2-2 An Incident Analysis Center "nicter" and its Social Commitment

ETO Masashi and TAKAGI Yaichiro

We have been developing the Network Incident analysis Center for Tactical Emergency Response (nicter), whose objective is to detect and identify propagating malwares. The nicter mainly monitors darknet, a set of unused IP addresses, to observe global trends of network threats, while it captures and analyzes malware executables. By correlating the network threats with analysis results of malware, the nicter identifies the root causes (malwares) of the detected network threats. This paper describes the system architecture of each component in nicter. Additionally, this paper reports the achievements of the nicter and its derivational systems that have been practically introduced to other organizations as an actual social commitment.

Keywords
Network monitoring, Malware analysis, Correlation analysis, System architecture, Social commitment

1 Introduction

General use of the Internet has created various security threats as social issues. They have been growing in their volume and complexity due to the availability of more versatile services offered on the Internet. For example, we are observing various types of security incidents (i.e. security accidents) exemplified by unauthorized access to and DoS attacks against Web services, leakage of personal information or corporate confidential information, and phishing caused by a large volume of spam. Many of those incidents are partially caused by malware infecting user machines. To counter this situation, National Institute of Information and Communications Technology (NICT) has been developing an incident analysis center (i.e. Network Incident analysis Center for Tactical Emergency Response: nicter), which detects security incidents greatly impacting the Internet at an early stage, analyzes their root causes, and produces their solutions[1]-[3].

The nicter has two analysis paths: 1) the macro analysis system that analyzes events collected through wide-area darknet monitoring[4]-[8] and detects incidents and 2) the micro analysis system[9]-[16] that collects and analyzes malware specimens and extracts their behaviors (see Fig. 1). Analysis results gained through these two systems are further analyzed for their correlation by the correlation analysis system, which associates an event with a root cause for an incident. In other words, the macro analysis system detects how incidents occur on the Internet[17]. On the other hand, the micro analysis system captures how malware works as a possible cause of incidents[18]. Matching the analysis results of these two systems enables us to detect the cause of current incidents[19]. We can also work on the creation of countermeasures depending on detected malware. The analysis results of the macro analysis system, the micro analysis system, and the correlation analysis system are integrated into the incident handling system which provides analysts with an integrated Web and visibility interface. At the end of the day, analysts will create detailed reports of incidents.

In this paper we refer to the incident whereby the nicter detected attacks against SIP servers on July 9, 2010. Figure 2 shows the statistical data of the number of hosts and packets attacking 5060/UDP from April 1, 2010 to December 31, 2010 as they were observed through the nicter's darknet monitoring network. This incident shows an increase of attacks to the related port numbers on July 9, which means behaviors searching for VoIP/SIP servers with insufficient security measures in place enabling unauthorized usage of IP phones. The nicter detected this incident early and analyzed and identified root causes to provide information to relevant organizations, which enabled ISPs and other parties to issue alerts to their users. In addition to the nicter, the cyber police detected similar attacks based on this incident on the same day and issued alerts on July 14.

As explained above, the nicter leverages its unique wide-area network monitoring system and anti-malware technologies to detect incidents and identify their root causes at an early

