

Residential IP Proxy サービスを用いた位置情報・アカウント利用時のターゲット広告の調査

福田 ひかり†

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

近年、プロキシサービスの中でも住宅用ネットワークを利用したトラフィック中継を提供する Residential IP Proxy (以下 RESIP とする) が人気を博している。RESIP に利用されている住宅用 IP アドレスは、住宅用 IP アドレスの保有者が各プロキシサービスのネットワークに自主的に参加することで収集されているとしている。また、プロキシサービスプロバイダではアプリの開発者を対象に収益の手段として SDK (Software Development Kit) を提供している。

Mi らの調査 [11] よりアプリ利用者が RESIP ホストへの参加を承諾する際の同意書の文章はトラフィックの中継動作について曖昧であることや、Mi ら [10], 半澤ら [13], 住友ら [14] の調査により悪用されている可能性が判明した。先行研究 [15] のゲスト時における位置情報、ターゲット広告の調査により、RESIP ホストが RESIP 利用者に影響を与えることが明らかになったが、Google アカウントなどにログインした際に影響を受けるか不明である。そこで、Google アカウントなどにログインして RESIP を利用した際に RESIP ホストの影響を受けるか、位置情報とターゲット広告を用いて調査を行う。

2 Residential IP Proxy

2.1 システム構成

RESIP はプロキシサービスの 1 種であり、動的 IP アドレスである住宅用 IP が提供されることで検閲回避やスクレイピングなどを目的として利用されている。静的 IP アドレスに比べて一般利用者とプロキシ利用者の判別が付きにくいことでサイトに拒否されにくいことが特徴である。

Bright Data におけるプロキシサービスの構造を図 1 に示す。クライアント、プロキシゲートウェイ、プロキ

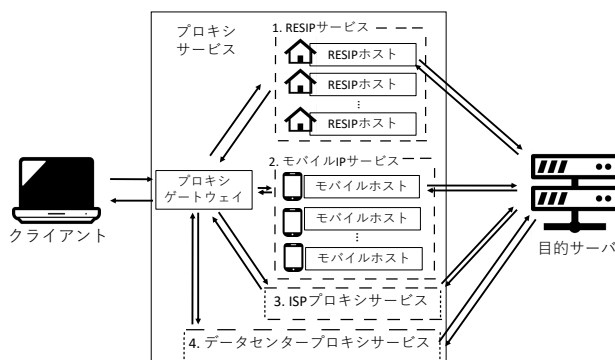


図 1 Bright Data におけるプロキシサービスの構成図

シホストから構成されている。プロキシを利用する際、ゲートウェイは RESIP 利用者から来た通信をに応じて RESIP サービス、モバイル IP サービス、ISP プロキシサービス、データセンタープロキシサービスへと任意に割り当てる。目的サーバとの通信は割り当てられたプロキシホストを経由して RESIP クライアントへ中継される。

Proxyrack と Oxylabs のシステム構成は Bright Data と同じである。Proxy-Seller は契約時に固定の RESIP ホストが割り当てられ、契約期間中は RESIP ホストの交換がない。プロキシ利用時はクライアントからゲートウェイを通らず RESIP ホストに接続し通信が中継される。

2.2 RESIP プロバイダ

本実験で利用した RESIP プロバイダは表 1 の Bright Data, Proxyrack, Oxylabs, Proxy-Seller の 4 社である。Bright Data, Proxyrack, Oxylabs の 3 社は先行研究 [10][11][13][15] にて調査されており、RESIP の所在国とホスト数が多いことから選定した。Proxy-Seller は他の 3 社がアクセスごとに任意にプロキシホストを割り当てることに対して、契約時に指定した地域のアドレスが固定して提供されることから選定した。

表 1 に RESIP プロバイダの主要な機能や特徴を示す。ここで提供しているプロキシの種類は、図 1 のプロキシゲートウェイを介して接続するサービスであり、Bright

