Cyberattack Countermeasure Technology to Support NTT's Security Business

Takeo Hariu, Daiki Chiba, Mitsuaki Akiyama, Takeshi Yagi, Yuhei Kawakoya, Yukio Nagafuchi, and Takaaki Koyama

Abstract

NTT Secure Platform Laboratories is researching and developing the world's most advanced technologies for countering cyberattacks to support NTT's security business. In this article, we introduce domain name analysis technology that can effectively detect and defend against malware infections, malware analysis technology to support managed detection and response services, and unified threat management solutions that use a resilient security engine and anti-malware blacklists.

Keywords: domain name analysis, MDR services, security orchestration

1. Domain name analysis technology

Domain names and the domain name system (DNS) are considered to be essential elements of today's Internet. A domain name is information written in the form "example.com" that is used for identifying the destination of communications when accessing a website or sending/receiving email. The DNS, meanwhile, is a mechanism for obtaining a mapping between a domain name and an Internet protocol (IP) address that is required for actual communications.

Unfortunately, in addition to this important use on the Internet, domain names and the DNS are also exploited for use as an infrastructure for mounting cyberattacks. For example, an attacker may generate new domain names daily to distribute malicious software (malware) or create a domain name similar to that of a legitimate service to deceive users through a phishing attack. An attacker may also manipulate a command-and-control server to issue malware-controlling instructions and use domain names and the DNS to mount cyberattacks such as a DDoS (distributed denial-of-service) attack, spam-mail distribution, or information theft.

NTT Secure Platform Laboratories has been researching and developing various technologies to create information related to malware infections (security intelligence), with the aim of contributing to the NTT Group's global security business [1]. These include various types of decoy systems (honeypots) for accurately and safely observing malware behavior at the time of an infection and malware dynamic analysis technology for analyzing malware behavior after an infection by having the system manipulate that malware. However, the ever-evolving nature of cyberattacks indicates that attackers have been devising measures to avoid such analysis technologies and countermeasures. As a result, attacks that cannot be identified by these technologies have emerged.

Consequently, to continue responding to cyberattacks as they become more advanced and sophisticated, NTT Secure Platform Laboratories has initiated the research and development (R&D) of attack analysis techniques that focus on the properties of malicious domain names exploited by an attacker as

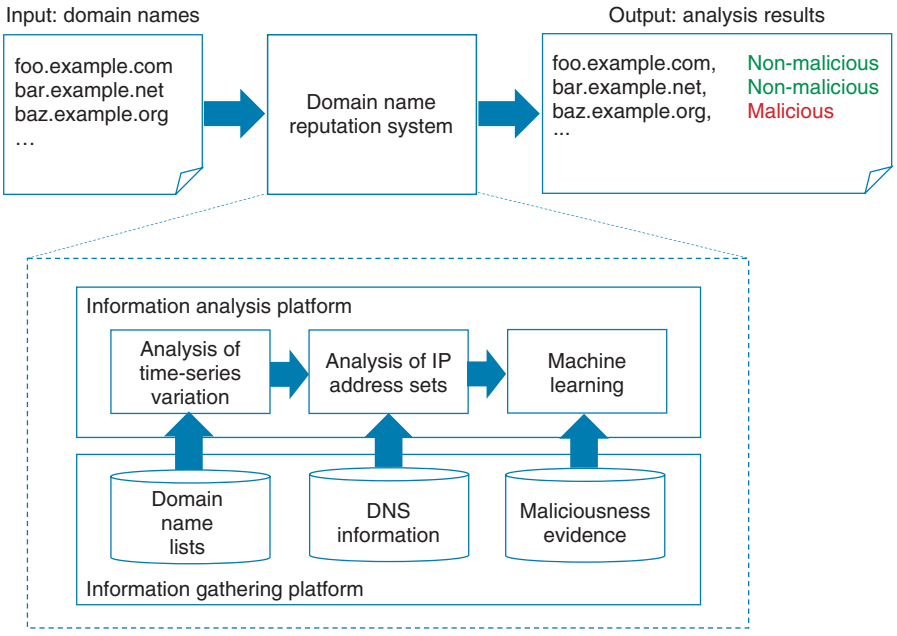


Fig. 1. Domain name reputation system.

an attack infrastructure, and has also begun acquiring more security intelligence. These R&D efforts include developing a domain name reputation system that can identify malicious domain names and a domain name categorization system that can generate information for effectively preventing cyberattacks based on malicious domain names. Cyberattack countermeasures including infection defense for preventing malware infections and identification of malware-infected hosts can be further improved using the security intelligence acquired by exploiting these two systems.

