

明治大学総合数理学部

2020 年度

卒 業 研 究

アドネットワークにおける広告効果指標の調査

学位請求者 先端メディアサイエンス学科

草野蘭之介

目次

第 1 章	はじめに	2
1.1	背景	2
第 2 章	アドネットワーク	3
2.1	インターネット広告	3
第 3 章	関連研究	4
第 4 章	予備実験	5
4.1	目的	5
4.2	方法	5
4.3	結果	5
4.4	考察	5
第 5 章	実験	8
5.1	目的	8
5.2	方法	8
5.3	結果	9
5.4	考察	11
第 6 章	おわりに	15
	参考文献	17
付録 A	取引件数の時間分布の相関を用いた Bitcoin 取引所のユーザのタイムゾーン推定	18
A.1	はじめに	18
A.2	実験概要	18
A.3	実験結果	20
A.4	評価	20
A.5	おわりに	21
	参考文献	24

第 1 章

はじめに

1.1 背景

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。

第2章

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 2.1 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idding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。

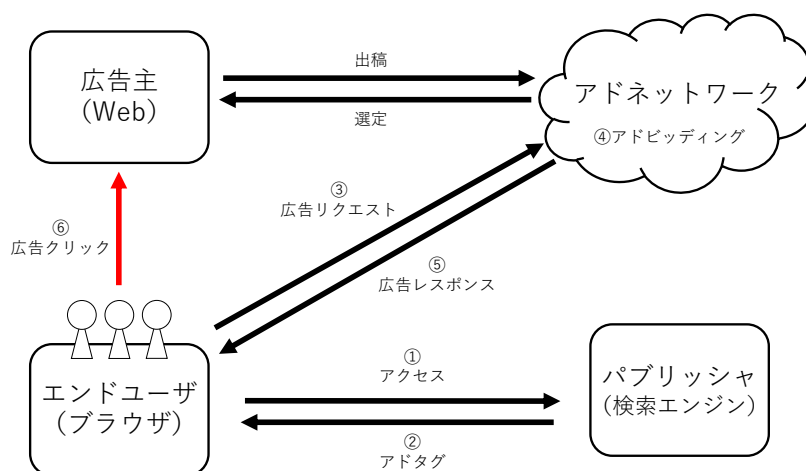


図 2.1 インターネット広告の仕組み

第 3 章

関連研究

金井らは、広告不正における、ユーザによる広告閲覧と広告クリックに着目した調査を行った [2]。既存研究では、広告を閲覧するユーザの立場から収集されたデータを分析した研究が多かった。これは広告不正の具体的な手口を調査する上では有効であるが、実際の広告不正全体の傾向を調査する上では不十分であるという問題点があった。そこで金井らは、広告ネットワーク上で観測された、ユーザによる広告閲覧および広告クリックのログを分析し、良性・悪性の広告ログの間にどのような特徴の差異があるかを調査し、広告不正の実態を効率的に検知・対策する技術の創出を目的とした。結果として、良性・悪性の広告ログの間には、パブリッシャサイトの URL やドメイン名、エンドユーザの IP アドレス、広告をクリックする座標などの観点から、特徴の差異があったと報告している。

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]。成長著しいインターネット広告市場では様々な収益モデルがあり、PPC (Pay-per-Click) 広告がメインとして挙げられるが、その PPC 広告に対する最大の脅威がクリック詐欺である。従来では、クリック詐欺に対抗するためにサーバ側からの不正検出技術を使用していたが、VPN や匿名ネットワークの場合、不正の検出が難しいという問題点があった。そこで Iqbal らは、ユーザ側からの調査が出来る新たな技術「FCFraud」を提案した。結果として、FCFraud は従来の技術の問題点であった、サーバ側からの技術を補完する技術として、クリック詐欺を行うボットネットの数そのものを減らす事ができたと報告している。