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

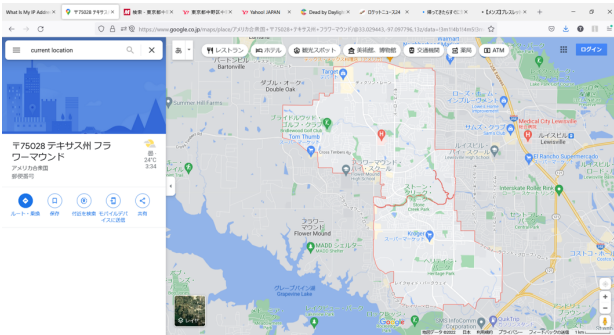


図2 サイト内検索による現在地推定例 (Google Maps)



図3 現在地情報に基づく現在地推定例



図4 プロキシなしのターゲット広告例

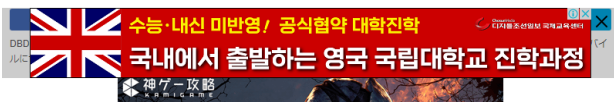


図5 プロキシありのターゲット広告例

Data は 4 種類に対して Proxyrack は 2 種類, Oxyllabs と Proxy-Seller は 3 種類ある。Proxy-Seller は契約時のみ利用する REIP ホストの所在国を指定することができるが, 他の 3 社は利用時に指定が可能である。また, Proxy-Seller の特徴は 3 社よりも RESIP の所在国数は少ないが 1 ヶ月分の料金が安いことである。

3 実験

3.1 目的

本実験は RESIP サービスを用いることで位置情報とターゲット広告に影響を与えるのかを調査し, RESIP サービスの影響の程度を判明させることを目的とする。Google アカウントなどを利用しログインした状態で調

査することで実際に使用される状況に近づけ, 言語や広告のパーソナライズがされることで RESIP の影響を受けにくい状態で RESIP による影響を受けるか調査する。

3.2 方法

3.2.1 実験 1: 位置情報の調査

位置情報の調査は 2022 年 8 月 27 日から 9 月 11 日にかけて, 表 2 の gmap, ymap, mapfan の 3 つのサイトを利用し行う。RESIP は各 RESIP プロバイダにおいてアメリカ, 日本, 韓国, インドの 4 か国に所在するものを 5 個ずつ計 80 個に対して行う。位置情報で用いる用語を表 4 に示す。ここで, 現在地の推定には, 図 2 内に“現在地”や“Current location”と入力し検索, サイト内の現在地推定ボタンをクリックするかの 2 種類がある。現在地検索結果の例を図 2 に示す。アメリカのテキサス州が表示されている。現在地特定が成功した例を図 3 に示す。中央に囲まれている現在地を示すマークが表示され実験地 (明治大学中野キャンパス) と一致している。

位置情報では 3 つのサイトを利用し位置情報を調査したが, ymap と mapfan の検索結果は特定のお店や地域が表示されるため現在地検索については gmap の結果のみを示す。

3.2.2 実験 2: 手動でのターゲット広告の調査

調査は 2022 年 8 月 27 日から 9 月 11 日にかけて, ターゲット広告へプロキシの影響があるかを 5 つの web サイトを対象として 4 つのプロキシプロバイダを用いて調査する。RESIP は各 RESIP プロバイダにおいてアメリカ, 日本, 韓国, インドの 4 か国に所在するものを 5 個ずつ計 80 個に対して行う。

通常では図 4 のように広告が日本語表記であるが, プロキシ利用時に影響を受けると図 5 のように日本語以外の広告がされる。ターゲット広告の調査で対象とした 5 つの web サイトは表 2 の yahoo, game, folk, rocket, smart である。

