Analysis on Malicious Residential Hosts Activities Exploited by Residential IP Proxy Services

Akihiro Hanzawa[1] and Hiroaki Kikuchi[1]

School of Interdisciplinary Mathematical Sciences, Meiji University,
Nakano, Tokyo, Japan
cs192013@meiji.ac.jp kikn@meiji.ac.jp

Abstract. A *residential IP Proxy* is a proxy service that provides a traffic relay using hosts on residential networks. Although the service providers claim that hosts voluntarily participate in the service and use it for various high-quality applications, in fact, the service provides avoiding detection and blocking by pretending as apparently benign users, they exploited the residential hosts to perform malicious acts such as DoS attacks. In 2019, Mi et al. studied that malicious hosts participating in the Residential IP Proxy service, and profiled the hosts, and clarified the infrastructure, scale, and malignancy of the such services. They found that most malicious activities were sending SPAMs and hosting fake websites that were performed by routers and WAP devices. However, residential WAP devices are commonly inside of firewall and these are not likely to be feasible in well managed residential networks.
To answer to the concern, in this paper, we analyze datasets of Residential-IP-Proxy hosts, collected by Mi et al. and report an analysis of the communication that Residential IP Proxies perform in Japan. We use NONSTOP, the analysis platform, provided by the Information Technology Research Organization, in the analysis. Our analysis found that most of devices used in Japan were mobile laptop PCs and port-scanning was the most frequent malicious activity. Consequently, more RESIP hosts are becoming involved in serious threat and we need countermeasures aimed at minimizing the abuse of RESIP hosts.

1 Introduction

Recently, a new service called *Residential IP Proxy as a Service* (RPaaS) have been provided in the market of proxy Internet connection via proxy hosts. Table 1 lists the major RPaaS service providers. RPaaS plays a useful role in enabling users access to arbitrary sites without any restriction. For example, Luminati, the largest Residential IP Proxy (RESIPs) service provider, is located in the United States, but has many clients who reside in Turkey, and who may be trying to avoid Turkey's network censorship. Web proxy services are studied for many researchers. Chung et al. studied a paid proxy services to be manipulating contents [13]. A measurement to reveal the purpose of proxy services was conducted by Weaver et al in [14].

In [1], Mi et al. reported that the presence of likely compromised hosts as residential IPs, identified from 6.18 million unique IPs, distributed over 238 countries and 52,905 ISPs. Among the hosts, they identified 237,029 IoT devices and 4,141 hosts running PUP networks. The traffic relayed via the RESIP involved ad clicking, SPAM messaging, and malicious IP hosting activities. They found that these malicious activities were performed by routers and WAP devices in residential networks. However, residential WAP devices are commonly inside of firewall and are not vulnerable to be compromised if these are under control.

Hence, our analysis of this study is motivated by the following questions.

1. What kinds of networks do RESIPs belong to (residential, institutional, or academic networks?)
2. How are RESIPs distributed geometrically in Japan, countryside, or metropolitan regions?
3. Who are the major RPaaS providers?
4. What is the impact of malicious RESIPs?
5. For what purposes are the RESIPs abused (advertisement, phishing, port scanning, or exploring)?

Our objective is to answer above research questions by investigating up-to-date RESIP activities.

To answer questions 1) and 2), we investigate the detailed properties of the RESIP addresses. For each of the IP addresses detected by Mi et al. [1] in 2017, we examine the geolocation query using the GeoLite2 city database [3] from MaxMind, Inc. We use the Registration Data Access Protocol (RDAP) service provided by the Asia–Pacific Network Information Center (APNIC) [4] to identify the domain and registry to which RESIP addresses belong.

To answer questions 3) to 5), we need to observe the malicious packets sent from the RESIP addresses. We, therefore, use the darknet database, NONSTOP [6], serviced by the National Institute of Information and Communications Technology (NICT). Using NONSTOP, we examine whether suspicious addresses detected as RESIPs had performed port-scanning to NICT's darknet. Since a darknet is unknown and unused network segment, we regard any packets designated for the darknet as malicious.