1.1 Domain name reputation system

At NTT Secure Platform Laboratories, we make use of information we collect as well as publicly available information to evaluate domain-name maliciousness from many angles. The domain name reputation system that we have developed enables us to identify and output malicious domain names used by attackers from input domain names (**Fig. 1**).

An attacker may generate new malicious domain names daily and may regularly change the domain name used in attacks to mount cyberattacks while evading countermeasures. For example, an attacker may attempt to evade countermeasures by generating a large number of malicious domain names that are valid for only a short period of time using a mechanism called a domain generation algorithm (DGA) or by reregistering legitimate domain names originally used for other purposes.

Amid such activity, this reputation system focuses on domain-name lifecycles from registration to expiration and analyzes domain-name characteristics that change due to cyberattacks as a time-series variation pattern. This approach enables accurate identification of domain names owned and used by an attacker. For example, we consider the case in which an attacker reregisters the domain name "example.com" originally used as a legitimate site as soon as it expires and then uses it as a malware distribution site. Our reputation system can detect the expiration of a domain name and subsequent changes in its use to identify a malicious domain name before the occurrence of an attack (**Fig. 2**).

In addition, this reputation system simultaneously performs analysis based on sets of IP addresses corresponding to a domain name. For example, given the input domain name "foo.example.com" and its parent domain name "example.com," this system examines sets of IP addresses corresponding to these domain names (**Fig. 3**). In particular, the use of security intelligence accumulated by NTT Secure Platform Laboratories makes it possible to refer to information on IP addresses used in past cyberattacks and to identify trends unique to attackers who operate malicious

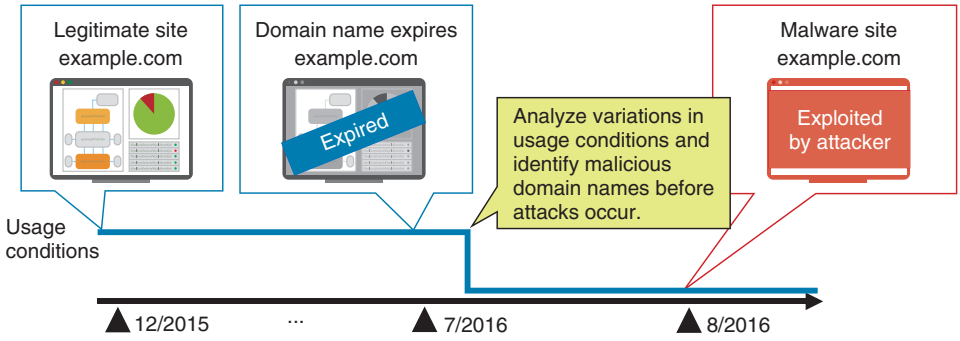


Fig. 2. Time-series variation analysis of domain name usage.

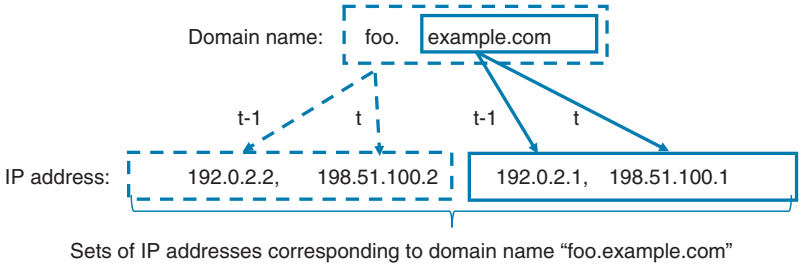


Fig. 3. Analysis of IP address sets corresponding to domain name.

domain names. Consequently, by using machine-learning techniques based on the results of analyzing IP address sets in this way and the results of analyzing time-series variation patterns as described above, we have achieved a system for calculating and predicting the possibility that a domain name is being exploited for malicious purposes.

