SOC Report

1.1 Introduction

IIJ announced its new security business brand wizSafe[*1] on October 31, 2016 and is constantly working to bring about a world in which customers can use the Internet safely. As part of such efforts, we release up-to-date information on security threats observed at our SOC in blog format via wizSafe Security Signal[*2]. This includes some information on threats identified through IIJ's Data Analytics Platform. For an overview of the Data Analytics Platform, see Internet Infrastructure Review (IIR) Vol. 38[*3].

Here, we give an overview of analysis using the Data Analytics Platform. The logs collected on the platform naturally include those from security devices such as firewalls, IPS/IDS, and antivirus solutions provided as IIJ services, as well as logs of DNS queries, Web access, incoming/outgoing email, and so forth. Characteristically, these logs contain only a tiny amount of abnormal traffic (threats) among a large amount of normal traffic. We therefore need to think about how to go about aggregating and visualizing the data so that we can identify threats clearly.

Section 1.2 describes information on threats revealed via the Data Analytics Platform in 2018, and Section 1.3 describes new initiatives using the Data Analytics Platform. The observations for 2018 are summarized in wizSafe Security Signal[*4].

1.2 Observational Data

First, we look at activity identified using the Data Analytics Platform that is particularly noteworthy. This information is taken from wizSafe Security Signal posts from last year.

1.2.1 Attacks Involving Cryptocurrencies

Attempts to monetize attacks using cryptocurrencies attracted attention in 2018. Analysis on IIJ's Data Analytics Platform also revealed several cases of attackers attempting to exploit cryptocurrencies.

The first example involves manipulating a website to embed a mining script. Our SOC's observations uncovered multiple cases of mining scripts embedded in websites that do not appear to have been put there intentionally by the website

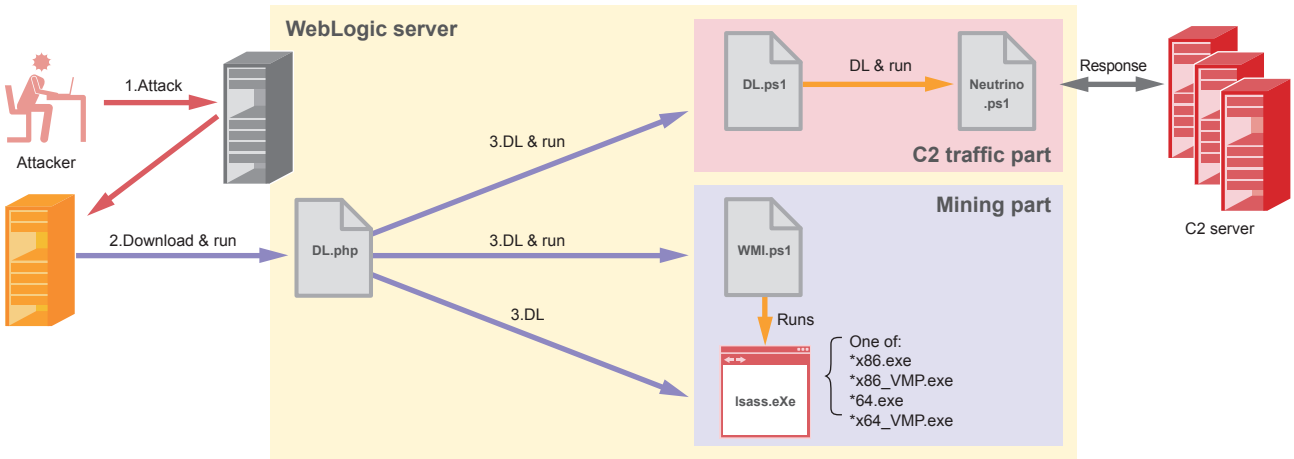

Figure 1: Overview of GhostMiner Attack

*1 IIJ announces new security business brand wizSafe (https://www.iij.ad.jp/en/news/pressrelease/2016/1031.html).
*2 wizSafe (in Japanese at https://wizsafe.iij.ad.jp).
*3 IIJ, Internet Infrastructure Review (IIR) Vol. 38 (https://www.iij.ad.jp/en/dev/iir/038.html).
*4 wizSafe, "wizSafe Security Signal 2018 Annual Summary" (in Japanese at https://wizsafe.iij.ad.jp/2019/03/601/).

administrator. By exploiting website vulnerabilities and so forth, attackers can embed mining scripts into Web pages. When a user views a tainted site, the user's computer runs the cryptocurrency mining script, and any mined proceeds go to the attacker.

