納税の広告が掲載されており, そのサイトで広告料を不正に横取りしていた事件を報道した [4]. これは, 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. インターネット広告では, 広告枠に対する広告の掲載とその効果指標の報告を, 全てアドネットワーク業者が行っており, これは従来のメディアとの大きな違いである. アドネットワークが 出稿した広告の表示回数を表すインプレッション数とクリックされる割合の水増しを行うという不正が問題になってきている.

そこで本研究では, アドネットワークの掲載する広告とその効果指標の正しさを広告主の観点から調査することを目的とする. 代表的なアドネットワークに広告を出稿して, クリックして流入したユーザの数, 行動などを PHP で取得することにより, 正しく広告が配布されているかを調査する.

2 アドネットワーク

2.1 インターネット広告

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存在する.

インターネット広告の仕組みを図 1 に示す. まず, エンドユーザがパブリッシャ (Web サイト) を訪れた際 (1), その Web サイト上に仕込まれたアドタグ (HTML 内の<script>タグなど) がブラウザ上に読み込まれる (2). その後, アドタグについて実行される JavaScript からア

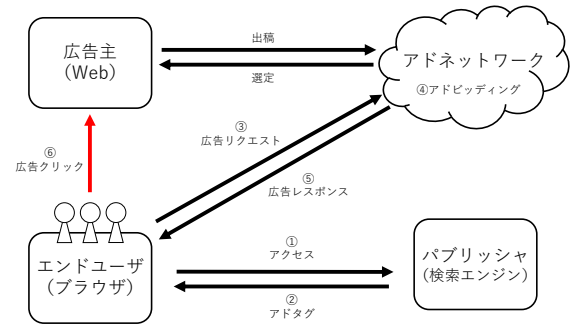


図1 インターネット広告の仕組み

ドネットワークに対して広告のリクエストが送られる (3). この際, アドネットワーク上では, アドビディングと呼ばれる広告枠の競売が行われる (4). この競売は, 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, 落札される. 最終的に, その広告枠を買った広告主の広告が, 閲覧しているページに表示される (5). エンドユーザは, 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, 閲覧回数 (インプレッション) や広告クリック回数が評価される.

2.2 先行研究

金井らは, 広告不正におけるクリック座標に着目した調査を行った [2]. 不正クリックには X 座標が共通で, Y 座標のみが異なるクリックが多く見られたことを報告している. Iqbal らは, クリック詐欺に対する新たな不正検出技術を提案している [3]. 従来は主にサーバ側からの調査であり, VPN や匿名ネットワークの場合, 不正の検出が難しいため, ユーザ側からの調査が出来る新たな技術を提案した.

3 実験

3.1 目的

インターネット広告による集客の実態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.

†Department of Frontier Media Science, School of Interdisciplinary Mathematical Science, Meiji University, Kikuchi Laboratory.

表1 取得した情報の一覧

取得元	データの例
アクセスログ	IP アドレス
	アクセス時刻
	リクエストメソッド
	ユーザエージェント
	リファラー
行動	滞在時間
	スクロール率

3.2 方法

次のように、広告を出稿し、アクセス数・行動などを観測する。2020/11/28 から 12/5 の7日間、Google 広告* で図3のようなディスプレイ広告を出稿した。広告をクリックしたユーザは、図4の実験用コンテンツページへ遷移する。広告クリックからのアクセスログを PHP で、遷移先コンテンツページでの行動を JavaScript で取得し、データベースに格納する。取得した情報の一覧を表1に示す。

取得したアクセスログの IP アドレスから地域を特定し、地域ごとのアクセス数を取得する。アクセス地域の判別には、MAXMIND 社の GeoLite2-City†のデータベースを利用する。

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 で取得する。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秒ごとに測定する。