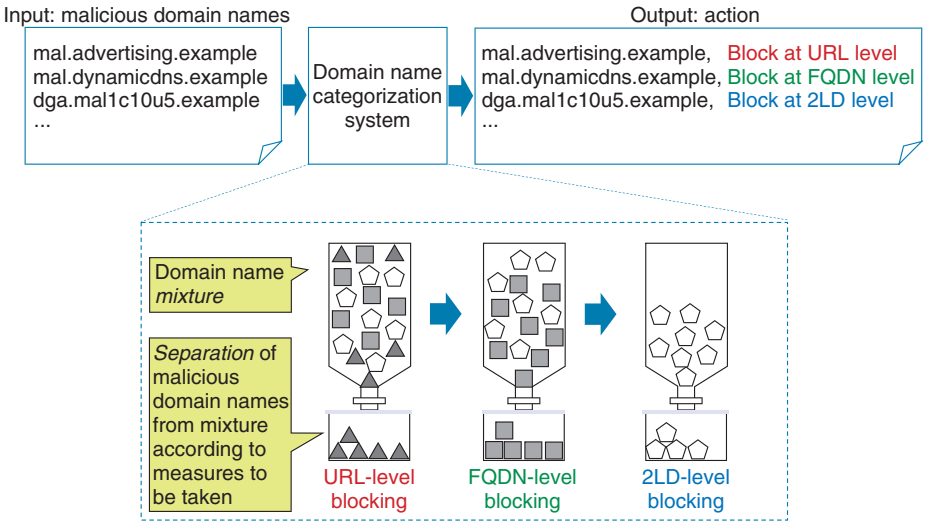
As a result of performing a large-scale evaluation using malicious domain names used in actual cyberattacks, this reputation system has been successful in predicting with good accuracy malicious domain names that have not been able to be identified using conventional techniques. The paper describing this system was presented at a prestigious international conference [2].

1.2 Domain name categorization system

This categorization system determines the history and circumstances behind domain-name generation and indicates the countermeasure that should be taken against individual malicious domain names. The domain name categorization system that we have achieved outputs the specific type of countermeasure that should be taken against each input malicious domain name to prevent cyberattacks (**Fig. 4**).

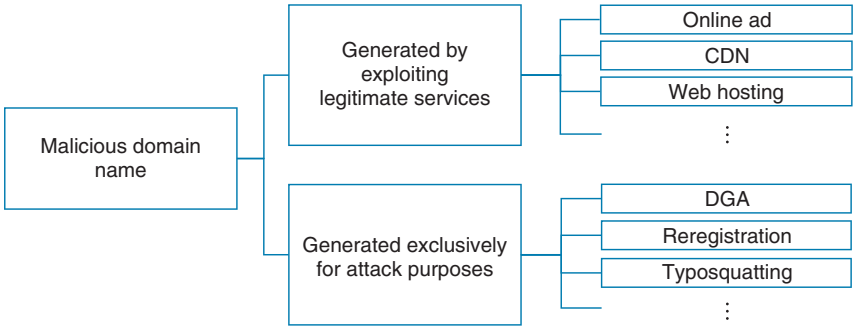
An attacker can avoid uniform countermeasures by generating malicious domain names with different characteristics. For example, malicious domain names include those that are generated by abusing mechanisms that are used by legitimate Internet services. As a result, if such domain names were to be simply blacklisted and blocked, legitimate users or the use of legitimate services might be mistakenly disturbed. However, some malicious domain names are prepared exclusively for the purpose of cyberattacks using techniques such as DGAs. In this case, blocking communications in units of domain names is the most effective approach. With the above information taken into account, even if many malicious domain names can be specified based on their use by attackers, the fact that those malicious domain names may have different generation structures means that those domain names themselves cannot be used effectively as security intelligence.

Under these conditions, we have come to realize the importance of indicating what type of action should be taken as a countermeasure to each malicious domain name in addition to simply indicating



FQDN: fully qualified domain name
2LD: second-level domain
URL: uniform resource locator

Fig. 4. Domain name categorization system.



CDN: content delivery network

Fig. 5. Analysis of generation structure of malicious domain names.

malicious domain names. We therefore set out to develop a domain name categorization system that systematically identifies the generation structure of each malicious domain name specified by the reputation system.

This categorization system systematically determines the generation structure of a malicious domain name that should be taken into account when applying a countermeasure (**Fig. 5**). Specifically, it divides malicious domain names into two main categories. The first category consists of malicious domain names generated by the malicious use of legitimate services. For example, the attacker may exploit online advertising services, CDN (content delivery network) services, or web hosting services for this purpose. Since a domain name used by such a service is inherently established in order to provide a legitimate service, it is necessary here to generate countermeasure information in units of URLs (uniform resource locators) instead of simply blocking certain domain names altogether to avoid erroneous interference with legitimate services.