第 4 章

予備実験

4.1 目的

実際にアドネットワークに広告を出稿することで、どのようなデータが得られるのか、および本実験で役立つような、広告の効果指標を発見することを目的とする。

4.2 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。2020/9/3 から 2020/9/4 の 2 日間、Google 広告*で図 5.2 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 5.3 の実験用コンテンツページへ遷移する。予備実験では、広告表示回数や広告クリック数などの Google 広告で取得できるデータを収集する。また、より詳細なユーザの振る舞いを調査するために、実験用コンテンツページ内に Google アナリティクス[‡]と紐づけるためのタグを挿入した。

4.3 結果

2 日間の集計結果は、表示回数が 24,270 回、クリック数が 227 回、クリック率が 0.94% だった。Google 広告および Google アナリティクスで取得できる主なデータを表 4.1 に、Google 広告、Google アナリティクスの管理画面の例を図 4.1、図 4.2 に示す。

4.4 考察

本実験で実際に役立つような広告効果指標として、スクロール率、ページ滞在時間、直帰率を発見した。スクロール率について、図 4.3 のように、パーセンテージが高いほどページ全体を見ていることが分かる。つまり、パーセンテージが高い人が多いほど興味がある人が多く見ていることが分かる。ページ滞在時間は、ユーザの一回のサイト訪問における滞在時間の平均時間を表しており、この時間が長いほどユーザが興味を示していることが分かる。直帰率については、サイト内の行動においてユーザが 1 ページしか閲覧せずに離脱した割合を示しており、この値が低いほど、ユーザはサイト内で様々な行動をすることが分かる。以上の 3 種類のデータが、良質なアドネットワークの判断基準になり得ると考えた。

表 4.1 予備実験で取得した情報の一覧

取得元	データの例
Google 広告	アクセス数 表示回数 デバイス 性別 年齢 地域 時間帯 曜日
Google アナリティクス	アクセス数 表示回数 滞在時間 直帰率 リファラー スクロール率

クリック数 ▼ 303	表示回数 ▼ 7.28万	平均クリック単価 ¥26	費用 ¥7890	⋮ ▼
----------------	-----------------	-----------------	-------------	-----

キャンペーン Sort: ステータス ↑↓

display-movie-research	費用 ¥7890	表示回数 7.28万	クリック数 303	平均クリック単価 ¥26	⋮ ^
------------------------	-------------	---------------	--------------	-----------------	-----

広告のプレビュー



< 1 / 1 >

掲載結果 費用 ▼



ステータス

① 一時停止中のキャンペーン

このキャンペーンは一時停止されており、現在広告が表示されていません。広告の掲載を開始するには、このキャンペーンを有効にしてください。

[キャンペーンを有効にする](#)

● 一時停止 ▼ 目標: ブランド認知度とリーチ タイプ: ディスプレイ 予算: ¥1,000/日 [キャンペーンの編集](#)

[キャンペーンを表示](#) 最適化スコア: -

図 4.1 Google 広告の管理画面の例



図 4.2 Google アナリティクスの管理画面の例

イベント ラベル ?	合計イベント数 ?	ユニーク イベント数 ?
	425 全体に対する割合: 100.00% (425)	342 全体に対する割合: 100.00% (342)
1. 10%	203 (47.76%)	147 (42.98%)
2. 20%	81 (19.06%)	54 (15.79%)
3. 30%	31 (7.29%)	31 (9.06%)
4. 40%	23 (5.41%)	23 (6.73%)
5. 50%	19 (4.47%)	19 (5.56%)
6. 60%	18 (4.24%)	18 (5.26%)
7. 70%	17 (4.00%)	17 (4.97%)
8. 80%	17 (4.00%)	17 (4.97%)
9. 90%	16 (3.76%)	16 (4.68%)

図 4.3 Google 広告におけるスクロール率の表示

第 5 章

実験

5.1 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。

5.2 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 5.2 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 5.3 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 5.1 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 5.1 に示す。