スクロール率は、コンテンツ全体の高さのうち、ユーザがスクロールした領域を割合で定める。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。

$$\text{スクロール率} = \frac{\text{最大スクロール量}}{\text{ページ全体の高さ} - \text{画面の高さ}}$$

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。これらの概要を図2に示す。

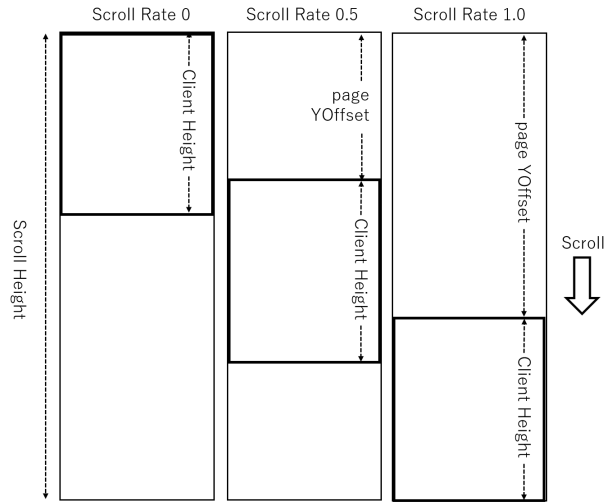


図2 スクロール率の取得指標



図3 実験に使用したディスプレイ広告



図4 実験サイト

3.3 結果

Google 広告に7日間出稿した結果を表2に示す。また、計測したアクセスログの件数とその分析結果を表3に示す。

地域ごとのアクセス数を GeoLite2 と Google 広告とで比較した結果を図5に示す。GeoLite2 では、304 アクセ

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

†GeoLite2 City, <https://dev.maxmind.com/geoip/geoip2/geoip2/>

表2 Google 広告の掲載結果

インプレッション (広告表示回数)	72,756
広告クリック数	303
クリック率 [%]	0.42

表3 Google 広告からのアクセス数と分析

アクセスログ数	414
ボットからのアクセス数	3
推定アクセス数	304

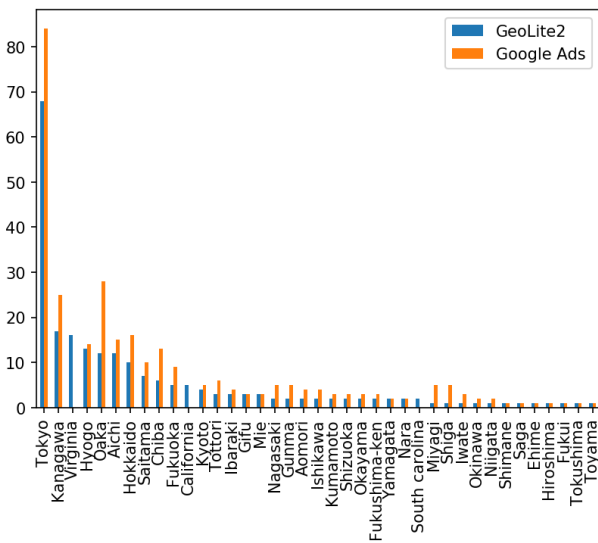


図5 アクセス地域の分布

ス中 79 件でデータなしという結果になった。GeoLite2 と Google 広告を比較すると、まず地域が特定されているアクセス数が GeoLite2 では 225 であるのに対し、Google 広告は 301 と多い。また、GeoLite2 は海外からのアクセスが合計 26 アクセスという結果であった。

時間ごとのアクセス数を図 6 に示す。0 時台が最も多く、12 時から 23 時にかけてが少なかった。

アクセスに使用されたデバイスを http のユーザーエージェントタグから判別した結果を図 7 に示す。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を JavaScript で取得した結果をそれぞれ表 5、4 に示す。304 のうち 113 のアクセスでデータが取得できなかった。データが取得できた 191 アクセスのうち 7 割強の 146 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。