The second category consists of malicious domain names generated exclusively for attack purposes. These would correspond, for example, to domain names generated by a DGA, domain names reregistered

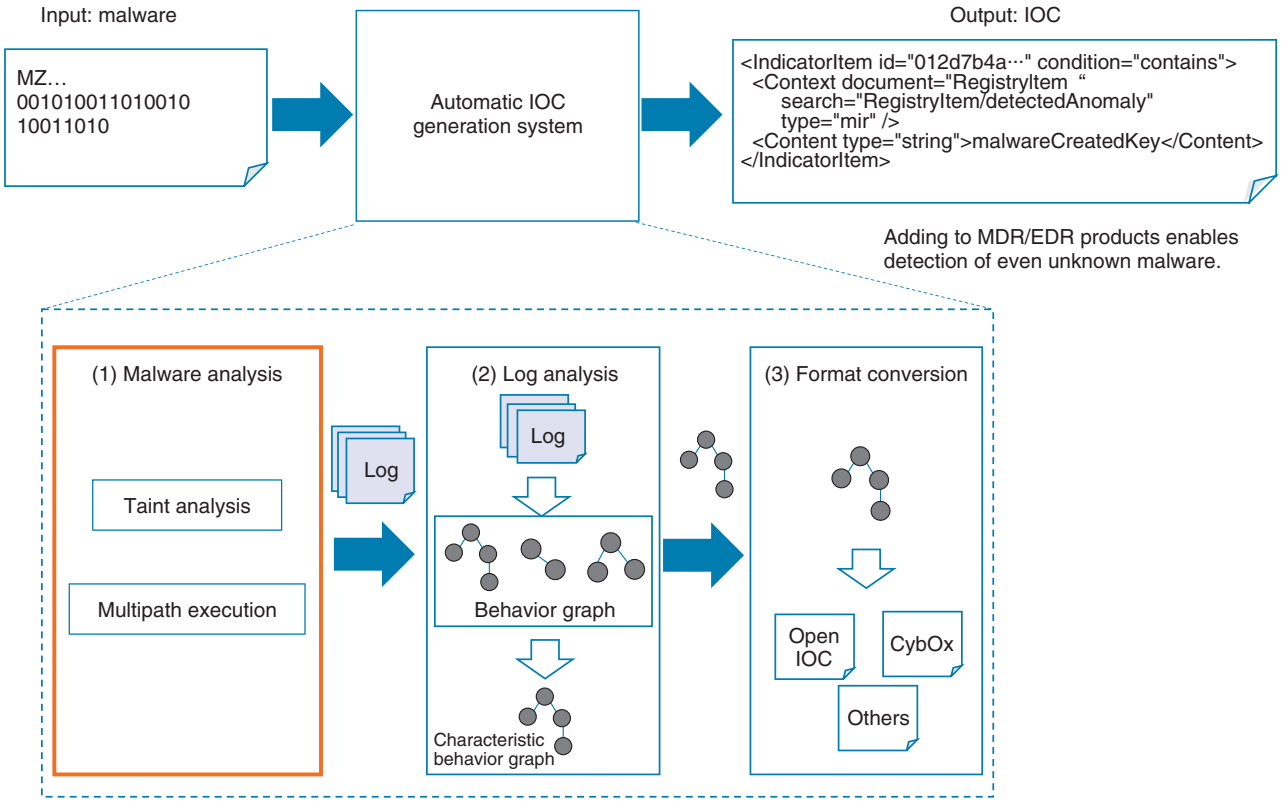


Fig. 6. Automatic IOC generation system.

by the attacker, or domain names generated by typo-squatting, which targets user mistyping. For such malicious domain names used only for attacks, damage from attacks can be prevented by proactively blocking those domain names.

With this categorization system that correctly determines the generation structure of malicious domain names, we have achieved a system that presents the optimal actions to be taken against specific domain names included in a so-called *mixture* consisting of a large number of malicious domain names having various generation structures. This system can provide the most effective countermeasures as security intelligence without having a negative effect on legitimate services.

Implementing the actions generated by this categorization system against actual malicious domain names has effectively prevented cyberattacks without inflicting any damage on legitimate services. The paper describing this system was presented at a major academic conference [3].

2. Malware analysis technology supporting managed detection and response (MDR)

The increasing sophistication of malware as seen in its use in targeted attacks is driving the expansion of conventional security monitoring at the network level and focusing attention on MDR, which includes response measures to attacks, and on endpoint detection and response (EDR), which includes the monitoring of behavior in a host.

To improve MDR services, NTT Secure Platform Laboratories has been researching and developing technology for automatically generating indicators of compromise (IOCs) as definition files that become the grounds for detecting the malicious behavior of malware in a host. This technology analyzes input malware using advanced malware analysis technology ((1) in **Fig. 6**), extracts characteristic behaviors of that malware ((2) in Fig. 6), and generates an NTT proprietary custom IOC based on that behavior ((3) in Fig. 6).