プロキシの影響を受けているターゲット広告は, 次の 3 つの条件のいずれかが満たされている広告と定義する。

1. 日本語以外の言語で書かれている。
2. 日本語表記であり, 海外在住者向けの広告である (日本商品の通信販売や日本在住と同様のイン

表 1 RESIP プロバイダ比較

プロバイダ名	Bright Data	Proxyrack	Oxylabs	Proxy-Seller
本社所在地	イスラエル	香港	リトアニア	キプロス共和国
プロトコル	HTTP/HTTPS, Socks4, Socks5	HTTP/HTTPS, Socks4, Socks5	HTTP, HTTPS	Socks5, HTTPS
1ヶ月分の料金	USD15.00~/GB	USD49.95~/10GB	USD15.00~/GB	USD1.64~/IP
RESIP の所在国数	195 か国	195 か国	195 か国	50 か国
提供している プロキシの種類	Residential proxies, ISP proxies, Datacenter proxies, Mobile proxies	Residential, Datacenter	Datacenter Proxies, Residential Proxies, Next-Gen Residential Proxies	Proxy IPv4, Proxy IPv6, Mobile Proxy LTE
プラン名	Residential Proxies	25 Private Unmetered Residential Ports	Residential Proxies	Proxy IPv4
RESIP が交換される 間隔の指定	× (IP を交換しない ことは可能)	○ (5, 10, 15, 30, 60 分)	○ (10 分)	×
RESIP 所在地指定	○	○	○	契約時に国を指定 変更不可
リクエストごとに RESIP を変更	○	×	○	×

表 2 調査対象の web サイト

実験内容	web サイト	本実験での名称	URL	API
位置情報	Google Maps	gmap	https://www.google.co.jp/maps	Google Maps API
	Yahoo! 地図	ymap	https://map.yahoo.co.jp/	Yahoo! Open Local Platform
	MapFan	mapfan	https://mapfan.com	MapFan API
ターゲット広告	Yahoo! JAPAN	yahoo	https://www.yahoo.co.jp/	
	神ゲー攻略	game	https://kamigame.jp/	
	folk	folk	https://folk-media.com	
	ロケットニュース 24	rocket	https://rocketnews24.com/	
	Smartlog	smart	https://smartlog.jp/	

表 3 実験目的と日数, 方法

番号	日数	目的	方法
実験 1	2022 年 8 月 27 日～ 9 月 11 日 (16 日間)	位置情報が実験地から変化することを調査	gmap, ymap, mapfan にて現在地検索 と現在地特定を行う
実験 2	2022 年 8 月 27 日～ 9 月 11 日 (16 日間)	ターゲット広告の表示言語や 表示内容への影響を調査	5 つの web サイトで条件をみたく ターゲット広告が表示されるか調査する
実験 3	2022 年 11 月 24 日～ 12 月 1 日 (7 日間)	実験 2 の自動化による調査件数の増加と ターゲット広告の影響を調査	青山 [16] の自動観測プログラムを用いて 広告を取得し条件に合致するか確認を行う

ターネット環境の構築など)。

- 日本語表記であり, 現在地と異なる地名が入っている。

3.2.3 実験 3: 自動でのターゲット広告の調査

調査は 2022 年 11 月 19 日から 12 月 1 日にかけて, Python の Selenium を利用したシステム [16] を用いて手動と同じ動作で広告の観測を行う。条件と利用するサイトは実験 2 と同じものである。RESIP プロバイダ

は Bright Data, Proxyrack, Oxylabs を利用する。各 RESIP プロバイダにおいてアメリカ, 日本, 韓国, インドの 4 か国に所在する RESIP ホストを 25 個ずつ計 300 個に対して調査する。

3.3 結果

3.3.1 実験 1: 位置情報の調査結果

表 5 の○は全ての RESIP 利用時にプロキシの影響を受け表示結果が変化したものを示す。×は全ての

表 4 位置情報の調査で利用する用語

現在地検索	現在地特定がエラーまたはデバイスの位置情報にアクセスを拒否した際、図 2 の検索タブに“現在地”または“Current location”と検索すること
現在地特定	図 3 右下に囲われているマークをクリックして現在地を特定すること
現在地特定が成功	図 3 中央に囲まれているマークが表示され、実験地と一致すること
実験地	実験を行った地点. 国内 (明治大学中野キャンパス, 駿河台キャンパス, 自宅)
RESIP 所在地	調査する際 RESIP ホストの所在地として設定した国