Our contributions of this work are as follows.

- We have found new trends in RESIP host activities based on the darknet traffic observed in Japan. Our new findings is that the main devices used in Japan were mobile laptop PCs, whereas router, firewall and WAP devices were identified from the profiles in the previous study [1].
- We have identify the malicious activities performed by RESIP hosts. Our analysis shows that the most frequent activity was port-scanning to look for vulnerable hosts, whereas the heaviest traffic was associated with SPAM-related activities, according to Mi et al.'s work [1].
- Our analysis reveals that the RESIP hosts are distributed widely across all regions in Japan. The statistics for RESIP hosts show that hosts are mainly associated with residential and mobile ISPs.

Table 1. RESIP service providers and basic specifications

RESIP Provider	Fee (2017)	Fee (2019)	IPs [1]
Proxies Online (United States)	$25/Gb	certificate expired	1,257,418
Geosurf (Netherlands)	$300/month	$450–2000/month	432,975
ProxyRack (United States)	$40/month	$60–120/month	857,178
Luminati (United States)	$500/month	$12.5/GB+$500/month	4,033,418
IAPS Security	$500/month	site unavailable	

2 Residential IP Proxy

Residential IP proxy services are a new business. The RESIP providers control a large number of residential hosts to proxy their customers' communication with any destination on the Internet.

Figure 1 illustrates how the RESIP service model works. Three parties are involved here, namely, the RESIP client, the Proxy gateway and the Residential hosts. Once a client signs up with a RESIP service, it receives a gateway's IP address or URL for the service. The gateway forwards the client's requests to one of residential hosts, which sends the request to the target hosts that the client wishes to visit. The responses are sent back to the client via the same routing arrangements. The forwarding proxies are assigned randomly and are periodically updated to confound analysis of traffic.

According to the study [1], the followings were discovered from their crawling, and analysis.

- A total of 6,183,876 unique RESIP addresses were collected. Their classifier estimated that 95.22% of RESIPs were residential addresses and that 237,029 addresses (43.2%) were assigned to IoT devices.
- RESIP service providers claimed that their proxies were all common users who willingly join their network. However, none of the five major providers operated a completely consent-based proxy system.
- The new RESIP service became a booming business. Table 1 shows that most providers have increased their service fees in the two years from 2017 [1] to our work (2019). On the other hand, some providers have already abandoned the business.

3 Investigation Methodology

3.1 Datasets

Table 2 lists the four databases examined in this study.

rpaas dataset This comprises records containing of the detected RESIP address, and the duration of its activities for the five major RESIP providers:

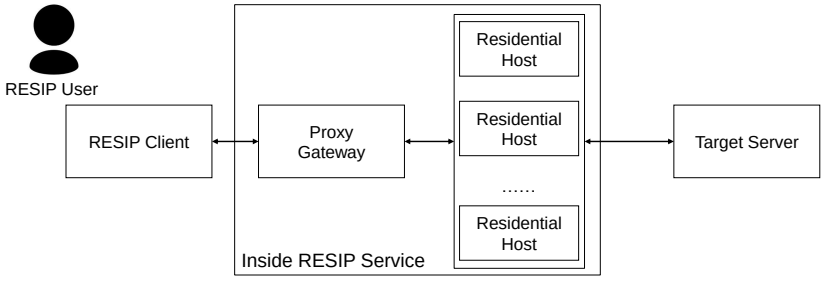

Fig. 1. RESIP service overview

Proxies Online (PO)[1] Geosurf (GO)[2], ProxyRack (PR)[3], Luminati (LU)[4] IAPS Security (IS)[5]The dataset of RESIP addresses and the source code of the profiling tools used are available at [2].

NICTER Darknet dataset NICT provides the source IP addresses sent to the NICT darknet of /20 block. Their analysis infrastructure, NONSTOP, provides the remote access to the attributes stored in packet headers, including capturing time, source and destination of the address and port, and the countries involved.