Applying such custom IOCs generated from malware collected from the networks of NTT customers

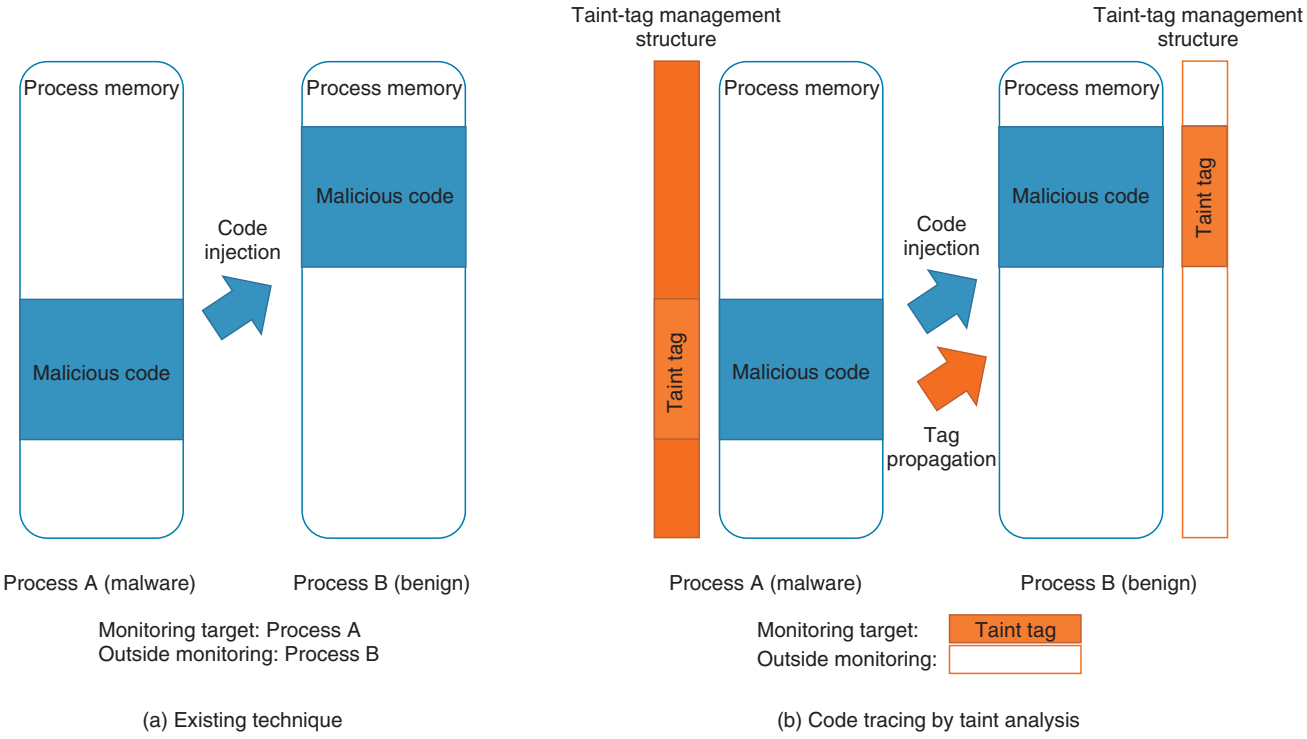


Fig. 7. Determination of analysis target by taint analysis.

enables us to detect specific attacks aimed at NTT customers that are unable to be detected by vendor-provided IOCs designed to defend against major attacks, which are widely seen in many places. In this way, we can provide MDR services that protect the networks and assets of NTT customers.

2.1 Problems with existing technology

Ordinary malware employs various anti-analysis techniques to avoid analysis and detection. These include code injection that injects a portion of malicious code into another process and virtual machine (VM) detection that detects whether the malware itself is running on a VM.

(1) Code injection

This technique injects code into a benign process (e.g., explorer.exe) to perform malicious actions within the process. Standard malware analysis and detection systems treat the processes of malware as monitoring targets, but benign processes are often outside the scope of monitoring, with the result that malicious behavior within benign processes may be overlooked. Additionally, even if benign processes are treated as monitoring targets, it may be difficult to distinguish between behavior driven by benign code in the process and that driven by malicious code in it (**Fig. 7(a)**).

(2) VM detection

This technique collects information on the environment in which the malware itself is running to determine whether that environment is a VM. If the malware determines that it is running on a VM, it terminates malicious activities and starts to behave in a different way (i.e., in a harmless manner) to deceive the analysis and detection system.