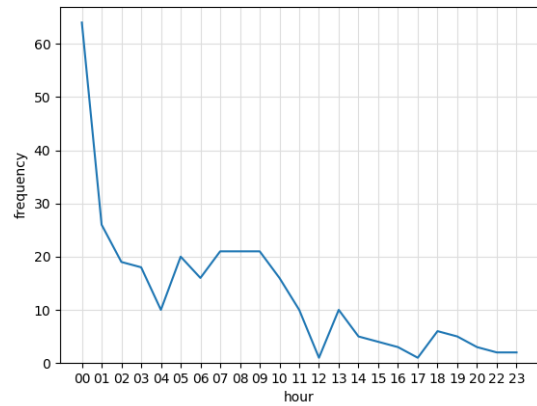


図6 アクセス時間帯の分布

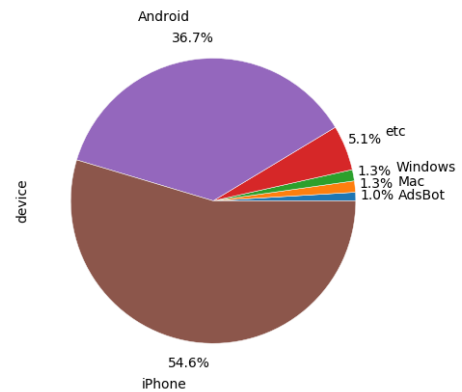


図7 デバイスごとのアクセス数

4 考察

4.1 広告クリック数とアクセス数の比較

Google 広告でのクリック数が 303 だったのに対し、表 3 より観測されたアクセスログ数は 414 であった。Bot からのアクセスの 3 件を除外すると 411 である。

さらに同一 IP アドレスからの 30 分以内のアクセスを 1 として数えると 107 減って 304 となる。この結果は、Google 広告のクリック数と 99.6% 一致する。

表 2、3 より広告クリック数が 303 だったのに対し、アクセスログ数は 414、ボットと 30 分以内の同一 IP アドレスからのアクセスを除いたアクセス数は 304 であった。同一アクセスと判断する条件を、5 分以内とすると 310、120 分以内とすると 301 となり、いずれにせよ 95% 以上一致する。

表4 スクロール率ごとのアクセス頻度

スクロール率	アクセス数
0.0	160
0.1	3
0.2	5
0.3	3
0.4	2
0.5	2
0.6	1
0.7	3
0.8	2
0.9	3
1.0	14

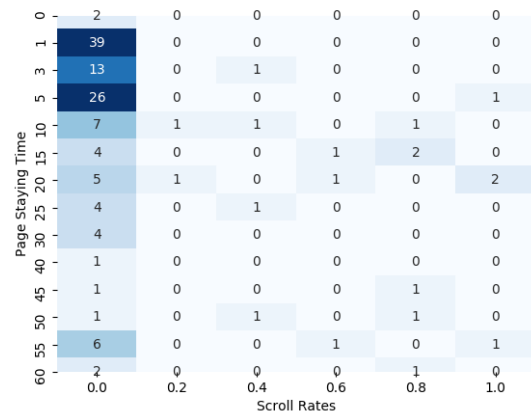


図8 ページ滞在時間(縦)とスクロール率(横)ごとのアクセス分布

表5 ページ滞在時間ごとのアクセス頻度

滞在時間(秒)	アクセス数
0	2
1	39
3	14
5	30
10	10
20	16
30	10
40	2
50	3
60	13
90	6
120	6
121以上	47

いと考えられる。

4.3 流入したユーザの振る舞い

ユーザの行動を JavaScript で取得したところ、約 37% 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。

取得できた 191 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図 8 に示す。表の左上(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(スクロール率 6 割、滞在 40 秒以上)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が、実際には左下の様に、スクロールはせずにページに長く滞在するユーザが多く見受けられた。

4.2 アクセスの地域、時刻、デバイス

Google 広告のアクセス地域特定も我々のプログラムと同様 IP アドレスから自動で判定される [5]。したがって図 5 より、アクセス数の差は使用する geoIP サービスによる誤差で生じると考えられる。

日本のインターネットの利用時間帯は総務省によると朝 6 時台から 23 時台までが 10% を超える時間帯である [6]。特に 12 時と 20 時から 22 時は 20% を超えて最も利用される時間帯である。しかし本実験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 10 人以上と比較的多い時間帯であった。よって、広告は必ずしもネット利用時間の分布に比例して提示されていない

5 おわりに

本論文では、Google ディスプレイ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 99.6% 一致した ($n = 303$)。一方、アクセスしたユーザの 21% が 1 秒以内にページを閉じ、76.4% 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。また、アクセス地域は判別に使用する API によってある程度の誤差が生じることも明らかになった。

ただ、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の実施など、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] 金井 文宏, 千葉 大紀, 高田 雄太, 秋山 満昭, 八木 毅, 波戸 邦夫, “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”, 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.
- [3] 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.
- [4] “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.)
- [5] “地理データについて”, (<https://support.google.com/analytics/answer/6160484?hl=ja>, 2020 年 12 月参照.)
- [6] 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.)
- [7] ADSTAGE, (<https://www.adstage.io/>, 2020 年 12 月参照.)