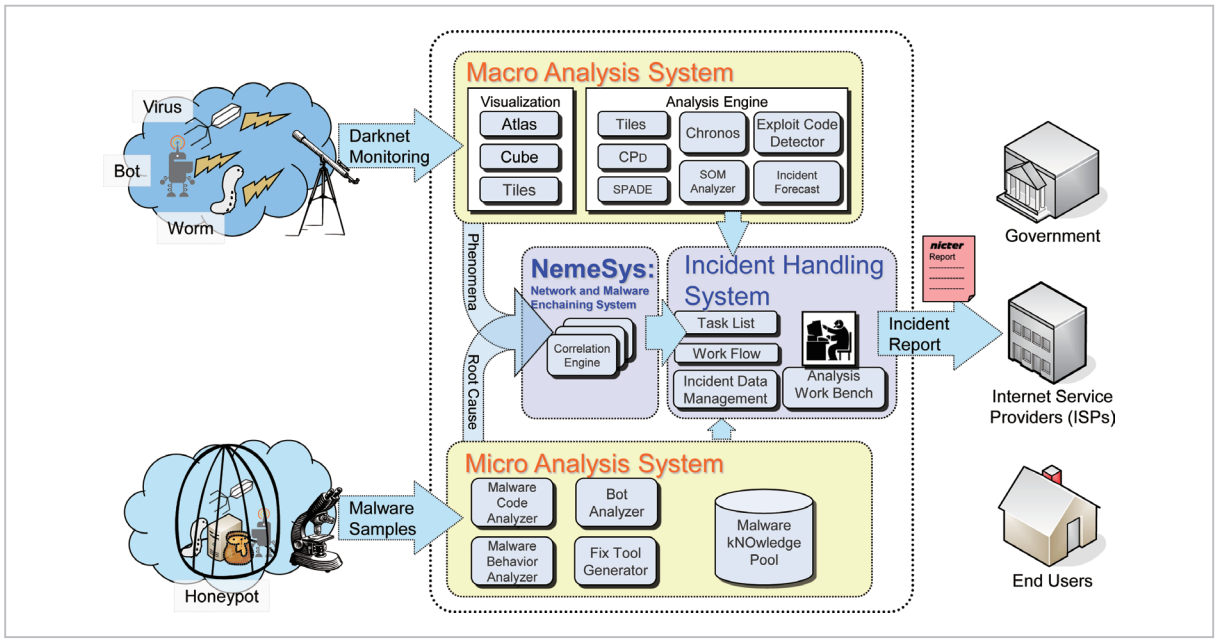

Fig.1 *Overall procedures of the nicter*

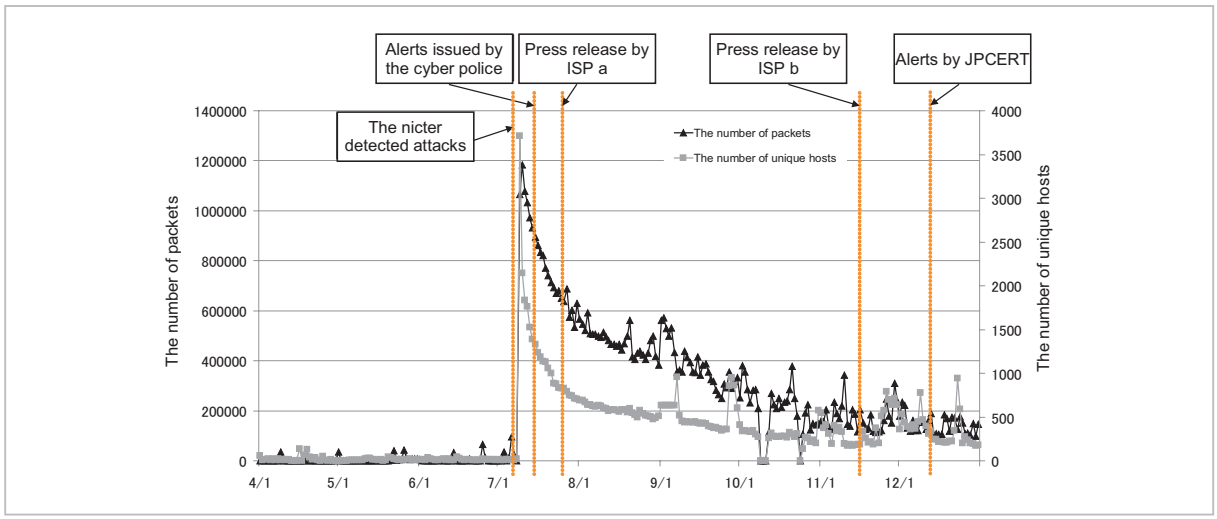

Fig.2 *Detection of attacks against SIP servers through darknet monitoring (04/01/2010-12/31/2010)*

stage.