GeoLite2 City dataset This is a geolocation database provided by MaxMind Inc. The attribute information includes countries, region, latitude and longitude.

APNIC whois dataset APNIC is one of the five Regional Internet Registries (RIRs) offering a Whois directory service to resources of IP addresses and domain names, and Autonomous System number (ASN). These information are provided in JSON format object from Registration Data Access Protocol (RDAP) [7].

[1] Proxies Online. `http://proxies.online`.
[2] Geosurf: Residential and data center proxy network. `https://www.geosurf.com/`.
[3] Proxyrack. `https://proxyrack`.
[4] Luminati: largest business proxy service.
[5] IAPS security.

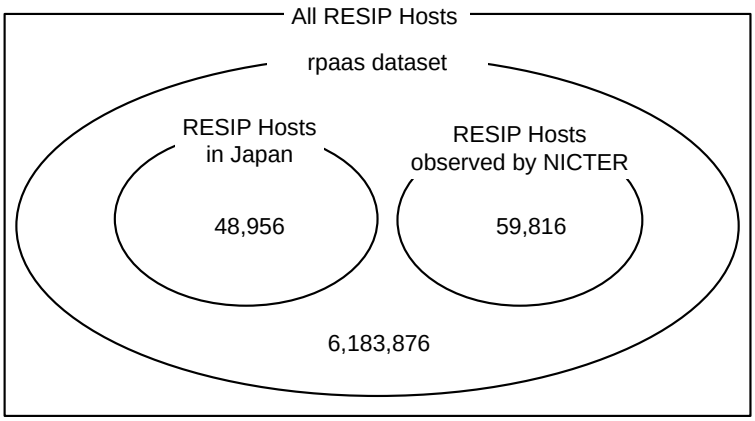

Fig. 2. Relationships among some subsets of RESIP addresses

3.2 Attributes of RESIP hosts

To investigate the attributes of RESIP hosts, we focus on those RESIP addresses that are under the management of Japanese organizations for which we know the region, name of organization, and address blocks used. The steps were as follows.

1. Lookup GeoLite2 city dataset for RESIP addresses to identify the addresses belonging to Japanese regional networks (JP). Estimate the prefecture names for the addresses.
2. Perform `nslookup` query to the extracted address to find the domain information.
3. Use RDAP query to obtain the CIDR block information and the registration organization.

3.3 Suspicious Traffic from RESIP hosts

Assume that any host whose source address has been captured in the NICT darknet is performing port scans to look for new vulnerable hosts. We use the NICT NONSTOP service on the first and last days for which a RESIP address has been detected. We examine if the target RESIP address has been observed. Of so, we identify the corresponding port numbers that indicate the type of service the host is interested in.

Table 2. Resources in this study

	rpaas dataset	NICTER Darknet dataset	GeoLite2 City dataset	APNIC whois dataset
Year	2017	2017	2019	2019
Details	The list of IP addresses participating RESIP service collected in [1]	Source and destination data of packets observed on /20 darknet by the National Institute of Information and Communications Technology	IP address and geographic information database provided by MaxMind, Inc.	IP address and domain database operated by APNIC registry
Records	6,183,876	About 150 billion		
Usage	Published	Access from NICTER NONSTOP	Datebase access from Python	RDAP request from Python

4 Results

4.1 Attributes of RESIP hosts in Japan

Fig. 2 illustrates the relationships between address subsets, RPaas datasets, and the target addresses; in a Venn diagram. Among the RESIP addresses (RPaas dataset), we found 48,956 IP addresses managed by Japanese organizations.

Table 3 lists the top 10 prefectures (states) as well as the numbers of RESIP addresses with regard to RESIP providers. Tokyo is the greatest in the number of RESIP addresses. The most common RESIP provider in Japan is ProxyRack (18,502 addresses).

Table 4 and 5 shows the top ten domains (with third level) and the ISPs, respectively. The biggest RESIP owner was NTT Communication Corp. , which is known as the largest IPS under which the greatest RESIP domain `ocn.ne.jp` is management of.