2.2 Taint analysis

We apply taint analysis to trace injected code and correctly determine the behavior of executed malware (**Fig. 7(b)**).

Taint analysis is a type of data flow analysis technology that traces the movement of specific data targeted for monitoring in a host. Specifically, it attaches an identifier called a taint tag to the targeted data. This taint tag is managed outside of the usual program execution environment (such as within a VM monitor). In the event that the data affixed with the taint tag are moved or copied, the taint tag as well will be propagated to the move or copy destination. Data flow throughout the host can be analyzed through

repeated propagation of this tag.

In short, even when malware injects its own code into a benign process, using taint analysis in this way to trace data belonging to malware makes it possible to trace that injected behavior and identify malicious code copied into a benign process. This scheme can also identify the behaviors resulting from the execution of the malicious code—that is, the code affixed with the taint tag—and correctly distinguish such behaviors from those generated by the execution of benign code [4].

2.3 Multipath execution

We use multipath execution technology to analyze the multiple paths taken by a malware program and exhaustively extract malware behaviors.

Multipath execution is technology that follows and analyzes the multiple paths that can be taken by a program. An ordinary program can take a variety of paths (specified by *if* statements). Program processing is achieved by changing program behaviors according to such branch conditions. In multipath execution, the analysis system records the branch destination selected when program execution reaches a branch. Then, on completion of that program execution, the system executes the program again and adjusts the state of execution by selecting a branch destination different from the previously selected one. In this way, the system can execute a different path every time the program is executed.

Thus, even if malware should select an execution path that is different from usual if it detects a VM, using multipath execution in this way makes it possible to select other execution paths on reanalysis and extract behavior original to that malware. This technology can exhaustively extract malware behavior.

2.4 Future development

Going forward, we plan to develop technology for tuning IOCs generated using the above technology so they conform with individual endpoints. This will enable our custom IOCs to be used in diverse endpoint products provided by vendors. We also plan to conduct field trials using our custom IOCs with the aim of enhancing MDR services.

3. Security orchestration

In this section, we review two key efforts underway to strengthen security against and recovery from cyberattacks.

3.1 Countering cyberattacks using unified threat management (UTM) in small and medium-sized businesses

Cyberattacks against public organizations and companies continue to evolve as reflected by targeted attacks and ransomware. Countermeasures to cyberattacks are necessary for all enterprises regardless of size. These countermeasures generally take the form of detection and blocking based on virus definition files and signature updates using security appliances. As security consciousness grows even among customers operating small and medium-sized businesses, the introduction of low-priced UTM[*1] is increasing.

In these circumstances, as an effort to deal with cyberattacks targeting small and medium-sized businesses that make up a majority of Japanese enterprises, we introduce a security system that links UTM with a resilient security engine developed by NTT Secure Platform Laboratories and an example of actually using this system within the NTT Group.

3.2 Security orchestration activities

At NTT Secure Platform Laboratories, we have been carrying out R&D of security orchestration technology to achieve rapid recovery from cyberattacks. These efforts have resulted in the development of the Resilient Security Engine (RSE) for information and communication technology (ICT) businesses [5, 6]. This RSE is installed inside a datacenter or on a company network to collect log data from various types of security appliances such as a firewall or WAF (web application firewall). Analyzing this collected information enables detection of attacks, automatic execution of measures on those security appliances based on analysis results, and presentation of attack detection to operators. In addition, the RSE recommends countermeasures, which enables a quick response to cyberattacks and reduces the burden on operators.

The RSE is not limited to dealing with attacks from the outside. With the aim of blocking external access to malware that has already infiltrated an enterprise, as in targeted attacks and zero-day attacks, it includes a function for extracting high-priority blacklists from massive threat-intelligence platforms provided by security vendors at a volume that can be set in firewalls and UTM appliances. This has the same effect

*1 UTM: An appliance that integrates multiple security functions such as firewalls, anti-virus measures, an intrusion prevention system, anti-spam measures, and web filters.