We have been working more than five years to conduct highly substantive research and development to develop socially viable technologies. The technologies developed through the nicter have resulted in multiple achievements, including technology transfer and information delivery to the society. This paper describes the system architecture of each component in the nicter in detail as well as introducing successful concrete applications of the nicter and its derivational systems that have been leveraged in multiple ways and have enabled social commitment.

2 Overview of the nicter

The nicter includes the correlation analysis system that analyzes results gained through 1) the macro analysis system that analyzes events collected through wide-area monitoring of network traffic and detects/analyzes incidents and 2) the micro analysis system that collects and analyzes malware specimens and extracts their behaviors so that it can associate incidents with malware behaviors (see Fig. 1).

The macro analysis system uses sensors distributed in organizations located in Japan and overseas to monitor darknets. Darknets are networks composed of unused IP addresses out of all addresses disclosed to the Internet. Approximately 140,000 IP addresses are available for this purpose. Darknets do not contain regular hosts (e.g. servers and clients) using these IP addresses. Thus, all traffic monitored can be considered as the results of unauthorized communications or erroneous setup of devices. Furthermore, we can say that darknet monitoring is the most appropriate measure to conduct large-scale monitoring and detect scan by malware (infection propagation) due to its ease of configuration. In particular, the technique to passively record incoming packets in darknets without responding to the source is called black hole monitoring, which is used on the greatest scale by the nicter. The traffic collected this way is analyzed by multiple analysis engines based on each attacker host, enabling the nicter to extract attack characteristics and other features of each host.

On the other hand, the micro analysis system uses such techniques as honeypots, dummy email accounts, and Web crawlers to collect malware specimens. The malware specimens collected this way are injected into dynamic/static malware analysis system for the extraction of characteristics for each malware behavior.

Then, all the macro and micro analysis results are saved into the integrated database called the Malware kNOwledge Pool (MNOP). The macro-micro correlation analysis is made possible through the effective query execution into the MNOP by the NEtwork and Malware Enchaining SYStem (NemeSys). The role of the NemeSys is to query the macro and micro analysis results saved on the MNOP and identify root causes for events detected by the macro analysis system. Specifically, a rapid increase of traffic for specific port numbers on the macro analysis system prompts the NemeSys to output a list of malware names that are highly likely to have caused the event based on the macro-micro correlation analysis. Identifying root causes for incidents in this way enables us to capture trends on the Internet, including currently popular attack techniques across the Internet and the rapid infection propagation of unknown malware.

Finally, human operators use the incident handling system (IHS) to verify these results and publish incident reports to be distributed to related organizations.

3 Macro analysis system

The macro analysis system is composed of a pool of sensors distributed on a wide area, multiple visibility engines, and an analysis system. Sensors monitor darknets in each location mainly through black hole monitoring, integrating obtained traffic into an analysis center. The analysis center uses real-time visibility engines to visualize the collected traffic, enabling operators to conduct a visual check to detect evident attacks by malware. On

the other hand, an automatic analysis engine automatically detects security events (e.g. new attack patterns and a rapid increase in traffic) and stores analysis results on a database as necessary.

The nicter currently uses black hole monitoring to monitor multiple networks with subnets (e.g. /16 and /24). Through this process, scan traffic by malware (e.g. SYN packets by TCP and echo requests by ICMP) have mainly been detected. In addition, attack monitoring sensors called low-interaction honeypots are also deployed.

3.1 Data bus architecture

Effectively responding to incidents requires us to discover suspicious events from obtained data and rapidly come up with countermeasures. To enable this sort of real-time analysis, the nicter has designed and developed its unique data transmission method using IP multicast (i.e. data bus architecture) to effectively provide data to multiple analysis engines and databases.

The data bus system shown by Fig. 3 converts each packet obtained by sensor modules into specific formats composed of minimum data required for analysis (e.g. IP headers, TCP headers, timestamps, and packet length) and transmits them to a gate module located in an analysis center through a VPN line. The gate module transmits the integrated packets to a unique segment (i.e. data bus) as UDP packets based on a multicast technology. The nicter leverages this configuration to enable all analysis engines and databases to obtain data for analysis at the same time. This architecture also adds analysis engines and databases in case of load increase of each system caused by an increase in traffic to simplify the improvement of scalability of the system.