Table 6 shows the numbers of RESIP addresses classified by the type of network. Following the domestic convention in Japan, the second level of a domain indicates the characteristics of the network, e.g., "ne" (**ne**twork service), "or" (**or**ganization), "ad" (**ad**ministrative) and so on. Table shows that the "ne" domain (usually used for residential networks) has the greatest number of RESIP addresses in Japan.

Note that 91 addresses are for "ac" (academic network, such as universities), nine are for "co" (companies), and one is for "go" (government). Obviously, these addresses are not residential and have not yet detected via Mi et al.'s analysis [1].

We should comment on the accuracy of the datasets. First, the estimated country is not always consistent. For example, 43 domains with a `.ru` top-level domain were estimated with Tokyo in the GeoLite2 City database. The undetermined domain (`pinspb.ru`) has some webpages written in Russian and was classified as Russian in [10] but Israel in [11].

Table 3. List of top 10 prefectures for RESIP hosts with service providers. PO: Proxies Online, GS: Geosurf, PR: ProxyRack, LU: Luminati, IS: IAPS Security

Prefecture	RESIPs	%	PO	GS	PR	LU	IS	Fraction of mobile phone and PHS users(%) [8]
Tokyo	12,766	26.1	2,709	84	4,442	5,027	4	26.0
Kanagawa	3,094	6.3	721	17	1,145	1,087	0	6.4
Aichi	2,940	6.0	715	15	1,163	942	0	5.2
Osaka	2,917	5.9	769	17	1,148	880	1	6.7
Saitama	2,544	5.1	605	14	1,082	754	0	4.7
Tiba	1,912	3.9	484	32	726	557	0	4.0
Hyogo	1,722	3.5	460	21	693	493	0	3.5
Hukuoka	1,266	2.5	426	9	436	320	0	4.0
Sizuoka	1,083	2.2	251	7	484	308	0	2.2
not found	6,619	13.5	1,741	52	2,108	2,507	8	
Total	48,956	100	11,918	304	18,502	16,325	13	100

Table 4. List of TOP 10 TLD+2 domains for RESIP hosts

TLD+2	RESIPs	%
ocn.ne.jp	7,468	15.2
au-net.ne.jp	5,616	11.4
plala.or.jp	2,900	5.9
dion.ne.jp	2,528	5.1
not found	2,441	4.9
so-net.ne.jp	1,966	4.0
mesh.ad.jp	1,935	3.9
eonet.ne.jp	1,305	2.6
home.ne.jp	1,209	2.4
nttpc.ne.jp	1,116	2.2
Total	48,956	100

4.2 Traffic from RESIPs

Fig. 4 shows the daily numbers of packets observed in the NICTER darknet. There were a total 1,683,440 packets sent from 59,816 RESIP addresses. The results show that the durations detected in Mi et al.'s analysis [1] has the intersection with the NICTER datasets.

Table 7 lists the top 10 RESIP addresses in terms of the cumulative observed packets. Note that the very busy activities (62,669 scans) were performed by only a few RESIP hosts. The durations of port-scanning from these 10 hosts are plotted in Fig. 5, where the scans are indicated at the IP addresses along the Y-axis.

Table 8 shows the top 10 destination port numbers specified by RESIP hosts. The corresponding services are given in the table. For example, the Telnet service