RESIP 利用時にエラーとなり結果が表示されなかったことや結果が RESIP の所在国ではなかったことを示す。

現在地検索では Bright Data と Proxy-Seller の韓国とインドに所在する RESIP ホスト利用時には変化が見られなかった。firefox 利用時において現在地情報へのアクセスが許可されていた際には現在地検索で該当する結果が得られず、現在地情報へのアクセスが拒否されていた際には該当する結果が得られた。また、現在地特定では Firefox 利用時において、どのプロキシプロバイダを利用して現在地特定が成功した。

3.3.2 実験 2：手動でのターゲット広告の調査結果

表 6 に結果を示す。ここで、列が観測方法、行が訪れたサイトを示す。標的数はプロキシの影響を受けたのみならず条件のいずれかを満たしたターゲット広告の数であり、全広告数は表示されたターゲット広告の総数、率は全広告数のうち標的数の割合である。

yahoo は日本以外からのアクセスを制限しているため、yahoo 内で表示される広告が Yahoo! JAPAN 広告のみであることで広告の表示がされないため標的数、総広告数、率が他のサイトに比べて低い。同様に Yahoo! JAPAN 広告も日本ではない RESIP ホスト利用時には表示されない。Oxylabs における“×”は広告が表示されなかったことを示している。最も影響を受けなかったのは Bright Data を利用して rocket を訪れた組み合わせの 0.12 であり、最も影響を受けたのは Proxy-Seller を利用して smart を訪れた組み合わせである。プロバイダの比較をすると Bright Data が 0.42 で他の 2 社に比べて影響を受けていない。また、Proxy-Seller が rocket を除いた 3 つのサイトにおいて影響を受けた割合が最も高い。

3.3.3 実験 3：自動でのターゲット広告の調査結果

実験 3 の広告調査結果を表 6 に示す。Oxylabs の“×”は広告が表示されなかったことを示し、Proxy-Seller の“—”は実験を行っていないことを示す。Bright Data

は smart の 0.34, Proxyrack は game の 0.54 が最も率が高い結果となった。自動で 0.50 を超えたものは Proxyrack の game のみである。Bright Data は rocket 以外のサイト、Proxyrack では smart で手動より自動の率が低い。Proxyrack の yahoo を除いた全てにおいて総広告数は手動より自動が低い。

3.4 評価

3.4.1 自動の手動に対する相対誤差

表 8 に Bright Data と Proxyrack の標的数、全広告数、率における自動から手動を引いた相対誤差を示す。両方のプロバイダで標的数と全広告数が手動よりも減った。

Proxyrack の率の相対誤差が 0.00 であることに對して、標的数と全広告数の相対誤差は全広告数のほうが 1.3 小さいため、自動のほうが手動よりも影響を受けたターゲット広告が多いことがわかる。

3.4.2 考察

位置情報の検索について、Google Maps の現在地を認識する仕組み [8] はインターネット接続の IP アドレス、過去の使用履歴、ラベルを付けた場所、デバイスから位置情報を取得によって行われている。過去の使用履歴は地名を含んだ検索履歴を指し、ラベルを付けた場所は Google アカウントログイン時に Google Maps で自宅や職場を設定できることを指している。

Bright Data の検索結果は、直前の別の RESIP での検索結果と同じ地域が表示され、RESIP ホスト所在地と異なる場所であった。これは過去の使用履歴による影響を受けた結果であり、RESIP の影響は受けなかった。RESIP の影響を受けなかった原因は不明である。自動で全広告数が減ったことの原因として一部の広告が取得できていないこと、広告が短時間の表示であること、広告が表示されるまでに時間がかかることが考えられる。