3.2 MacS DB

The data bus architecture explained in Section **3.1**. was created to rapidly transmit data to real-time analysis engines. However, some engines are not real-time and analyze events by processing data accumulated for a certain time period (e.g. several minutes or hours) in batches. This type of analysis engines are the most suitable for database systems with data accumulated over a long period of time instead of real-time data transmission systems (e.g. data bus architecture). We also need to note that since darknet monitoring tries to detect unused IP addresses, the data traffic volume is smaller in comparison with live networks used by ordinary servers or client hosts. However, it is difficult for ordinary databases that collect and accumulate all packets to fully function due to a high load required. Thus, we have developed a special database system (MacS DB) for the nicter so that all packets can be accumulated in real-time.

Figure 4 shows the high-level overview of this system. The system is composed of the master that obtains traffic data from a data bus, the slaves that replicate data from the master and respond to data lookup requests from analysis modules, and the load balancer that distributes the load of slaves and maps lookup

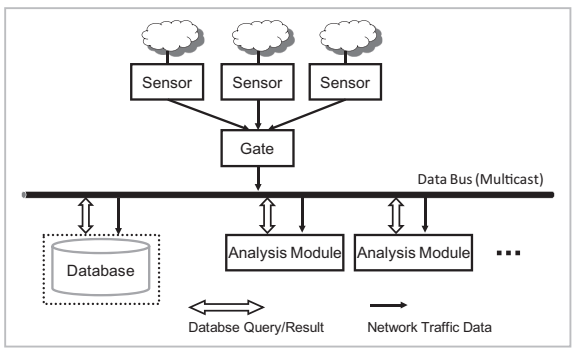

Fig.3 Data bus architecture

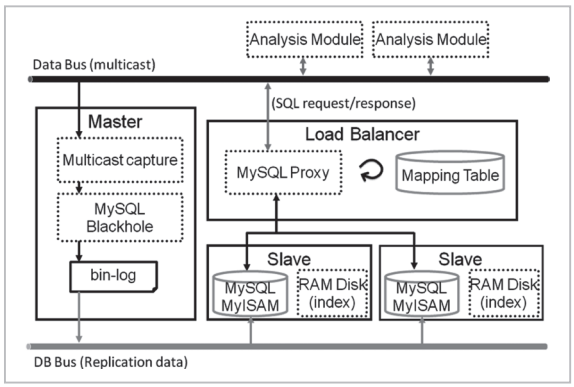

Fig.4 MacS DB

requests to slaves based on the information on a mapping table. Furthermore, the DB bus was deployed to separate traffic data on the data bus from the network traffic created by data replication. It is not effective to manage the large volume of data accumulated each day based on a single database. This system creates a database for each sensor and creates and manages a table for each day.

Generally speaking, load-balancing a database is usually done by the separation of updates from lookups. This system has also been designed so that it can balance the load of each system by separating the master that obtains and inserts traffic data and slaves that respond to lookup requests. MySQL, the open source software we used for this system, allows for the selection of a table engine named Blackhole. The Blackhole engine does not have any storage structure and discards all INSERT statements. However, MySQL outputs queries against the Blackhole engine into log files as usual, enabling the slaves to execute those queries through replication. This system has allowed the master to specify the Blackhole engine so that the load to update indexes and insert data to tables can be excluded. Furthermore, the master has configured binary logs required for replication. The SQL statements output through this configuration are added to binary log files and executed by slaves through replication. The mechanism allows the master to dedicate itself to obtaining traffic data from a data bus, meeting the requirement of high-speed processing by the master.

In addition to these features, various arrangements including placing index areas on RAM disks to reduce the I/O load of slaves have enabled the MacS DB to achieve high-speed data processing of up to 65,000 pps (packets per second). This is a sufficient efficiency considering that the maximum data transmission speed of the gate is around 50,000 pps.