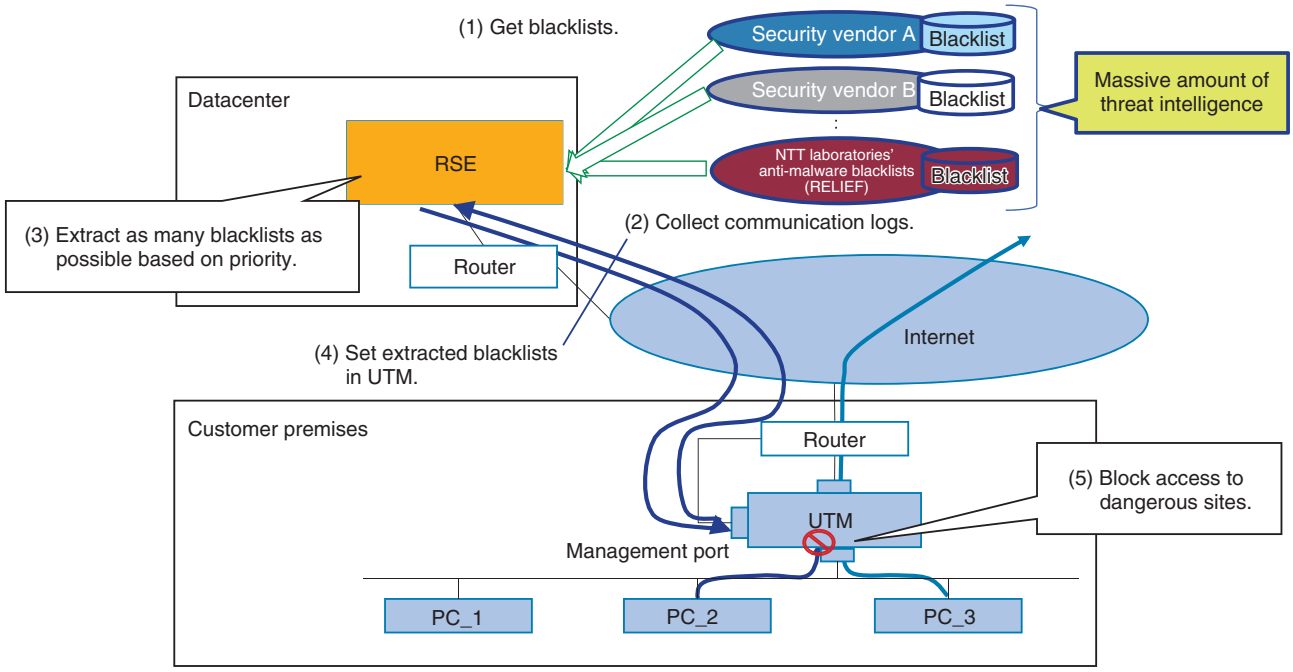


Fig. 8. Blacklist delivery to UTM by RSE.

as setting threat intelligence from multiple security vendors (**Fig. 8**).

3.3 UTM solutions using RSE and use by an NTT Group company

To respond to the escalating security needs of small and medium-sized businesses, NTT EAST maintains and operates UTM appliances under contract with customers. Here, the RSE delivers anti-malware blacklists (RELIEF)[*2]—the threat intelligence platform of NTT Secure Platform Laboratories—to a UTM appliance managed by NTT EAST to add those blacklists to those already possessed by the UTM. This scheme enables customers using the UTM managed by NTT EAST to enjoy even safer use of the network. In this way, the RSE helps differentiate the NTT Group from the security services of other companies and enables the provision of safe-and-secure value-added services in a rapidly growing UTM market.

4. Future development

Going forward, we plan to expand the application of security appliances and undertake the development of new security measures that include not only blacklists but also other types of information.

References

[1] T. Hariu, K. Yokoyama, M. Hatada, T. Yada, T. Yagi, M. Akiyama, T. Ikuse, Y. Takata, D. Chiba, and Y. Tanaka, "Security Intelligence for Malware Countermeasures to Support NTT Group's Security Business," NTT Technical Review, Vol. 13, No. 12, 2015. https://www.ntt-review.jp/archive/ntttechnical.php?contents=ntr201512fa3.html

[2] D. Chiba, T. Yagi, M. Akiyama, T. Shibahara, T. Yada, T. Mori, and S. Goto, "DomainProfiler: Discovering Domain Names Abused in Future," Proc. of the 46th Annual IEEE/IFIP International Conference on Dependable Systems and Networks (DSN 2016), pp. 491–502, Toulouse, France, June/July 2016.

[3] D. Chiba, M. Akiyama, T. Yagi, T. Yada, T. Mori, and S. Goto, "DomainChroma: Providing Optimal Countermeasures against Malicious Domain Names," Proc. of the 41st IEEE Annual Computer Software and Applications Conference (COMPSAC 2017), pp. 643–648, Turin, Italy, July 2017.

[4] Y. Kawakoya, M. Iwamura, E. Shioji, and T. Hariu, "API Chaser: Anti-analysis Resistant Malware Analyzer," RAID 2013, Lecture Notes in Computer Science, Vol. 8145, pp. 123–143, Springer, Berlin, Germany, 2013.

