Design and Evaluation of a Diffusion Tracing Function for Classified Information among Multiple Computers

Nobuto Otsubo, Shinichiro Uemura, Toshihiro Yamauchi, and Hideo Taniguchi

Graduate School of Natural Science and Technology, Okayama University, Japan

{yamauchi,tani}@cs.okayama-u.ac.jp

Abstract. In recent years, the opportunity to deal with classified information in a computer has increased, so the cases of classified information leakage have also increased. We have developed a function called "diffusion tracing function for classified information" (tracing function), which has the ability to trace the diffusion of classified information in a computer and to manage which resources might contain classified information. The classified information exchanged among the processes in multiple computers should be traced. This paper proposes a method which traces the diffusion for classified information among multiple computers. Evaluation results show the effectiveness of the proposed methods.

Keywords: Prevention of information leaks, Network security, Log management

1 Introduction

The improvement in computer performance and propagation in various services has increased the opportunity to deal with classified information, such as customer information. According to the analysis [1] of personal information leakage incidents, it has been reported that leaks often happen by inadvertent handling and mismanagement, which account for approximately 57% of all known cases of information leakage. In addition, several employees often share classified information. To trace the status of classified information in a computer and to manage the resources that contain classified information, we proposed a diffusion tracing function for classified information (tracing function), which manages any process that has the potential to diffuse classified information [2].

In this paper, we propose a method that uses the tracing function to trace the classified information being exchanged among multiple computers in internal network and to prevent information leakage outside internal network.

adfa, p. 1, 2011.

2 Requirements of diffusion tracing function for classified information among multiple computers

By tracing how classified information is diffusion, a computer can know which resources contain classified information. However, the tracing function [2] only traces the status of classified information in a computer. Thus, we propose a method that uses tracing function to trace classified information being exchanged among multiple computers in internal network and to prevent information leakage outside internal network.

In order to prevent information leakage outside the internal network, function needs to centrally manage the classified information that exists in the client computers in network. Moreover, it is necessary to determine whether the client computers are installed the tracing function or not, in order to prevent the diffusion of classified information to not installed computers.

We split the computers in network into a managed network and a non-managed network. Computers that need to handle classified information would be in the managed network, and the tracing function would be installed on them. A diffusion tracing function for classified information among multiple computers (tracing function for networks) must meet the following requirements:

(1) Computers that are installed the tracing function can be distinguished from computers that are not.
(2) The location and the flow of classified information in the managed network can be managed.
(3) The diffusion of classified information in the managed network can be instantly traced.
(4) Leakage of classified information out of the managed network can be detected.
(5) Leakage of classified information out of the managed network can be stopped in advance.

The tracing function for networks must also meet the following requests:

(1) Accurately trace the diffusion of classified information in a managed network.
(2) The processing overhead of tracing function for networks should be small.

3 Design

3.1 Overview of the Proposed Function

Figure 1 shows an overview of the function design to handle classified information in a managed network. A managed network consists of a single management server and multiple client computers installed the tracing function. The following describes what is achieved by using the tracing function for networks.

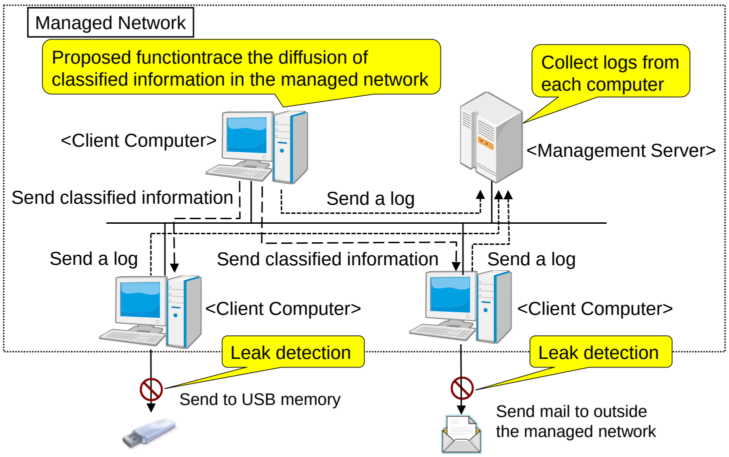

Managed Network
Proposed functiontrace the diffusion of classified information in the managed network
Collect logs from each computer
<Client Computer>
<Management Server>
Send classified information
Send a log
Send a log
Send classified information
Send a log
<Client Computer>
<Client Computer>
Leak detection
Leak detection
Send to USB memory
Send mail to outside the managed network

Fig. 1. Overview of diffusion tracing function for classified information among multiple computers

(1) Before sending data to another client computer, a process managed by tracing function (hereafter, managed process) in the client computer queries the management server to determine whether the tracing function is installed on the receiving computer. If the tracing function is installed, the transmission of classified information is allowed. If the function is not installed, the transmission is disallowed. This makes it possible to satisfy function requirements (1), (4), and (5).