4 Micro analysis system (MicS)

The micro analysis system conducts fully automated, in-depth dynamic/static malware analysis to capture the characteristics and behaviors of obtained malware. As explained before, the nicter collects malware specimens based on multiple honeypots, putting the malware into the micro analysis system and the final analysis results are stored into the MicS DB, which is the database dedicated to the micro analysis system. The micro analysis system is capable of analyzing one sample in five to ten minutes based on a single analysis set (meaning 150 to 250 samples analyzed each day). Currently the nicter deploys multiple analysis sets, capable of analyzing a total of 2,000 samples every day.

4.1 Collecting malware specimens

As explained above, the nicter operates multiple honeypots (e.g. high-interaction or low-interaction honeypots and Web crawlers) to collect malware specimens. The nicter has configured honeypots that behave as regular hosts on real machines (instead of virtual machines) as a type of high-interaction honeypot. One of the roles of these honeypots is to monitor the communication between them and botnet Command and Control (C&C) servers in case they are compromised by malware (especially bot). This mechanism allows all the messages received by these honeypots to be automatically recorded into monitoring servers and monitored by human operators 24/7, enabling us to capture various behaviors of malware. In case the honeypots are infected with malware, they can infect other hosts on the Internet. To prevent this behavior, these honeypots have an IDS deployed to protect them, creating a structure whereby suspicious communication other than C&C messages can be detected and the hard disk image of these honeypots can be automatically restored as required. In case these honeypots are actually infected with malware, they can be restarted at any time so that their disk image can be reinitialized. During the restart process, infected

hard disks and original (pre-infection) disks are compared, extracting only executable files and allowing for the capture of malware specimens. The nicter's high-interaction honeypots leverage these technologies to monitor malware behaviors and obtain malware specimens in a safe manner.

5 Macro-micro correlation analysis system (NemeSys)

The macro-micro correlation analysis system (NemeSys) has been developed so that it can analyze the attacks monitored by the macro analysis system in greater detail and detect their root causes. The NemeSys functions based on the profiling of the network behaviors of attacker hosts. Namely, it uses the technique to detect malware specimens with close proximity to monitored attacker hosts through the comparison of network behavior profiles between the hosts monitored by the macro analysis system and each malware extracted by the micro analysis system.

To rapidly and effectively complete this correlation analysis, we have developed the Malware kNOwledge Pool (MNOP). The MNOP stores summary information about hosts monitored across darknets and malware specimens collected by honeypots and other tools (e.g. monitoring date and time and IP addresses of attacker hosts) as well as all the analysis results by the macro and micro analysis systems. The NemeSys makes multiple queries about the MNOP and completes correlation analysis based on the mapping of this information. Before the MNOP was developed, we had to manually search for information about databases and log information distributed across multiple locations to conduct correlation analysis. The MNOP, however, has made this task both automatic and effective. Figure 5 shows the high-level architecture of the macro-micro correlation analysis based on the MNOP.

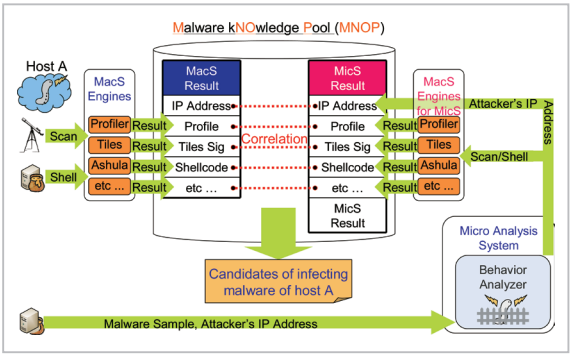

Fig.5 *High-level architecture of macro-micro correlation analysis system*

Traffic data collected by black hole sensors and honeypots based on the macro analysis system are analyzed by scan profilers, Tiles, shell code detection engines, all of which are also run on the macro analysis system, produce analysis results to be stored on the MNOP together with the IP addresses of attacker hosts. On the other hand, malware specimens collected by honeypots are analyzed by the micro analysis system. In particular, the network traffic of each malware extracted by the dynamic analysis system is first analyzed by the analysis engine of the macro analysis system and then its analysis results are stored on the MNOP. The procedures enable both monitored data based on the macro analysis system and network behaviors of malware specimens analyzed by the micro analysis system to be stored in the same data format (i.e. as the analysis results of the analysis engine), making it easier to compare between the two sets of data. The NemeSys queries this database to detect the malware specimens with characteristics matching the scan data of specific attacker hosts detected by black hole monitoring.