The above is an example of an attack aimed at clients, but we have also observed cryptocurrency mining attacks on servers. One specific example is an attack campaign[*5] called GhostMiner (Figure 1). The GhostMiner campaign was observed in March 2018 and exploits a vulnerability (CVE-2017-10271) in Oracle WebLogic Server. The vulnerability allows the execution of remote code, so the attacker ultimately attempts to get the Web server to mine cryptocurrency. We have also observed several other attempts to use remote code execution vulnerabilities to get servers to mine cryptocurrency[*6][*7].

Yet another example involves not mining but attempts to illegally transfer funds. In December 2018, we observed scanning activity (Figure 2) targeting the JSON-RPC protocol used in an Ethereum client[*8]. The scanning activity was looking for Ethereum clients that are accessible via the Internet due to a misconfiguration. We note that a number of conditions must be met for the funds transfer to actually complete successfully.

Cryptocurrencies are appealing to attackers because attacks on them can be monetized directly and because, depending on the type, they offer a high degree of anonymity. Also, any environment that has computational resources can be attacked, as evidenced by the variety of attacks geared to cryptocurrency mining, which target both clients and servers. We expect attackers to continue to target cryptocurrencies as one means of monetizing attacks.

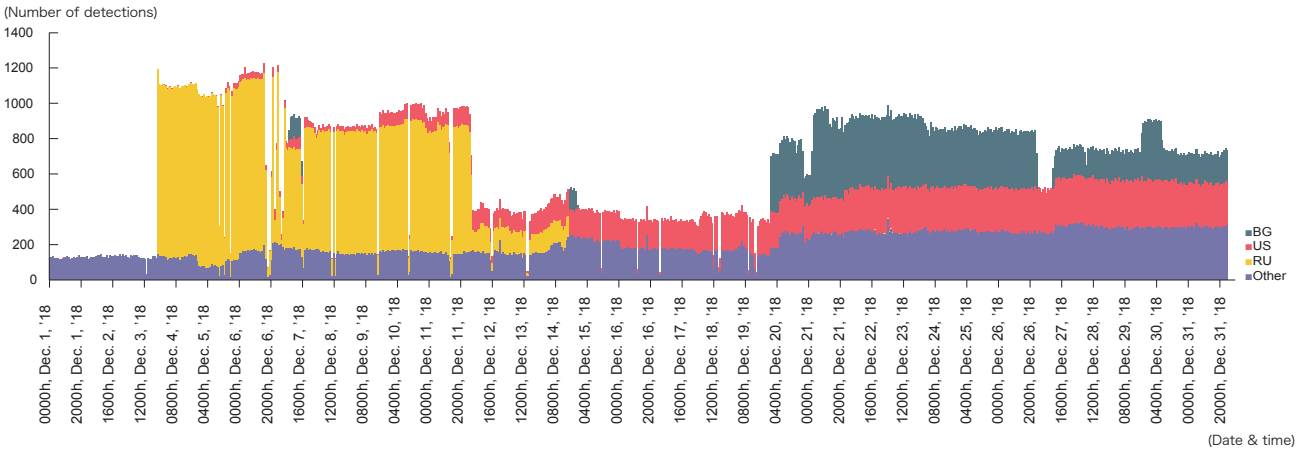

Figure 2: Scanning of 8545/TCP (Dec. 2018)

*5 wizSafe, "GhostMiner infections spreading" (in Japanese at https://wizsafe.iij.ad.jp/2018/04/323/).
*6 wizSafe, "wizSafe Security Signal January 2018 Observation Report" (in Japanese at https://wizsafe.iij.ad.jp/2018/02/247/).
*7 wizSafe, "wizSafe Security Signal February 2018 Observation Report" (in Japanese at https://wizsafe.iij.ad.jp/2018/03/286/).
*8 wizSafe, "Ethereum JSON-RPC scans observed" (in Japanese at https://wizsafe.iij.ad.jp/2019/01/541/).

1.2.2 SYN/ACK Reflection Attack

One peculiar example of a DDoS attack that the SOC observed in 2018 is a SYN/ACK reflection attack using 80/TCP. This was included in wizSafe Security Signal for September 2018[*9] (Figure 3). The attack sends TCP SYN packets with a spoofed source address to many addresses simultaneously, thereby effectively recruiting the resulting SYN/ACK packet responses to perform a DDoS attack on the source address.