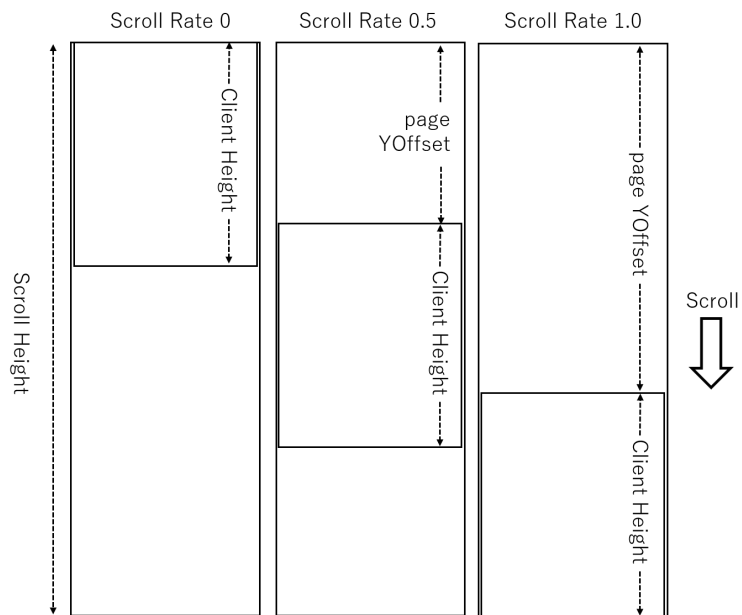


図 5.1 スクロール率の取得指標

表 5.1 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

表 5.2 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

5.3 結果

Google 広告に 7 日間出稿した結果を表 5.2 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 5.3 に示す。

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

†GeoLite2 City, <https://dev.maxmind.com/geoip/geoip2/geolite2/>

‡Google アナリティクス, <https://analytics.google.com/analytics/web/provision/?hl=ja#/provision>



図 5.2 実験に使用したディスプレイ広告



図 5.3 実験サイト

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 5.4 に示す。GeoLite2 では、304 アクセス中 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 5.5 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 5.6 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 5.5, 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

表 5.3 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

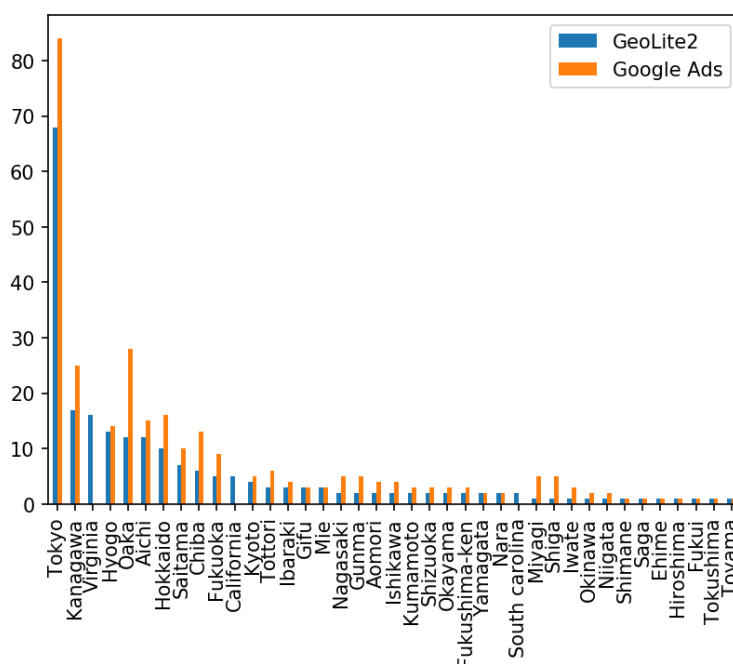


図 5.4 アクセス地域の分布

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。

5.4 考察

5.4.1 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 5.3 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 5.2, 5.3 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。