Proxy-Seller のインドと韓国に所在する RESIP ホストについては、一部の IP アドレスの所在地などを確

表5 位置情報の調査結果

RESIP プロバイダ	Microsoft Edge		Google Chrome		firefox	
	現在地検索	現在地特定	現在地検索	現在地特定	現在地検索	現在地特定
Bright Data	×	○	×	○	×	○
ProxyRack	○	×	○	×	○	○
Oxylabs	○	×	○	×	○	○
Proxy-Seller	○ (アメリカ, 日本) × (韓国, インド)	×	○ (アメリカ, 日本) × (韓国, インド)	×	○ (アメリカ, 日本) × (韓国, インド)	○
通常	○	○	○	○	○	○

表6 ターゲット広告調査結果

観測サイト		Bright Data		ProxyRack		Oxylabs		Proxy-Seller		平均	
		手動	自動	手動	自動	手動	自動	手動	自動	手動	自動
yahoo	標的数	0.0	0.0	0.0	0.0	×	×	0.0	—	0.00	0.00
	総広告数	0.3	0.0	0.3	1.3	×	×	1.6	—	0.0	1.3
	率	0.00	0.00	0.00	0.00	×	×	0.00	—	0.00	0.00
game	標的数	2.7	2.2	3.4	4.2	×	×	5.9	—	4.0	2.7
	総広告数	7.0	8.6	11.0	8.2	×	×	11.1	—	9.7	8.5
	率	0.39	0.24	0.29	0.54	×	×	0.50	—	0.31	0.32
folk	標的数	1.3	0.3	0.8	0.7	×	×	0.6	—	0.9	0.4
	総広告数	2.3	1.2	2.1	1.5	×	×	2.0	—	2.1	1.3
	率	0.53	0.24	0.35	0.38	×	×	0.32	—	0.40	0.27
rocket	標的数	0.3	0.5	1.9	1.1	×	×	0.5	—	0.9	0.7
	総広告数	5.2	2.6	4.9	2.6	×	×	4.8	—	5.0	2.6
	率	0.04	0.17	0.38	0.39	×	×	0.10	—	0.17	0.22
smart	標的数	3.3	1.5	2.7	1.8	×	×	3.2	—	3.1	1.6
	総広告数	4.8	3.9	4.8	4.7	×	×	4.5	—	4.7	4.1
	率	0.70	0.34	0.57	0.40	×	×	0.70	—	0.65	0.35
平均	標的数	1.9	1.1	2.2	1.9	×	×	2.6	—	2.2	1.1
	総広告数	4.8	4.1	5.7	4.1	×	×	5.6	—	5.4	3.6
	率	0.41	0.25	0.40	0.40	×	×	0.40	—	0.41	0.23

表7 ゲスト時 [15] との率比較

	ゲスト時 [15]						アカウント利用時					
	Bright Data		ProxyRack		平均		Bright Data		Proxyrack		平均	
	手動	自動	手動	自動	手動	自動	手動	自動	手動	自動	手動	自動
yahoo	0.05	0.00	0.00	0.00	0.03	0.00	0.00	0.00	0.00	0.00	0.00	0.00
game	0.74	0.02	0.48	0.01	0.65	0.10	0.39	0.24	0.29	0.54	0.31	0.32
folk	0.82	0.02	0.68	0.01	0.74	0.19	0.53	0.24	0.35	0.38	0.40	0.27
rocket	0.54	0.05	0.70	0.06	0.61	0.09	0.04	0.17	0.38	0.39	0.17	0.22
smart	0.83	0.02	0.52	0.03	0.68	0.20	0.70	0.34	0.57	0.40	0.65	0.35
平均	0.68	0.02	0.55	0.02	0.62	0.13	0.41	0.25	0.40	0.40	0.41	0.23