This SYN/ACK reflection attack was observed by the SOC on September 26, 2018, but it has also been observed on a small scale since October, and attacks of the same type are detected daily via the Data Analytics Platform. One feature of the DDoS attack observed on September 26 is that the source invokes the attack by sending a small amount of SYN packets to servers on which the 80/TCP port is open to the Internet. If an attacker sends a high volume of SYN packets to a single server, the administrator of the receiving server is liable to think that a TCP SYN flood attack[*10] is underway and block further traffic. If this happens, the attacker may not be able to realize an attack of the scale envisaged. Also, because only a small volume of SYN packets is received per server, one can infer that the attacker is probably sending out SYN packets far and wide.

The attack described above uses port 80/TCP, and servers on which 80/TCP is open can generally be considered to be Web servers. Hence, normal Web access traffic on such servers do contain a small amount of SYN packets, the type of packet used in a SYN/ACK reflection attack, and it is thus difficult to determine whether any of those packets are being used in an attack. In this example, we detected the attack by cross-analyzing the multiple customer firewall logs present on our Data Analytics Platform.

Because firewall logs reveal information about internal and external access, such attacks can be detected when multiple firewall logs indicate that 80/TCP responses are being generated for a single specific IP address (the spoofed IP address being attacked). However, this feature could indicate scanning activity rather than a DDoS attack. We therefore

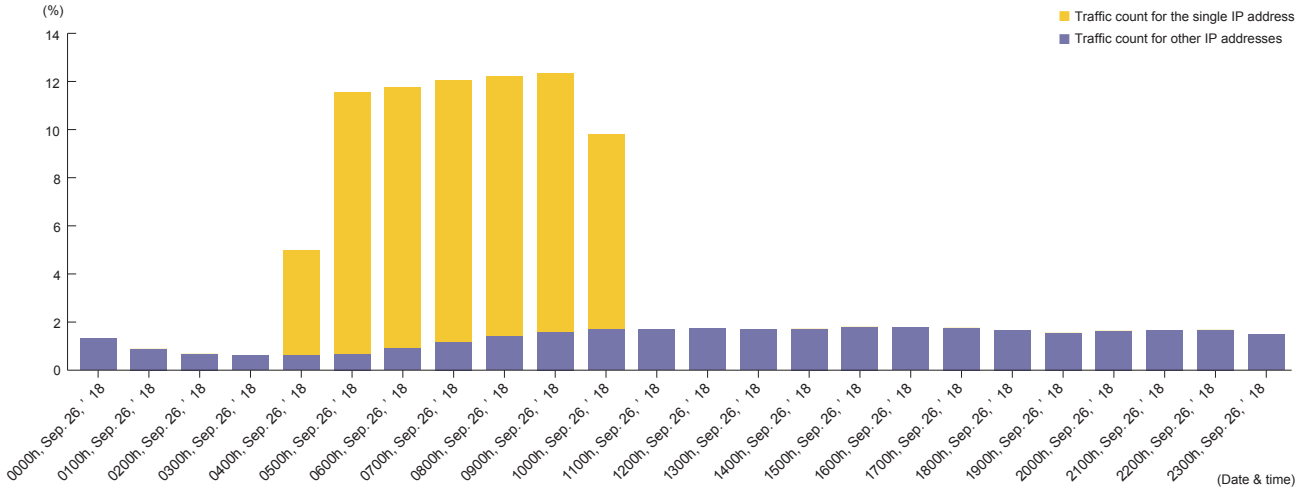

Figure 3: Increase in 80/TCP Traffic from a Single IP Address

*9 wizSafe, "wizSafe Security Signal September 2018 Observation Report" (in Japanese at https://wizsafe.iij.ad.jp/2018/10/470/).

*10 In a TCP SYN flood attack, the attacker sends a large amount of SYN packets—requests used to establish a TCP connection—to the target system, causing it to prepare for a large number of connections and thereby wasting processing power, memory, etc.

differentiate between DDoS attacks and scanning activity based on total bytes sent/received, duration, and so on as calculated from the firewall logs.

The SYN/ACK reflection attack that we wrote about in September 2018 was also observed in IIJ's honeypots. We reported about this in detail in an IIJ-SECT blog post (in Japanese): "SYN/ACK reflection attack using IoT devices as a springboard"[*11]. The post describes changes in the ports used in the attack and reveals that it is a complex DDoS attack that uses the UDP protocol, so we encourage you to read through it.