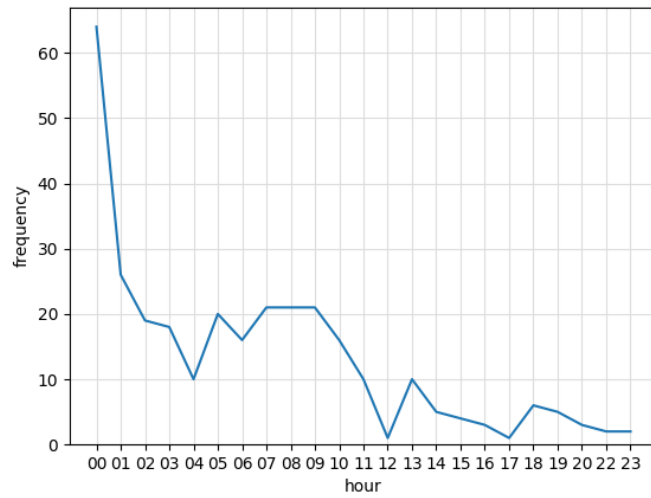


図 5.5 アクセス時間帯の分布

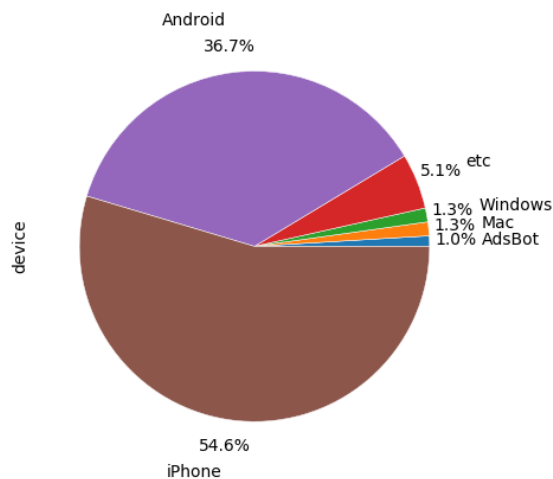


図 5.6 デバイスごとのアクセス数

5.4.2 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 5.4 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，

表 5.4 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

表 5.5 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

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と考えられる。

5.4.3 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 5.7 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

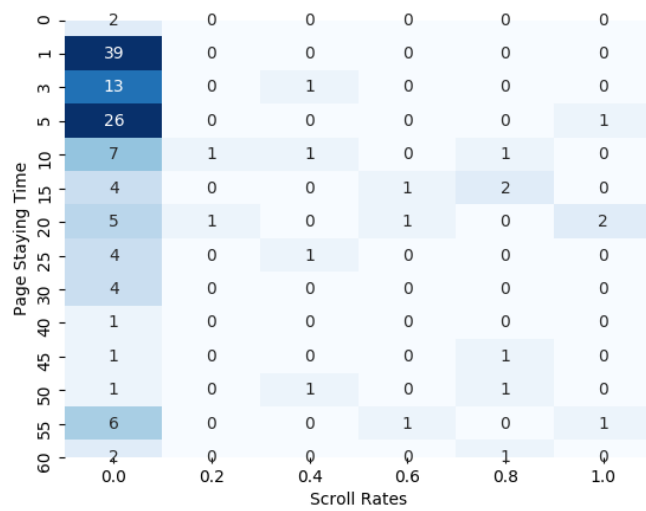


図 5.7 ページ滞在時間 (縦) とスクロール率 (横) ごとのアクセス分布

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。

第 6 章

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。

謝辞

本研究を行うにあたり，多くの方より御指導いただきました．特に明治大学総合数理学部先端メディアサイエンス学科，菊池浩明教授に深く感謝申し上げます．共に研究に取り組んでくださった柴山りな氏，山崎孝順氏，予備実験等に協力してくださった菊池研究室の皆様並びに先端メディアサイエンス学科の方々に深く感謝の意を表するとともに，謝辞とさせていただきます．

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;30:e1871, 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] 石川善一郎, 奥牧人, 河野崇, “Web リスティング広告における基本広告データを用いたコンバージョン分析と予測”, DEIM Forum, 2017
- [8] ADSTAGE, (<https://www.adstage.io/>, 2020年12月参照.)