6 Social commitment of the nicter

6.1 NIRVANA

The macro analysis system of the nicter visualizes the traffic collected through darknet monitoring in real-time. In particular, a tool called Atlas calculates geographical positions of senders and receivers based on the IP address information included in attacker packets and visualizes them on a global map, enabling us to grasp global trends of network attacks.

NIRVANA is an implementation of the

Fig.6 *Visualization of the Internal NICT network by NIRVANA*

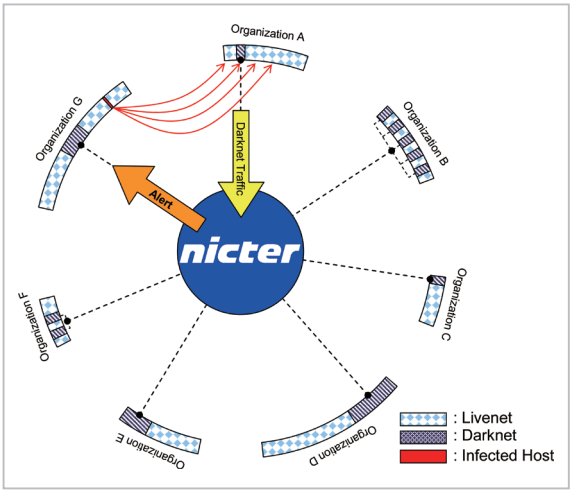

Fig.7 *DAEDALUS mechanism*

nicter's technology in visibility and traffic monitoring applied to the monitoring of specific networks instead of the overall Internet network. Unlike Atlas, NIRVANA captures network traffic flows on a specified network topology instead of a global map. The procedure has enabled us to visually capture network incidents occurring in our internal networks as well as misconfiguration and bottlenecks in our network devices (see Fig. 6).

NIRVANA periodically collects information on routing and other features from routers and switches existing on networks and leverages the information to determine the overall flows of packets across networks. The system has made it possible for us to capture unintended changes of routing caused by network device trouble almost in real-time.

We have been using NIRVANA for the internal network of NICT so that we can detect multiple network incidents, misconfiguration of our devices, and communication bottlenecks in our daily operations. The technology has also been transferred to private corporations to help them investigate their large-scale internal networks covering their multiple domestic locations.

6.2 DAEDALUS

Generally speaking, it is difficult for darknet monitoring to detect the status of real networks that include users and servers since it deals with the unused IP address space. In contrast, the nicter proposes DAEDALUS, a technology to protect real networks through darknet monitoring. DAEDALUS has a quite simple structure that sends alerts to the network of partnering organizations as the nicter's darknet receives scan data from their addresses.

As Fig. 7 shows, the nicter borrows a part of unused IP addresses (shaded areas of Fig. 7) from multiple partnering organizations as its darknet. All the traffic monitored in this area is integrated into the nicter. As malware is detected in a specific organization, it usually scans inside and outside the organization to propagate the infection. If the scan is detected in a darknet, the sender host is likely to be infected with malware. Thus, DAEDALUS has the basic mechanism to issue alerts to the network administrator of the organization. The example of Fig. 7 shows that the host infected with malware in the network of Organization G scans against Organization A to propagate infection. DAEDULUS detects the event through the darknet of Organization A and issues alerts including the address information of attacker hosts to the network administrator of Organization G.

This mechanism is only made possible since it can cover a wide area through its enhanced darknet monitoring network. It is critical that we work with more organizations in the future to further expand our monitoring network.