1.2.3 Resurgence of Attacks Targeting Known Vulnerabilities

One notable of the 2018 analysis performed on the Data Analytics Platform is that some attacks targeting vulnerabilities that have already been disclosed and for which patches have been made available have re-emerged after being dormant for some time. One example is malware that exploits a vulnerability (CVE-2017-11822) in the Microsoft Office Equation Editor.

The vulnerability that the malware exploits is a buffer overflow issue with the Microsoft Office Equation Editor that allows remote code execution. Microsoft issued a patch that fixes this vulnerability in November 2017. As a workaround, users can also disable the Equation Editor as a means of avoiding this attack without applying the patch.

We observed an attack targeting this vulnerability via the Data Analytics Platform in September 2018, almost a year after the fix had been issued (Figure 4)[*12]. The attackers sent malware that exploits the vulnerability as an email attachment. We think this was a deliberate attempt to target systems in which a fix was never applied or in which the Equation Editor had only temporarily been disabled as a means of avoiding the attack, and that it was timed for when awareness of this vulnerability had faded somewhat.

In addition to the Microsoft Office vulnerability discussed here, we have observed multiple other similar cases via the Data Analytics Platform[*13]. A lesson to be learned from these observations is that whether one implements a fundamental

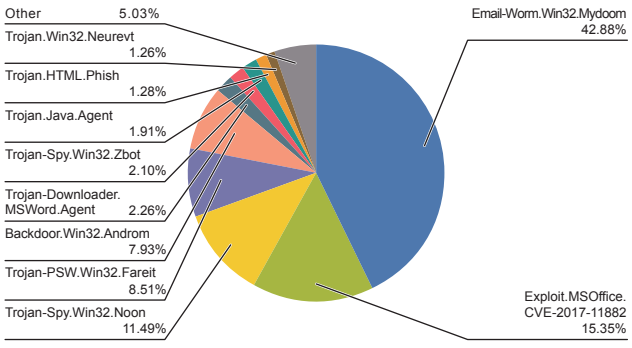

Figure 4: Breakdown of Malware Types Detected in Received Emails (Sep. 2018)

*11 IIJ-SECT Security Diary, "SYN/ACK reflection attack using IoT devices as a springboard" (in Japanese at https://sect.iij.ad.jp/d/2019/02/128021.html).
*12 wizSafe, "wizSafe Security Signal September 2018 Observation Report" (in Japanese at https://wizsafe.iij.ad.jp/2018/10/470/).
*13 wizSafe, "wizSafe Security Signal November 2017 Observation Report" (in Japanese at https://wizsafe.iij.ad.jp/2017/12/184/).

fix to a vulnerability or, depending on the circumstances, any of the various workarounds available, it is crucial that such measures are kept in place indefinitely.

1.3 Detecting Malicious Transmissions Using Machine Learning

Characteristically, the data analyzed on our Data Analytics Platform include only a tiny amount of abnormal traffic (threats) among a large amount of normal traffic. Efforts are being made to use machine learning to discover such threats. The main task handled is that of detecting anomalies from imbalanced data. Here, we describe two such projects that are underway, along with the challenges they face.

1.3.1 Application to DNS Query Data

The domain names of the C2 (command & control) servers used by malware may be generated algorithmically using a DGA (domain generation algorithm). The domain names generated by DGAs differ widely depending on the type of algorithm and the parameters used when running it. This can make it difficult to blacklist the malware's servers ahead of time or to create an expression for the detection signature.

So in this project, we aim to solve the problem by combining the Data Analytics Platform's DNS query data with machine learning. We take this approach because tasks that humans find difficult to construct rules for can be amenable to machine-learning solutions. A desirable property of machine learning algorithms is that they can autonomously acquire the ability to classify anomalies when provided with data containing features that are effective in identifying those anomalies. For example, in IIR Vol. 41, we looked at URL strings and described an approach to identifying rogue sites using neural networks[*14].

We know of several attempts to use machine learning to detect DGAs, including some that have already been put into real-world use. Currently, we are engaged in research that follows on from the FANCI (Feature-based Automated NXDomain Classification and Intelligence)[*15] system announced at USENIX Security '18. As the conference paper on FANCI explains, the system combines domain features inspired by those used in past research with a machine learning algorithm known as random forests, and it generalizes very well.