付録 A

取引件数の時間分布の相関を用いた Bitcoin 取引所のユーザのタイムゾーン推定

A.1 はじめに

近年注目されている暗号資産 Bitcoin は、匿名性とオープンソースによる透明性が高いことが特徴である。しかし、各ユーザのトランザクション情報は公開され、共有されているので、その情報から各ユーザの属性を推定することが出来るのではないかと考える。実際に、Bitcoin に関する事件はこれまでにいくつか起きており、2014 年のマウントゴックス事件、2015 年の BTER ハッキング事件や 2018 年の CoinCheck 事件などがある。これらはシステムの脆弱性や杜撰な管理体制などが原因で生じたと考えられる。

2015 年に Dupont ら [1] は取引時刻に着目し、取引時間分布からアドレスを管理するユーザの居住地のタイムゾーンが特定できることを示した。また、井垣ら [2] は Dupont らの研究をもとに、Bitcoin のオンラインフォーラムである Bitcointalk のユーザに対して、77% のタイムゾーン推定が成功することを示した。

本研究では、ある国の取引所に属するユーザは、本当にその国に居住しているのかどうかを明らかにすることを目的とする。アドレスに関する取引データを Bitcoin のブロックチェーンから取得し、世界 80 か所の取引所の取引時間分布とそれらに属するアドレスの取引時間分布の相関を求め、ユーザの属性推定を試みる。

A.2 実験概要

A.2.1 実験目的

本実験の目的は、取引所のユーザ集合の属性国を推定することである。図 A.1 に推定の原理を示す。取引所 E に属するユーザ 3 の取引時間分布 a_3 とインターネット利用時間分布 T_1, T_2, T_3 を比較して、最も相関の大きな時刻 T_1 を推定する。その取引所全ユーザ a_1, a_2, a_3 の推定時間分布を求める。

A.2.2 実験データ

本実験で使用したデータは以下の 3 点である。

1. 各取引所の取引データ ([3])
2. アドレス毎の取引データ ([3])

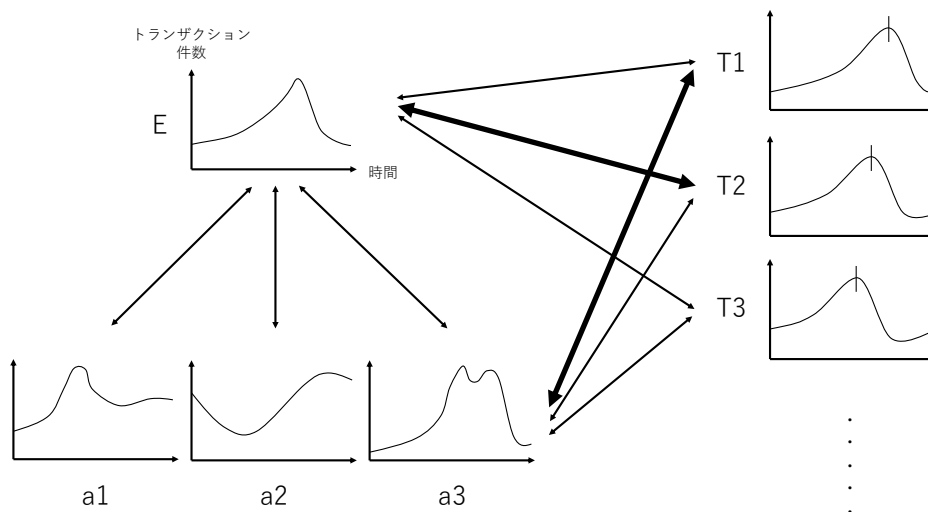


図 A.1 時刻推定の原理

3. インターネット利用時間帯データ ([4])

A.2.3 実験手法

データの取得

各取引所の取引データと各取引所に属するアドレス毎の取引データを用い、取引についての時間分布を求める。各取引所の取引データは Bitcoin のオンラインフォーラムである Wallet Explorer[3] のトップページの Exchanges 欄に掲載されている各取引所の取引データをスクレイピングにて取得した。各取引所に属するアドレス毎の取引データも、同様に Wallet Explorer からスクレイピングにて取得した。

また、総務省ホームページの 2014 年から 2019 年の調査結果報告書 [4] より、日本におけるインターネットの利用時間帯データを取得し、6 年間の平均を UTC にシフトしたデータを参照用とする。総務省ホームページを参照した理由としては、日本において Bitcoin 取引はスマートフォンやパソコン上で行われるため、インターネットの利用分布と関連性があると考えた。