[5] T. Koyama, K. Hato, H. Kitazume, and M. Nagafuchi, "Resilient Security Technology for Rapid Recovery from Cyber Attacks," NTT Technical Review, Vol. 12, No. 7, 2014. https://www.ntt-review.jp/archive/ntttechnical.php?contents=ntr201407fa3.html

*2 RELIEF: The proprietary threat intelligence platform developed by NTT Secure Platform Laboratories consisting of blacklists generated by honeypots and dynamic analysis that includes malicious sites not easily discovered by other companies.

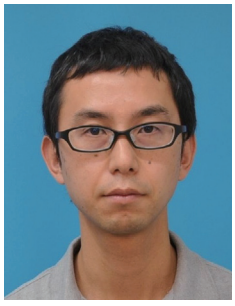
[6] T. Koyama, B. Hu, Y. Nagafuchi, E. Shioji, and K. Takahashi, "Security Orchestration with a Global Threat Intelligence Platform," NTT Technical Review, Vol. 13, No. 12, 2015.
https://www.ntt-review.jp/archive/ntttechnical.php?contents=ntr201512fa4.html

Takeo Hariu
Senior Research Engineer, Supervisor, Cyber Security Project, NTT Secure Platform Laboratories.
He received an M.S. in electro-communications from the University of Electro-Communications, Tokyo, in 1991. Since joining NTT in 1991, he has been engaged in network security R&D. He is a member of the Institute of Electronics, Information and Communication Engineers (IEICE) and the Institute of Electrical Engineers of Japan (IEEJ).

Daiki Chiba
Researcher, Cyber Security Project, NTT Secure Platform Laboratories.
He received a B.E., M.E., and Ph.D. in computer science from Waseda University, Tokyo, in 2011, 2013, and 2017. Since joining NTT in 2013, he has been engaged in research on cybersecurity through data analysis. He won the Research Award from the IEICE Technical Committee on Information and Communication System Security in 2016 and the Best Paper Award from the IEICE Communications Society in 2017. He is a member of the Institute of Electrical and Electronics Engineers (IEEE) and IEICE.

Mitsuaki Akiyama
Research Engineer, Cyber Security Project, NTT Secure Platform Laboratories.
He received an M.E. and Ph.D. in information science from Nara Institute of Science and Technology in 2007 and 2013. Since joining NTT in 2007, he has been researching and developing network security techniques, focusing especially on honeypots and malware analysis.

Takeshi Yagi
Senior Research Engineer, Cyber Security Project, NTT Secure Platform Laboratories.
He received a B.E. in electrical and electronic engineering and an M.E. in science and technology from Chiba University in 2000 and 2002. He also received a Ph.D. in information science and technology from Osaka University in 2013. Since joining NTT in 2002, he has been engaged in the research and design of network architecture and traffic engineering. His current research interests include network security, especially honeypots and security-data analysis based on machine learning. He is a member of IEICE, IEEE, and IEEJ.



Yuhei Kawakoya
Senior Research Engineer, Cyber Security Project, NTT Secure Platform Laboratories.
He received a B.E. and M.S. in science and engineering from Waseda University, Tokyo, in 2003 and 2005. He has been involved in R&D of computer security since 2005. He is a member of the Information Processing Society of Japan (IPSJ) and IEICE.

Yukio Nagafuchi
Senior Research Engineer, Secure Architecture Project, NTT Secure Platform Laboratories.
He received a B.S. and M.S. in science and engineering from Saga University in 1996 and 1998. He joined NTT in 1998 and has been engaged in designing and developing network systems, VoIP (voice over Internet protocol) network systems, virtual network systems, and network security systems. His research interests lie in the area of security orchestration systems for ICT and Internet of Things (IoT) environments. He is a member of IEICE and IPSJ.

Takaaki Koyama
Senior Research Engineer, Supervisor, Secure Architecture Project, NTT Secure Platform Laboratories.
He received a B.A. and M.M.G. in media and governance from Keio University, Kanagawa, in 1994 and 1996. He joined NTT Software Laboratories in 1996 and has been studying software CALS (Continuous Acquisition and Life-cycle Support). Since 1999, he has also been studying GMN-CL (Connectionless networking technologies for Global Megamedia Networks)—a kind of IP-virtual private network technology—and developing network security equipment and operation systems. His research interests have recently extended to security orchestration systems for ICT and IoT environments. He is a member of IPSJ.