As the first step in our follow-up research, we aim to stick to the methodology described in the paper as much as possible and use the DNS query data available from our Data Analytics Platform. This first step is intended to assess whether the methodology can be applied unmodified to the Data Analytics Platform's data. We do this because any given methodology will not necessarily produce the same results when different data are used. Next, if we determine that the methodology cannot be applied as is, and that it is possible to investigate why and implement a solution, we intend to work toward a practical implementation that may include additional performance enhancements. Potential performance enhancements could, for example, come from the use of gradient boosting decision trees, a popular method in recent years, or ensemble learning that combines undersampling and bagging.

The volume of data passing through and processed by the Data Analytics Platform is large, however, so as a matter of practicality, the model needs to have high throughput. We aim to strike a balance between the increasing computational load that results from the use of more complicated models and workflows and the performance enhancements that can be obtained, and with some fine tuning, we ultimately aim to build the system into our Data Analytics Platform to provide one means of detecting these anomalies.

1.3.2 Application to Web Proxy Data

One other project we are pursuing aims to detect communications sent to C2 servers by malware in Web proxy data. We are currently running validation tests with the objective of applying the methodology presented by IIJ engineers at Black Hat Europe 2018[*16] to our Data Analytics Platform's logs. The methodology presented at Black

*14 IIJ, Internet Infrastructure Review (IIR) Vol. 41 (https://www.iij.ad.jp/en/dev/iir/041.html).

*15 USENIX, "FANCI: Feature-based Automated NXDomain Classification and Intelligence" (https://www.usenix.org/conference/usenixsecurity18/ presentation/schuppen).

*16 Black Hat, "Deep Impact: Recognizing Unknown Malicious Activities from Zero Knowledge" (https://www.blackhat.com/eu-18/briefings/schedule/#deep-impact-recognizing-unknown-malicious-activities-from-zero-knowledge-12276).

Hat Europe 2018 is described later in this edition under "Focused Research (1): Deep-Learning Analysis of Logs to Detect Malicious Communications".

The project uses convolutional neural networks, which are commonly applied to image recognition tasks, to discern trends in normal traffic and anomalous traffic (with a C2 server). The key here is learning model performance and its evaluation.

If there were, say, a model capable of producing 95% accuracy or better, this would generally be regarded as good performance. But because the volume of logs collected on the Data Analytics Platform is enormous, 1% of this data is not the sort of volume that a human could process by eye. Even if false positives do arise, the model needs to provide a level of accuracy that is tolerable when put into operation. This, of course, is the case when only machine learning is used, and approaches that reduce false positives through non-machine-learning systematic processing are also conceivable.

Aside from accuracy, we also need to be aware of differences in the distributions of the datasets we are dealing with. It is quite possible that the distributions of datasets used by reportedly well-performing machine learning models presented at conferences, academic events, and the like are characteristically different from the dataset distributions encountered by our SOC, so follow-up research is needed.

In view of the above, the SOC takes the overall design and operation of systems that use machine learning models into consideration, conducting follow-up research and working to improve the accuracy of machine learnings models, and is focused on building systems that improve quality without putting any additional load on current security operations.

One attempt to improve accuracy entails feature engineering. There is a strong perception that feature engineering involves adding features expected to be effective on the basis of data analysis, but other approaches also exist. For example, various statistics can be calculated from existing features, combined with the data from which they were derived, and then used for learning and evaluation. We will also use various other methods to repeatedly add and evaluate features as we work to enhance model.

1.4 Conclusion

In this edition, we provided an overview of analysis using the Data Analytics Platform, went over some actual observations from 2018, and described our efforts with respect to machine learning. The final machine learning approach that we described has the potential to further expand the detectable range for threats where detection with conventional methods is difficult or subject to limitations. We are able to pursue these efforts entirely because we are able to use traffic logs received from customers on the Data Analytics Platform, subject to customer consent. Machine learning approaches require large volumes of data in particular, so it is fair to say that we are only able to pursue these efforts because we have access to these traffic logs via the Data Analytics Platform. We will continue to provide up-to-date information on threats through wizSafe Security Signal and the IIJ-SECT blog, and we will continue striving to bring about a world in which the Internet is even safer for customers to use.

Satoshi Kobayashi
Data Analyst, Security Operations Center, Security Business Department, Advanced Security Division, IIJ

Shun Morita
Data Analyst, Security Operations Center, Security Business Department, Advanced Security Division, IIJ