DAEDALUS has actually been operated

for more than two years by a domestic organization. It has so far detected multiple malware infections and issued alerts to the network administrator to proactively prevent the propagation of malware infection within the organization.

6.3 Wide-area sensor deployment in Japan

NICT has been collaborating with several organizations including domestic universities to work on the commissioned research entitled "Research on technologies enabling a rapid analysis of incidents in a wide area" since FY2008 to enhance the scope of nicter monitoring and its analysis capabilities (see Fig. 8). This initiative has added 70,000 IP addresses to the darknet monitoring network by the nicter, paving the way for monitoring networks of a larger-scale.

On the other hand, we have been serving the improvement of network security of various partnering organizations by providing them with visibility engines and various statistical data available through the above-mentioned nicter and alerts issued by DAEDALUS. The DAEDALUS has already detected malware infections of multiple organizations and misconfiguration of network devices even within a short period of its implementation through its proof of concept experiments started in FY2010, attracting many favorable evaluations from our partnering organizations.

The framework we have created based on this commissioned research will continue after the completion of the research, allowing us to further increase the number of organizations collaborating with us and to enhance monitoring networks.

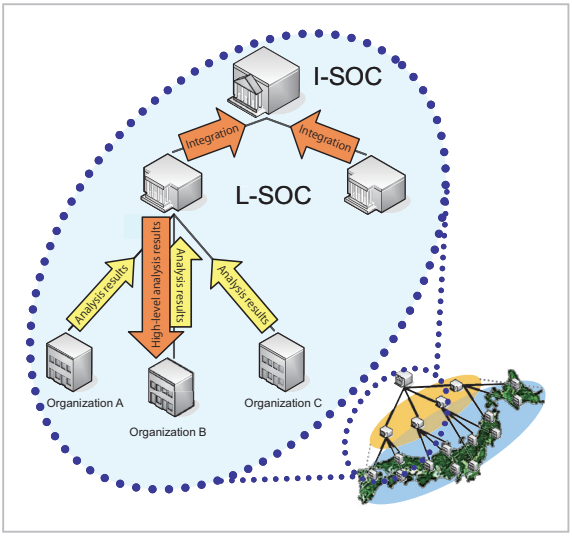

Fig.8 *Overview of commissioned research in FY2008*

7 Conclusion

We have been working on a wide area of research on network security encompassing darknet monitoring and malware analysis through the nicter, which has produced technologies being applied to our society in various manners.

On the other hand, more complexity in network environments is expected to lead to further sophisticated attacks on networks including malware. For example, security research in IPv6 environments still includes a lot of areas that need further clarification, requiring us to work on systematic countermeasures before we start a full-scale implementation. Furthermore, multiple Web services and P2P applications targeting the smartphone market and SNS are at the peak of their implementation, posing a challenge for us to address security issues created by these platforms and service applications.

The nicter is going to deal with security issues required in these latest environments, enabling us to work on research and development so that we can contribute to our society with its concrete results.

References

1 K. Nakao, K. Yoshioka, D. Inoue, and M. Eto, "A Novel Concept of Network Incident Analysis based on Multi-layer Observations of Malware Activities," The 2nd Joint Workshop on Information Security