Table 5. Top 10 domains for RESIP hosts

Organization	Domains	RESIPs	%	Share of FTTH users (%) [9]
NTT Communication Corporation	ocn.ne.jp, plala.or.jp	10,941	22.3	34.2
KDDI CORPORATION	au-net.ne.jp, dion.ne.jp	8,301	16.9	12.8
Japan Nation-wide Network of Softbank Corp.	bbtec.net, access-internet.ne.jp	7,781	15.8	
Japan Network Information Center	nttpc.ne.jp, mesh.ad.jp	4,756	9.7	
Sony Network Communications Inc.	so-net.ne.jp, ap.nuro.jp	2,544	5.1	
OPTAGE Inc.	eonet.ne.jp	1,274	2.6	5.4
BIGLOBE Inc.	mesh.ad.jp	1,230	2.5	
Jupiter Telecommunication Co.,Ltd	home.ne.jp	1,209	2.4	
Chubu Telecommunications Co.,Inc.	commufa.jp	1,125	2.2	
ARTERIA Networks Corporation	ucom.ne.jp, vectant.ne.jp	965	1.9	2.3
Total		48,956	100	

designated for the well-known port number 23 was observed in 613,606 packets, which accounts for 36.4% of the total.

Any possible relationship between the designated port number and the duration of the scan would be brought out by the scatterplot of Fig. 6. Note that no significant correlation between the target of the service and its duration can be seen. However, major services MSSQL and SMTP are constantly observed.

4.3 Discussion

Let us remark each of questions.

Table 6. Counts of RESIP hosts for the various network types (second-level domains)

2LD	RESIPs	%
ne	28,824	74.1
or	4,340	11.1
ad	2,208	5.6
ac	91	0.2
co	9	
go	1	
gr	1	
ed	1	
Total(.jp)	38,946	100

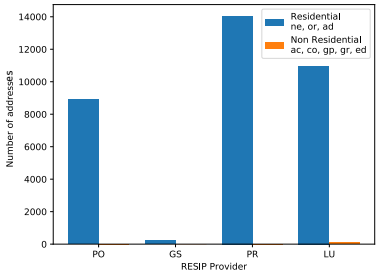

Fig. 3. Number of RESIP addresses for network types (second-level domains)

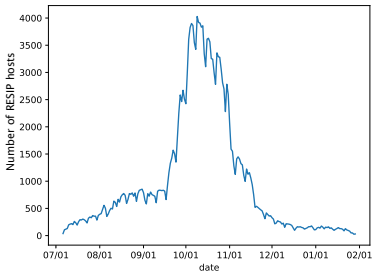

Fig. 4. Daily counts of RESIP hosts observed in NICTER darknet

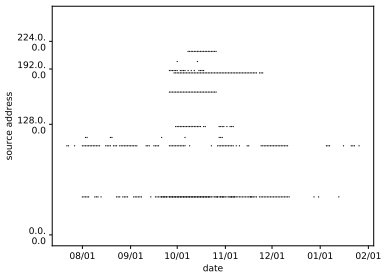

Fig. 5. Active durations for RESIP source addresses

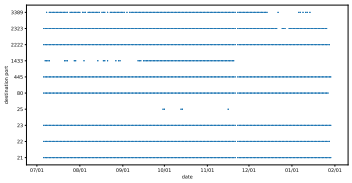

Fig. 6. Active durations for destination port numbers

1) For the various kinds of networks, we found that 90.8% of the RESIP hosts could be classified as residential, based on Table 6 and Fig. 3. The subdomains “ne”, “ad”, and “or” were the most used in RESIP proxies. According to the domain name convention, these are known to be residential. Note that some exceptions “ac” and “co” domains, assigned for academic and company business, were also found. We consider that, for mobile laptop computers with a RESIP library installed, the installation was without consent of their owners and was being operated for malicious purposes.

2), 3) Tables 3, 4 and 5 confirm that the RESIPs are distributed widely in all prefectures (regions) and that the distribution matches the statistics for cell phone users. This implies that residential and mobile ISPs are the main RESIP hosts in Japan, which differs with the earlier observation [1] that most RESIP devices (69.8%) could be identified as routers, firewalls, or WAP devices. Table 5 shows no skew in the relationship between RESIP hosts and the number of ISPs.

