IP ブラックリストを用いた Residential IP Proxy ホスト検知手法の提案

北原 拓海 †

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，住宅用の IP アドレスを使用したプロキシアプリである Residential IP Proxy(RESIP) の利用が盛んになってきている．主な用途としては，IP アドレスに基づく検閲や，データスクレイピングによるアクセス制限を回避する場合などが挙げられる．

しかし RESIP はそのような用途の他にも，自身の身元を秘匿することを利用した不正アクセスや攻撃の踏み台としても悪用が疑われている．Mi らは 2017 年に RESIP ホストの 95% が住宅用の IP アドレスであり，その内の 43% が IoT 機器のものであることを報告している [2]．半沢らは国内のダークネットを観測し，所有する機器が意図せずに RESIP ホストとなり悪意を持った第三者に利用されている可能性があることを指摘している [3]．そのためユーザーは所有する機器が悪用されていることを検知して防ぐことが重要である．Tosun らは端末で取得したパケットの特徴を分析してホストで稼働している RESIP アプリを検知するアルゴリズム [4] を提案している．しかし，誤検知の頻度が高く，精度に問題があった．

そこで本稿では，RESIP について，従来の方法とは異なる IP アドレスのブラックリストを作成することで主要な RESIP アプリの検知を提案する．また，実験に基づく検出精度を報告する．

2　準備

本研究では Hola VPN[5], Proxyrack[6], Honeygain[7] の 3 つの RESIP アプリについての調査を行った．

2.1　Hola VPN

Hola VPN は 2008 年頃にイスラエルで開発された RESIP アプリである．無料でクライアントとなる場合には自身のネットワークのリソースを提供して Brightdata の RESIP ホストになる旨がソフトウェアの利用時に表示される．

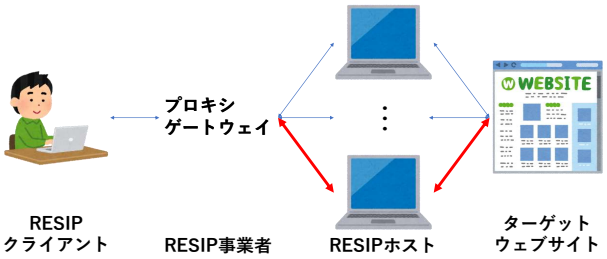

図 1　実験環境

2.2　Proxyrack アプリ

Proxyrack は 2014 年に開発された有料の RESIP アプリである．Proxyrack の RESIP ホストとなることで報酬を受け取る事 Proxyrack アプリを配布をしている．

2.3　Honeygain

Honeygain は 2018 年に開発された，RESIP ホストとなることで仮想通貨などの報酬を受け取ることが可能なアプリである．Honeygain は RESIP アプリの Oxylab との独占契約を締結しており，Honeygain を使用して RESIP ホストになったユーザのネットワークリソースは Oxylab を利用する RESIP クライアントに提供される．

3　提案手法

3.1　予備調査

自身が RESIP ホストになっていることを判断するには，ホストでパケットを観測してその通信路の情報を調査する方法がある．RESIP ホストとなって様々なアドレスにアクセスする際，RESIP 事業者のゲートウェイとの定期的な通信が行われる．

そこで本調査では 3 つの RESIP アプリのホストとなって各アプリについて 5 分× 100 回の通信を観測して IP アドレスを収集し，各 RESIP アプリの通信の特徴を定量化する．

---

†Takumi Kitahara, Proposal on Node Detection to be used as RESIP Host based on IP Black list, Department of Frontier Media Science, School of Interdisciplinary Mathematical Science, Meiji University, Kikuchi Laboratory.

表 1　調査した RESIP の開始年，対応プラットフォームと使用している SDK

名称	開始年	対応プラットフォーム	RESIP プロバイダ
Hola VPN	2008 頃	Windows, Android, iOS, Mac	Brightdata
Proxyrack	2014	Windows	Proxyrack
Honeygain	2018	Windows, Android, iOS, Mac	Oxylab

入力 調査対象 PC で取得した通信先 IP アドレス *ip*，IP アドレスのリスト *list*

1. RESIP アプリを起動して 100 秒待機する
2. 5 分間パケット *ip* を取得する
3. *ip* がプライベート IP アドレスではなければリスト *list* に記録する
4. RESIP アプリを終了する
5. 1. から 4. を 100 回繰り返す

図 2　観測手順

表 2　観測ログの例

Time	IP	Count
2022/11/23_5:21:56	20.190.141.32	73
	20.54.89.15	59
	20.205.248.73	28
	20.54.89.106	24
	52.191.219.104	24
	52.148.82.138	17
	20.72.205.209	12
	52.109.8.44	12