各データの比較

各取引所の取引データの時間分布とインターネット利用時間のデータを比較し、相関係数のピーク時との誤差を求める。各取引所に属するアドレス毎の取引データについても同様にして、推定分布を求める。

ユーザの属性特定

取引所が属する国のタイムゾーンを正解と定義する。正解と取引所の推定タイムゾーンの比較により、差が 2 時間以内のものを推定成功とする。

更に、世界 80 か所の取引所に属する全てのアドレスとインターネット利用時間のデータを同様に分析する。アドレスの推定タイムゾーンと、取引所の推定タイムゾーンとの誤差が 2 時間以内であれば、そのアドレスの

表 A.1 インターネット使用時間データと Coincheck の誤差表

比較対象	差	トランザクション件数
全年度	-1	987
2016	-1	477
2017	-1	320
2018	+4	100
2019	+6	90

表 A.2 推定値と正解との誤差

推定値と正解との差	取引所数
0	19
1	36
2	10
3	4
4以上	11

ユーザはその国に居住している可能性が高いと考える。

A.3 実験結果

表 A.1 では、インターネット利用時間データと Coincheck データの誤差表である、両データを UTC に揃え、インターネット利用時間データを 1 時間ずつずらしていき、相関が高い時のズレの値を誤差値としている。インターネット利用時間は UTC+9 であるので、全年度、2016 年度、2017 年度の 3 つのデータはそれぞれ UTC+8 となる。

表 A.3 に主要な取引所の推定タイムゾーンとその取引所の取引数などの統計量を示す。

図 A.2 では、取引所 Exchanging.ir の年毎の取引数分布を可視化した図である。縦軸は取引件数、横軸は時間を表す。この取引所では 2015 年において取引件数が一番多く、約 1 万 3 千件行われていた。なお、この図では年間取引件数が 1000 回以上の年のみの可視化を行っており、図の右上に年毎の総取引数を記載している。

図 A.3 は、取引所 Coincheck(全年度) とインターネット利用時間の相関係数の分布である。正解 T は UTC+9 であるが、推定 T^* は UTC+8 であった。

図 A.4 は、取引所 BX.in.th とそれに属する全アドレスの相対誤差分布を示す。図 A.5 は、取引所 AnxPro とそれに属する全アドレスの相対誤差分布を示す。誤差範囲は +12 から -11 である。アドレスは取引回数が 15 回以上行われたアドレスが対象である。15 回以上取引が行われたアドレスは、取引所 BX.in.th において 6848 個存在し、取引所 AnxPro において 1546 個存在した。

A.4 評価

取引所 BX.in.th は、推定値が正解と同じ値であった。また、図 A.4 より取引所 BX.in.th に属するアドレスは、取引所の推定値と同値をとるアドレスが多いことがわかる。同値(差が 0)であるアドレスは 1044 個。差が 1 であるアドレスは 1619 個。差が 2 であるアドレスは 928 個。取引所 AnxPro は、推定値 -7、正解は +8 と大幅にずれている。しかしながら、図 A.5 より取引所 AnxPro に属するアドレスは、取引所の推定値と同値