4) Tables 7, Fig. 4 and 5 demonstrate that constant port-scanning was performed from RESIP hosts. In contrast to the report [12], there are now many

Table 7. List of top 10 RESIP addresses for the frequency of observations in darknet

Address	Days	RESIP provider	# Packets
43.249.57.255	8	ProxyRack	62,669
187.120.17.2	34	Proxies Online Geosurf	35,353
200.170.223.50	7	Luminati	21,676
103.29.97.2	8	Proxies Online Geosurf Luminati	17,004
165.73.122.29	14	Luminati	16,127
212.90.62.209	5	Luminati	15,142
43.248.73.6	90	Proxies Online Geosurf Luminati	13,425
190.57.236.230	18	Luminati	13,388
112.196.77.202	27	Proxies Online Geosurf	13,061
125.99.100.22	10	Proxies Online Luminati	12,952

Table 8. List of top 10 destination port numbers in frequencies

Destination port	Service	# Packets	%
23	Telnet	613,606	36.4
445	SMB	399,250	23.7
21	FTP	193,917	11.5
1433	MSSQL	144,928	8.6
80	HTTP	97,780	5.8
22	SSH	49,767	2.9
2323	(Telnet)	43,310	2.5
25	SMTP	21,732	1.3
2222	(SSH)	16,838	0.1
3389	RDP	9,782	0.5

cyberattacks form identified RESIP hosts. Therefore, we can infer that the threat from RESIP service is becoming more serious.

5) Table 8 shows that the major RESIP activities were related to port-scanning. This observation is not consistent with the result from Mi et al.'s work [1], which claimed that the most frequent activity was ad mail (SPAM) at 36.55%. Our analysis shows that the SPAM traffic accounts for only 1.3 % of activity and that its duration is limited, as shown in Fig. 6.
   This may be a feature of Japanese networks, where ad messages are shifting from email to SNSs. Another possible reason might be limitations in the observation. Our estimations were based on the darknet, which carries only a small fraction of the Internet traffic. We need additional investigations to be able to distinguish clearly between the objectives of RESIP hosts.

5 Conclusions

We have studied RESIP host activities detected from networks under the control of organizations in Japan, which accounts for 0.79% of the all Internet RESIP hosts. Our analysis of 1,683,550 RESIP packets observed from the darknet revealed that 90.8% RESIP were residential and the RESIP proxies were distributed evenly across all prefectures and IPSs. New finding is that most of devices that became RESIP hosts in Japan were mobile, whereas routers, firewalls and WAP devices were identified from the profiles in the previous study [1]. Another distinct aspect of the RESIP behavior is the distribution of malicious activities. In [1], the SPAM and malicious website hosting were the most common (36.5% and 32.7%, respectively), whereas the SPAM traffic accounted for only 1.3% of all traffic in our analysis. We found that port-scanning was the most

frequent malicious activity. Despite these evolving trends, we conclude that more RESIP hosts are becoming involved in serious threat and we need countermeasures aimed at minimizing the abuse of RESIP hosts.