3.2　観測方法

パケットの観測と通信先 IP アドレスを収集する．Python で作成した観測手順を図 2 に示す．本観測では RESIP アプリを起動した後 100 秒待機してからホストの通信を 5 分間観測し，その後 RESIP アプリを再起動することで，確立された接続をリセットするようにした．

表 2 に本ツールを動作させた時のログを示す．通信の観測を開始した時刻と，5 分間で観測したパケットの数で降順に並べられた IP アドレスリストを出力し，次の 5 分間の観測結果を下に追加する．図 2 の例は RESIP アプリを起動しない状態で取得したログの一部が表示されている．この 5 分間で一番通信が多く観測された IP アドレスは 20.190.141.32 であり，その量は 73 個であったことを示している．

また第 1 オクテットが 20 または 52 の IP アドレスが多く観測されているが，これらはいずれも Microsoft のアドレスである．

入力 調査対象 PC で取得したパケット *p*，通信先 IP アドレスのリスト *a*

1. 5 分間パケット *p* を取得する
2. プライベート IP アドレス以外の通信先 IP アドレスをリスト *a* に記録する
3. *a* をブラックリストと照合する

図 3　提案プログラムのアルゴリズム

4　RESIP 検知プログラムの開発

4.1　概要

本ツールは前述の観測ツールで収集された IP アドレスを元に作成されたブラックリストを使用し，RESIP アプリとの関係が疑われる通信先との通信を検知する．Python で作成したこの提案方式のアルゴリズムを図 3 に示す．

4.2　実験方法

本実験では調査対象の RESIP アプリの通信を観測して定期的にアクセスを行う IP アドレスを記録することで，各 RESIP アプリの通信の特徴と RESIP 検知プログラムのブラックリストに登録すべきアドレスについて調査を行う．実験で使用したアプリを起動する OS は Windows 10 である．調査した RESIP アプリは Hola VPN, Proxyrack, Honeygain の 3 つである．

(1) 表 1 の 3 つの RESIP アプリついて，5 分のパケット収集を 100 回行う．

(2) 実験 1 で収集した IP アドレスについて，継続的に通信が行われていたものを IP ブラックリストに記録し，作成した RESIP 検知プログラムの精度を調査する．

表 3　実験で使用した RESIP アプリ及び収集した IP アドレスとパケット数

使用した RESIP	IP	パケット
なし (通常時)	48	1786
Hola VPN	141	11192
Proxyrack	325	89416
Honeygain	703	208820

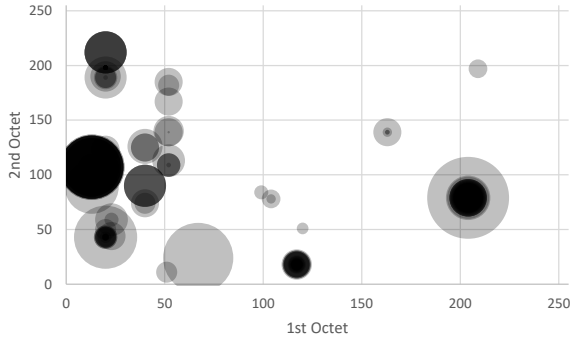

図 4　RESIP ホストではない状態での通信先と通信量

4.3　実験 1 結果 (通信先)

4.3.1　通常時

RESIP を起動しない状態での端末の通信先と通信量を図 4 に，100 回の観測で継続的に通信を行っていることが確認できた通信先を表 4 に示す．観測は 2022 年 11 月 24 日 5:53-14:06 に東京都の自宅から，家庭内 LAN に接続した Windows 10 端末で行った．

通常時の通信では 5 分 x100 回の観測で 48 のアドレスから 1786 のパケットを観測した．

継続的に通信を行っていた IP アドレスは，100 回中 27 回観測した 204.79.197.239 が最多だった．また通信先 IP アドレスを第 2 オクテットまででまとめた結果も同様に 204.79.x.x が最多であった．表 4 の上位の IP アドレスは Microsoft と CDN の EdgeCast(旧 Verizon) である．

4.3.2　Hola VPN

Hola VPN を起動した時の通信先と通信量を図 5 に，100 回の観測で継続的に通信を行っていることが確認できた通信先を表 5 に示す．2022 年 11 月 23 日 5:28-16:29 に観測した Hola VPN の通信は通常時の通信と比較すると通信先は 2.9 倍，パケットは 6 倍であり，通信先には通常時には見られなかった Dropbox, Amazon, DigitalOcean などのアドレスが多く見られた．

表 4　通常時の通信先 IP アドレスと whois 情報 (上位 10)

IP アドレス	whois
204.79.197.239	Microsoft Corporation (MSFT)
20.198.118.190	Microsoft Corporation (MSFT)
117.18.232.200	EdgeCast Networks Asia Pacific Network
13.107.5.93	Microsoft Corporation (MSFT)
20.43.132.130	Microsoft Corporation (MSFT)
204.79.197.200	Microsoft Corporation (MSFT)
20.212.97.243	Microsoft Corporation (MSFT)
117.18.232.240	EdgeCast Networks Asia Pacific Network
40.90.184.82	Microsoft Corporation (MSFT)
104.78.85.232	Akamai Technologies, Inc.