表 8 自動の手動に対する相対誤差

プロバイダ	標的数	全広告数	率
Bright Data	-0.8	-0.7	-0.16
Proxyrack	-0.3	-1.6	0.00

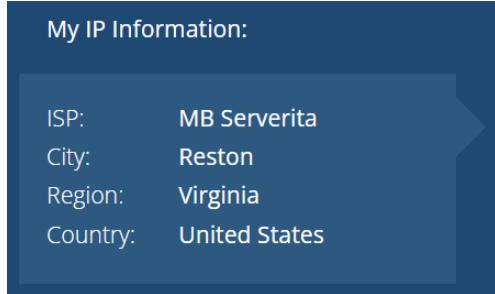


図 6 RESIP ホストの位置情報が異なった例

認できる web サイト [7] にて指定した国とは異なる国が表示された。韓国に所在する RESIP ホスト利用時に RESIP ホストの所在地として韓国以外の地域が表示された例を図 6 に示す。web サイト [7] で表示された RESIP ホストの所在地と現在地検索結果が同じ地域であったため、一部サイトにて RESIP ホストに紐づいている地域が実際の所在地と異なることが現在地検索が失敗した原因だと考える。

4 おわりに

実験より、アカウント利用時においても RESIP の影響を受けることがわかった。位置情報に関しては現在地推定の方法やブラウザによる差があることがわかった。ターゲット広告に関しては、Oxylabs 以外の調査において表示された広告のうち、RESIP ホストの影響を受けたターゲット広告が確認された。

本稿では RESIP サービスが RESIP 利用者へ与える影響を調査した。今後は位置情報について RESIP 所在地が現在地として表示される条件の解明や継続した RESIP の調査を課題とする。

参考文献

[1] Bright Data, (<https://brightdata.com/>, 2022 年 12 月参照).

[2] Proxyrack, (<https://www.proxyrack.com/>, 2022 年 12 月参照).

[3] Oxylabs, (<https://oxylabs.io/>, 2022 年 12 月参照).

[4] Proxy-Seller, (<https://proxy-seller.com/>, 2022 年 12 月参照).

[5] Bright Data, “Monetize inactive app users” (<https://brightdata.com/sdk>, 2021 年 12 月参照).

[6] ProxyRack, “Android Monetization SDK - ProxyRack” (<https://www.proxyrack.com/android-sdk/>, 2021 年 12 月参照).

[7] WhatIsMyIPAddress.com, (<https://whatismyip-address.com/>, 2022 年 12 月参照).

[8] Google, “ポリシーと規約” (<https://policies.google.com/technologies/location-data>, 2022 年 12 月参照).

[9] iplocation.net, (<https://www.iplocation.net/>, 2022 年 12 月参照).

[10] X. Mi, X. Feng, X. Liao, et al., “Resident Evil: Understanding Residential IP Proxy as a Dark Service”, 2019 IEEE Symposium on Security and Privacy (SP), pp. 1185-1201, 2019.

[11] X. Mi, S. Tang, Z. Li, et al., “Your Phone is My Proxy: Detecting and Understanding Mobile Proxy Networks”, NDSS Symposium 2021, pp.1-18, 2021.

[12] A. Tosun, M. De Donno, N. Dragoni and X. Fafoutis, “RESIP Host Detection: Identification of Malicious Residential IP Proxy Flows”, 2021 IEEE International Conference on Consumer Electronics (ICCE), pp. 1-6, doi: 10.1109/ICCE50685.2021.9427688, 2021.

[13] 半澤映拓, 菊池浩明, “Residential IP Proxy サービスに悪用される住宅用ホストの調査”, コンピュータセキュリティシンポジウム 2019, pp. 918-925, 2019.

[14] 住友孝彰, 菊池浩明, “Residential IP Proxy サービスを悪用した不正行為の調査”, 第 84 回情報処理学会全国大会, pp. 556-556, 2022.

[15] 福田ひかり, 井窪竜矢, 菊池浩明, “Residential IP Proxy サービスを用いた位置情報・ターゲット広告の調査”, 第 84 回情報処理学会全国大会, pp. 557-558, 2022.

[16] 青山綾佳, “Selenium を用いたターゲット広告の観測システムの開発”, 2022 年度菊池研究室卒業論文, 2023.