References

1. X. Mi, X. Feng, X. Liao, B. Liu, X. Wang, F. Qian, Z. Li, S. Alrwais, L. Sun, and Y. Liu, "Resident Evil: Understanding Residential IP Proxy as a Dark Service", IEEE Symposium on Security and Privacy (SP), vol. 1, pp. 170–186, 2019.
2. RPaaS: Characterizing Residential IP Proxy as a Service. `https://rpaas.site/`.
3. MAXMIND: GeoLite2 Free Downloadable Databases. https://dev.maxmind.com/geoip/geoip2/geolite2/.
4. APNIC: Whois search. https://www.apnic.net/about-apnic/whois_search/.
5. D. Inoue, M. Eto, K. Yoshioka, S. Baba, K. Suzuki, J. Nakazato, K. Ohtaka, and K. Nakao, "nicter: An Incident Analysis System Toward Binding Network Monitoring with Malware Analysis," In Information Security Threats, Data Collection, and Sharing, WISTDCS'08. WOMBAT Workshop on, IEEE, pp. 58–66, 2008.
6. T Takehisa, M Kamizono, T Kasama, J Nakazato, M Eto, D Inoue, K Nakao, Utilization of Secure Remote Analysis Platform for Cybersecurity Information (NONSTOP), Computer Security Symposium2014, volume: 2, pp. 207-214, 2014. (in Japanese)
7. APNIC: Registration Data Access Protocol. https://www.apnic.net/about-apnic/whois_search/about/rdap/.
8. Telecom data book 2018(Compiled by TCA)
9. Statistics on Internet traffic in Japan, MIC, http://www.soumu.go.jp/main_content/000523384.pdf.
10. IPInfoDB. https://ipinfodb.com/.
11. RIPE NCC: whois Database. https://www.ripe.net/manage-ips-and-asns/db/support/documentation/glossary/whois-database.
12. National Institute of Information and Communications Technology NICTER Observation Report 2017 (in Japanese)https://www.nict.go.jp/cyber/report/NICTER_report_2017.pdf.
13. T. Chung, D. Choffnes, and A. Mislove, "Tunneling for transparency: A large-scale analysis of end-to-end violations in the internet", In Proceedings of the 2016 ACM on Internet Measurement Conference, pp 199–213. ACM, 2016.
14. N. Weaver, C. Kreibich, M. Dam, and V. Paxson, "Here be web proxies", In International Conference on Passive and Active Network Measurement, pp 183–192. Springer, 2014.

A 47 prefectures for RESIP

Table 9 shows the number of RESIP addresses for each of 47 prefecture of Japan, with numbers for major five service providers.

Table 9. List of 47 prefectures for RESIP hosts with service providers. PO: Proxies Online, GS: Geosurf, PR: ProxyRack, LU: Luminati, IS: IAPS Security

Prefecture	RESIPs	PO	GS	PR	LU	IS
Tokyo	12,766	2,709	84	4,442	5,027	4
Kanagawa	3,094	721	17	1,145	1,087	0
Aichi	2,940	715	15	1,163	942	0
Osaka	2,917	769	17	1,148	880	1
Saitama	2,544	605	14	1,082	754	0
Chiba	1,912	484	32	726	557	0
Hyogo	1,722	460	21	693	493	0
Hukuoka	1,266	426	9	436	320	0
Sizuoka	1,083	251	7	484	308	0
Hokkaido	1,061	324	9	448	225	0
Kyoto	997	213	0	438	310	0
Mie	638	115	1	300	208	0
Hiroshima	589	168	2	257	138	0
Gifu	584	118	1	299	139	0
Ibaragi	568	107	1	264	179	0
Okinawa	543	89	3	153	284	0
Tochigi	473	125	1	186	134	0
Gunma	432	112	1	144	158	0
Nagano	418	80	0	172	144	0
Niigata	409	95	0	200	100	0
Shiga	380	99	1	131	135	0
Miyagi	372	104	5	150	97	0
Okayama	316	89	0	129	97	0
Nara	302	74	1	121	85	0
Kumamoto	297	98	1	108	82	0
Ehime	271	94	0	97	68	0
Yamaguchi	242	79	0	97	57	0
Fukushima	241	72	0	113	42	0
Kagawa	227	60	2	128	27	0
Toyama	216	57	0	92	56	0
Ishikawa	210	61	0	65	73	0
Yamanashi	201	54	0	81	62	0
Oita	186	48	1	77	52	0
Wakayama	177	51	0	88	34	0
Aomori	168	34	0	85	46	0
Fukui	159	36	1	57	61	0
Kagoshima	157	41	1	72	38	0
Kouchi	154	53	0	67	27	0
Yamagata	148	31	1	68	38	0
Iwate	139	36	0	61	32	0
Akita	131	36	0	60	31	0
Nagasaki	128	25	0	49	52	0
Saga	126	38	0	54	28	0
Tokushima	125	37	0	46	33	0
Miyazaki	124	33	0	51	36	0
Tottori	94	38	0	37	14	0
Shimane	90	13	0	30	43	0
not found	6,619	1,741	52	2,108	2,507	8
Total	48,956	11,918	304	18,502	16,325	13