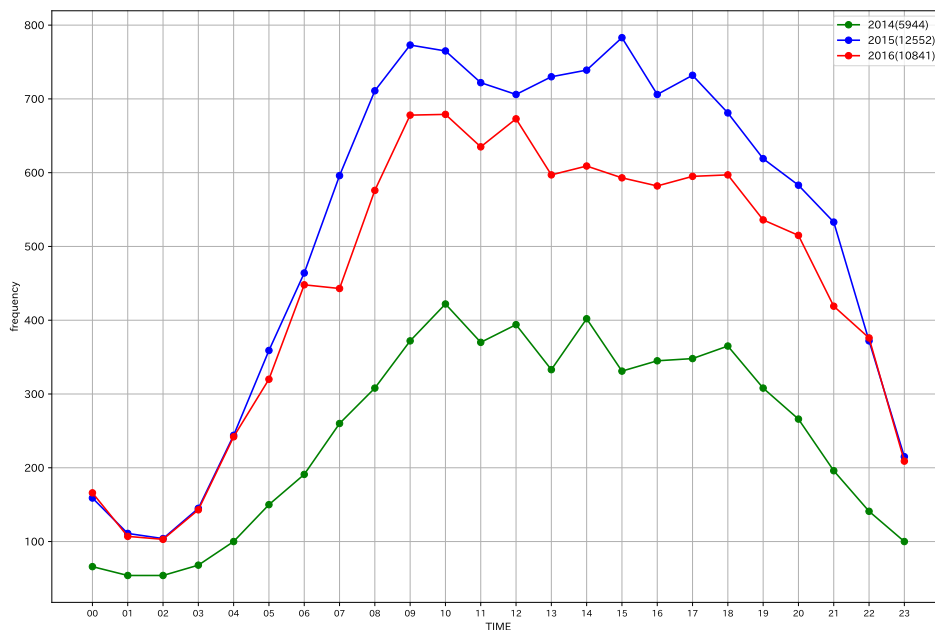


図 A.2 取引所データの年毎の取引数分布

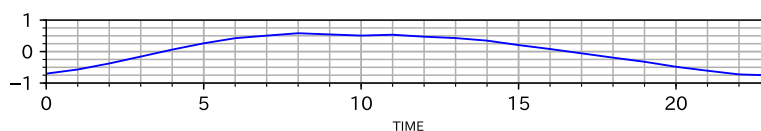


図 A.3 Coincheck の相関による時間分布

をとるアドレスは少なく、本来の正解である UTC+8 付近の推定値であるアドレスが多いことがわかる。

表 A.3 の実験結果の推定値と正解の比較により、80 取引所中差が 0 の取引所が 19 箇所あり、差が 1 の取引所が 36 箇所あり、差が 2 の取引所が 10 箇所となった。

A.5 おわりに

本研究では、推定値を求めるにあたってインターネット利用時間を使用し、取引所の利用者の所属国を推定した。推定値と正解の差が 2 の範囲内である取引所は、80 取引所中 65 取引所であった。このことから、取引所のユーザはその取引所が属する国に生活する可能性が高く、インターネット経由で Bitcoin 取引所を利用するユーザが多いと考えられる。