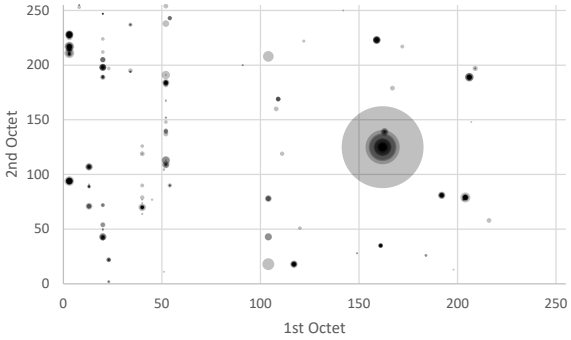

図 5　Hola VPN の通信先と通信量

表 5　Hola VPN の通信先 IP アドレスと whois 情報 (上位 10)

IP アドレス	whois
162.125.80.18	Dropbox, Inc. (DROPB)
3.94.72.89	Amazon Technologies Inc.
3.228.177.90	Amazon Technologies Inc.
3.228.36.186	Amazon Technologies Inc.
3.94.40.55	Amazon Technologies Inc.
206.189.231.23	DigitalOcean, LLC (DO-13)
40.70.229.150	Microsoft Corporation (MSFT)
20.198.119.84	Microsoft Corporation (MSFT)
192.81.214.145	DigitalOcean, LLC (DO-13)
159.223.133.120	DigitalOcean, LLC (DO-13)

4.3.3　Proxyrack

Proxyrack を起動した時の通信先と通信量を図 6 に，100 回の観測で継続的に通信を行っていることが確認できた通信先を表 6 に示す．2022 年 11 月 24 日 7:49-

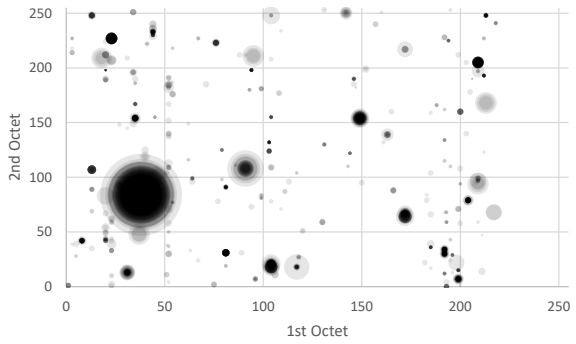

図 6　Proxyrack の通信先と通信量

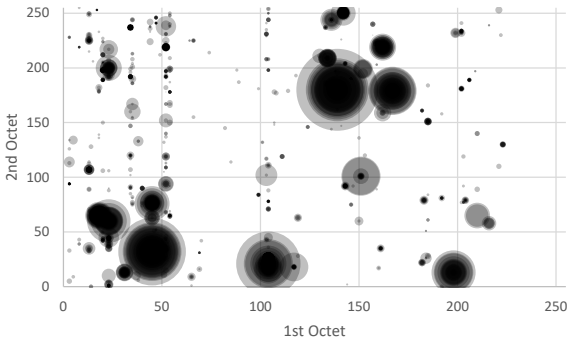

図 7　Honeygain の通信先と通信量

表 6　Proxyrack の通信先 IP アドレスと whois 情報 (上位 10)

IP アドレス	whois
38.84.70.82	PSINet, Inc. (PSI)
209.205.197.226	24 SHELLS (TS-74)
23.227.143.219	24 SHELLS (TS-74)
192.30.45.30	VeriSign Global Registry Services
192.34.234.30	VeriSign Global Registry Services
23.227.142.26	24 SHELLS (TS-74)
44.233.186.238	Amazon.com, Inc. (AMAZO-4)
104.21.57.231	Cloudflare, Inc. (CLOUD14)
199.7.54.30	VeriSign Global Registry Services
213.248.242.79	Nominet UK

表 7　Honeygain の通信先 IP アドレスと whois 情報 (上位 10)

IP アドレス	whois
34.237.55.225	Amazon Technologies Inc.
20.198.118.190	Microsoft Corporation(MSFT)
104.26.12.49	Cloudflare, Inc.(CLOUD14)
104.26.13.49	Cloudflare, Inc.(CLOUD14)
104.16.248.249	Cloudflare, Inc.(CLOUD14)
20.198.119.143	Microsoft Corporation(MSFT)
20.198.119.84	Microsoft Corporation(MSFT)
104.16.249.249	Cloudflare, Inc.(CLOUD14)
104.16.123.96	Cloudflare, Inc.(CLOUD14)
23.60.109.197	Akamai Technologies, Inc.