(JWIS07), pp. 267–279, 2007.
2 D. Inoue, M. Eto, K. Yoshioka, S. Baba, K. Suzuki, J. Nakazato, K. Ohtaka, and K. Nakao, “nicter: An Incident Analysis System Toward Binding Network Monitoring with Malware Analysis,” WOMBAT Workshop on Information Security Threats Data Collection and Sharing (WISTDCS 2008), pp. 58–66, 2008.
3 K. Nakao, K. Yoshioka, D. Inoue, M. Eto, and K. Rikitake, “nicter: An Incident Analysis System using Correlation between Network Monitoring and Malware Analysis,” The 1st Joint Workshop on Information Security (JWIS06), pp. 363–377, 2006.
4 M. Bailey, E. Cooke, F. Jahanian, J. Nazario, and D. Watson, “The Internet Motion Sensor: A distributed blackhole monitoring system,” Proceedings of the 12th ISOC Symposium on Network and Distributed Systems Security (NDSS), pp. 167–179, Citeseer, 2005.
5 SANS Internet Storm Center, http://isc.sans.org/
6 F. Pouget, M. Dacier, and V.H. Pham, “Leurre.com: On the Advantages of Deploying a Large Scale Distributed Honeypot Platform,” E-Crime and Computer Conference (ECCE’05), 2005.
7 D. Moore, C. Shannon, G.M. Voelker, and S. Savage, “Network telescopes: Technical report,” CAIDA, April 2004.
8 M. Bailey, E. Cooke, F. Jahanian, A. Myrick, and S. Sinha, “Practical darknet measurement,” Information Sciences and Systems, 2006 40th Annual Conference on, pp. 1496–1501, IEEE, 2007.
9 N. Provos, “Honeyd-a virtual honeypot daemon,” 10th DFNCERT Workshop, Hamburg, Germany, 2003.
10 C. Leita, M. Dacier, and F. Massicotte, “Automatic handling of protocol dependencies and reaction to 0-day attacks with ScriptGen based honeypots,” Recent Advances in Intrusion Detection, pp. 185–205, Springer, 2006.
11 N. Provos, “A virtual honeypot framework,” Proceedings of the 13th conference on USENIX Security Symposium-Volume 13, p. 1, USENIX Association, 2004.
12 E. Alata, V. Nicomette, M. Kaâniche, M. Dacier, and M. Herrb, “Lessons learned from the deployment of a high-interaction honeypot,” Dependable Computing Conference, 2006, EDCC’06. Sixth European, pp. 39–46, IEEE, 2006.
13 R. Isawa, S. Ichikawa, Y. Shiraishi, M. Mori, and M. Morii, “A Virus Analysis Supporting System-For automatic grasping virus behavior by code-analysis result,” Joho Shori Gakkai Shinpojiumu Ronbunshu, 1(13): 169–174, 2005.
14 D. Inoue, M. Eto, K. Yoshioka, Y. Hoshizawa, R. Isawa, M. Morii, and K. Nakao, “Micro analysis system for analyzing malware code and its behavior on nicter,” Symposium on Cryptography and Information Security (SCIS) 2007, IEICE, Jan. 2007.
15 C. Willems, T. Holz, and F. Freiling, “Toward automated dynamic malware analysis using cwsandbox,” IEEE Security & Privacy, pp. 32–39, 2007.
16 N. Solutions, Norman sandbox whitepaper, 2003.
http://download.norman.no/whitepapers/whitepaper/Norman SandBox.pdf
17 D. Inoue, K. Yoshioka, M. Eto, M. Yamagata, E. Nishino, J. Takeuchi, K. Ohkouchi, and K. Nakao, “An Incident Analysis System nicter and Its Analysis Engines Based on Data Mining Techniques,” 15th International Conference on Neuro-Information Processing of the Asia Pacific Neural Network Assembly (ICONIP 2008), 2008.
18 D. Inoue, K. Yoshioka, M. Eto, Y. Hoshizawa, and K. Nakao, “Malware Behavior Analysis in Isolated Miniature Network for Revealing Malware's Network Activity,” IEEE International Conference on Communications (ICC 2008), pp. 1715–1721, 2008.

19 K. Nakao, D. Inoue, M. Eto, and K. Yoshioka, “Practical Correlation Analysis between Scan and Malware Profiles against Zero-Day Attacks Based on Darknet Monitoring,” IEICE TRANSACTIONS on Information and Systems, 92(5): 787–798, 2009.

(Accepted June 15, 2011)

ETO Masashi, *Ph.D.*

Senior Researcher, Cybersecurity Laboratory, Network Security Research Institute

Network Security, Malware Analysis, Network Operation

TAKAGI Yaichiro

Technical Expert, Cybersecurity Laboratory, Network Security Research Institute

Network Security, Network Traffic Analysis