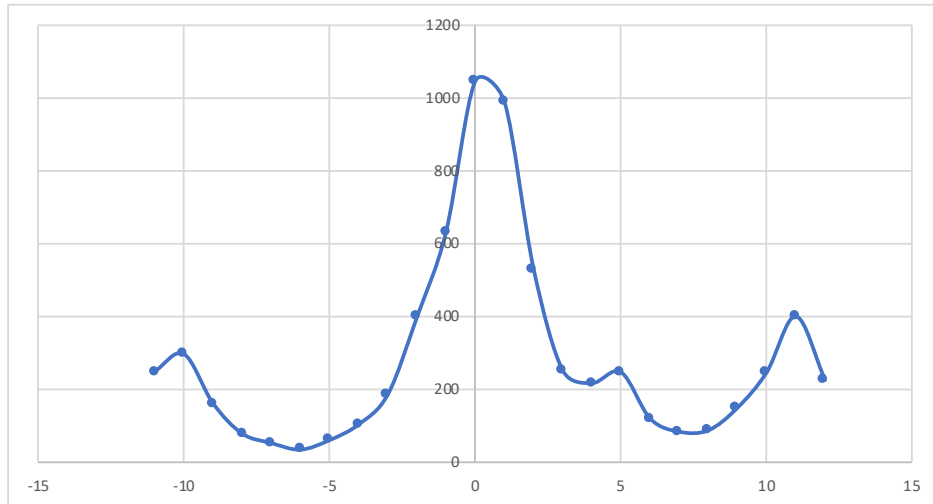


図 A.4 取引所 BX.in.th のアドレス分布

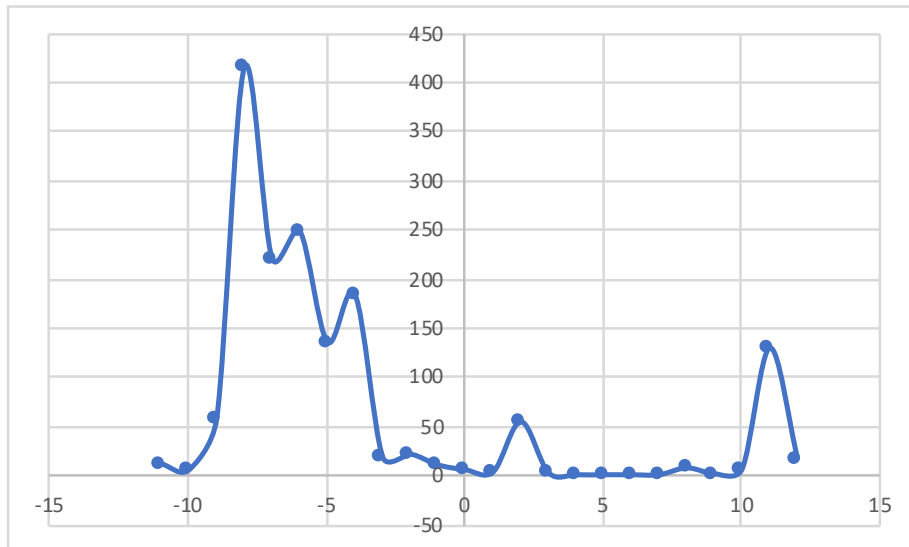


図 A.5 取引所 AnxPro のアドレス分布

表 A.3 タイムゾーン推定結果と取引数 (一部)

国	取引所名	UTC(正解)	UTC(推定値)	総取引数	総アドレス数
香港	oneCoin	+8	+7	1922	876
香港	AnxPro	+8	-7	326265	122166
ケイマン諸島	BlockTrades	-5	-5	52426	27748
タイ	BX.in.th	+7	+7	323191	48571
デンマーク	Ccedk	+1	0	2645	1040
シンガポール	CoinHako	+8	+10	136698	20588
イギリス	Coinmate	0	0	43042	5297
ドイツ	Coinomat	+2	+1	14688	5481
アメリカ	EmpoEx	-7	+1	2688	741
イギリス	Exmo	0	+1	90813	8430
ブラジル	FoxBit	-3	-4	91641	3670
シンガポール	FYBSG	+8	+5	63221	20449
中国	ChBtc	+8	+8	30083	2744
アメリカ	CoinCafe	-5	-6	18961	4547
カナダ	Coins-e	-5	-2	28437	11221
デンマーク	ExchangeMycoins	+1	+1	784	436
中国	Banx.io	+8	+1	10704	3621
ベトナム	BitcoinVietnam	+7	+6	6238	2276
ポーランド	Bitcurex	+1	+5	45710	4354
クロアチア	BitKonan	+1	+1	13014	4530
メキシコ	Bitso	-6	-5	3262	1188
中国	Btc38	+8	+8	82783	7585
オランダ	CleverCoin	+1	+1	9647	2614
中国	CoinTrader	+8	+6	48553	1497
トルコ	Exchanging.ir	+3	+2	35868	5
香港	Gatecoin	+8	+3	8294	1711
イギリス	Indacoin	0	+3	7457	2221
カナダ	QuadrigaCX	-8	-6	11535	4516
イギリス	Vaultoro	0	0	36447	13581
中国	Coingi	+8	+12	7457	812
ブラジル	UseCryptos	-3	-1	11535	700
イギリス	Cryptonit	0	+1	4605	1545
ポルトガル	Zyado	0	-1	1985	590

参考文献

- [1] J. Dupont, A. C. Squicciarini, “Toward De-Anonymizing Bitcoin by Mapping Users Location”, In Proceedings of Conference on Data and Application Security and Privacy(CODASPY’ 15), pp.139-141, ACM, 2015.
- [2] 井垣秀星, 永田倅大, 菊池浩明, “平均取引時間分布の相関を用いた Bitcoin ユーザのタイムゾーン属性の推定”, 情報処理学会第 81 回全国大会, pp.3 481-3 482, 2019.
- [3] Wallet Explorer (<https://www.walletexplorer.com>, 2019 年 6 月)
- [4] 総務省-情報通信メディアの利用時間と情報行動に関する調査報告書 (2019 年 11 月)