18:52 に観測した Proxyrack アプリの通信は通常時の通信と比較すると通信先は 6.8 倍, パケットは 50 倍であった．最も観測された回数が多かった IP アドレスは 100 回中 98 回観測された 38.84.x.x(PSINet) であり，その他には 24 SHELLS, VeriSign のアドレスが上位 10 アドレス中 6 つを占めた．

4.3.4　Honeygain

Honeygain を起動した時の通信先と通信量を図 7 に，100 回の観測で継続的に通信を行っていることが確認できた通信先を表 7 に示す．2022 年 11 月 28 日 1:30-13:22 に観測した Honeygain の通信は通常時の通信と比較すると通信先は 14.6 倍，パケットは 116 倍であった．最も観測された回数が多かった IP アドレスは 100 回中 89 回観測された 34.237.x.x(Amazon) であり，その他には Cloudflare のアドレスが上位 10 アドレスの半分を占めた．

4.4　実験 2　結果 (ブラックリストの作成)

調査 1 で収集した IP アドレスを元に RESIP ホスト検知プログラムに使用するブラックリストを作成した．登録した IP アドレスの割当国と whois 情報を表 8 に示す．登録した IP アドレスは，実験 1 での 100 回の観測のうち 80 回以上観測したアドレスの上位 16 ビットに限定した．ただし Honeygain のパケットの観測でのべ 190 回観測した 20.198.x.x は通常時でも観測されるアドレスのため，最終的にはその 1 つを除いた計 10 個の IP アドレスをブラックリストに登録した．

作成したブラックリストを用いて，収集したパケットを判定した結果を表 9 に示す．

RESIP ホストでない通常時のパケットで RESIP ホストであると誤判定される偽陽性は 100 回の実験では起こらなかった．

表 8　ブラックリストに登録した IP アドレスの割当国

IP アドレス	割当国	whois
3.228.x.x	アメリカ	Amazon Technologies Inc.
3.94.x.x	アメリカ	Amazon Technologies Inc.
162.125.x.x	アメリカ	Dropbox, Inc.
81.31.x.x	ドイツ	JAGEX
23.227.x.x	アメリカ	Leaf Group Ltd.
38.84.x.x	アメリカ	PSINet, Inc.
104.16.x.x	アメリカ	Cloudflare, Inc.
104.26.x.x	アメリカ	Cloudflare, Inc.
18.65.x.x	アメリカ	Amazon Technologies Inc.
34.237.x.x	アメリカ	Amazon Technologies Inc.

5　おわりに

本研究では RESIP アプリのホストとなる 3 つのアプリについて通信先 IP アドレスを観測することで，各 RESIP アプリが高頻度で定期的に通信する IP アドレスを確認した．またそれらの IP アドレスに基づいて，対象の端末が RESIP ホストとなっているかを高い精度で判別する方法を提案した．

参考文献

[1] Mirai ボットネットとは？ Cloudflare，(閲覧日：2022/11/29，https://www.cloudflare.com/ja-jp/learning/ddos/glossary/mirai-botnet/)

[2] Xianghang Mi, et al., “Resident Evil: Understanding Residential IP Proxy as a Dark Service”，2019 IEEE Symposium on Security and Privacy, 2019, pp. 1185-1201.

[3] 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，CSS2019，pp.918-925，2019.

[4] Altug Tosun, et al., “RESIP Host Detection: Identification of Malicious Residential IP Proxy Flows”，2021 IEEE International Conference on Consumer Electronics，2021.

[5] Get The Free and Premium Hola Online — Proxy Unblocker，(閲覧日 2022/11/15，https://hola.org/)

[6] Proxyrack: Buy Proxies HTTP, UDP, SOCKS Proxy，(閲覧日：2022/11/15，https://www.proxyrack.com/)

[7] Passive Income – Effortlessly — Honeygain，(閲覧日：2022/11/15，https://www.honeygain.com/)

表 9　RESIP 検知プログラムの精度

	提案手法			従来手法 [4]		
	Hola VPN	Proxyrack	Honeygain	Hola VPN	Proxyrack	Honeygain
TP	99	98	100	100	99	100
TN	100	100	100	88	88	88
Accuracy	0.995	0.99	1	0.94	0.935	0.94
F-score	0.994	0.989	1	0.943	0.938	0.943