(2) The tracing function writes a log in the client computer and transfers that log to the management server. The collected logs in the management server are used to manage the location and diffusion of classified information in the managed network. This satisfies function requirement (2).

(3) Before a managed process sends classified information, a managed process in the sending computer notifies sending classified information to the receiving computer. Then tracing function in the receiving computer mark a process receiving the classified information as managed process. This satisfies function requirement (3).

In order to satisfy requirements (1) and (3), communication functions must ensure that the receiving client computer is installed the tracing function before transmitting classified information. In the following sections, we propose a design of the function.

3.2 Communication method to pass the information to be managed

In order to control the communication of classified information, an application program, which monitors the communication of classified information (hereafter, communication monitor), is developed. Before computers exchange classified information, the communication monitor in the sending computer sends the destination port number and IP address of the receiver socket to the receiving

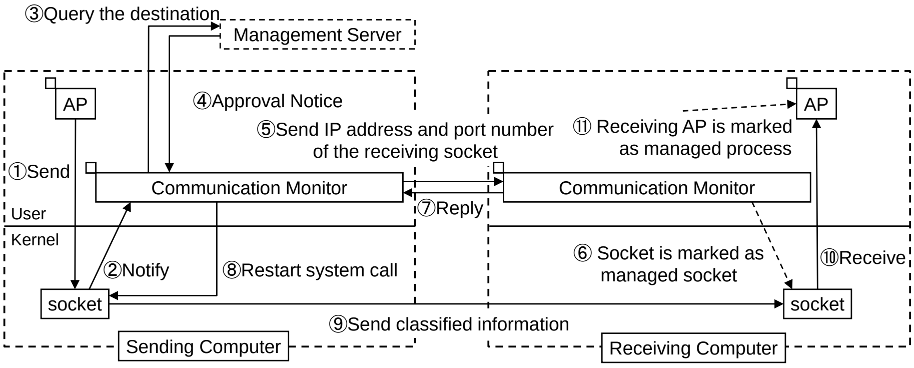


Fig. 2. Sending classified information to other computer AP

computer. The receiving computer uses port number to mark a receiver socket as managed socket. Then the receiving computer marks a process receiving classified information from managed socket as managed process.

The management server maintains the list of IP address of client computers installed the tracing function. The communication monitor in the sending computer queries whether a receiving computer has been registered in the management server list, and, if so, the transmission process is initiated.

Figure 2 shows the flow of process using the communication monitor to send the classified information from the sending computer to the receiving computer as described below:

(1) Send system call of managed process is invoked.
(2) The processing of the send system call is suspended, and the IP address and port number of the receiver socket is sent to the communication monitor.
(3) The communication monitor queries the management server whether the receiving computer has been installed tracing function.
(4) If the receiving computer has been installed, the management server notifies the communication monitor that it has been allowed to send classified information. If it is not installed, the management server notifies the communication monitor that it is not allowed to send classified information.
(5) If transmission of classified information is allowed, the communication monitor in the sending computer sends the port number of the receiver socket to the receiving computer. If transmission of classified information is not allowed, the function terminates the processing of the send system call.
(6) The communication monitor in the receiving computer marks a socket using the receiving port number as managed socket.
(7) The communication monitor in the receiving computer replies to the communication monitor in the sending computer that it is ready to receive the classified information.
(8) After receiving the confirmation, the communication monitor in the sending computer restarts the send system call that was suspended in step (2).

Table 1. File transfer time by FTP (ms)

	Data size of a transmitting file	
	1MB	10MB
Function before implementation	83	885
Function after implementation	233	2010
Overhead	150 (181%)	1125 (127%)

(9) Classified information is sent.

(10) The application program in the receiving computer initiates a receiving system call and receives the classified information from the managed socket.

(11) The receiving AP is marked as a managed process.

4 Evaluation

4.1 Overhead

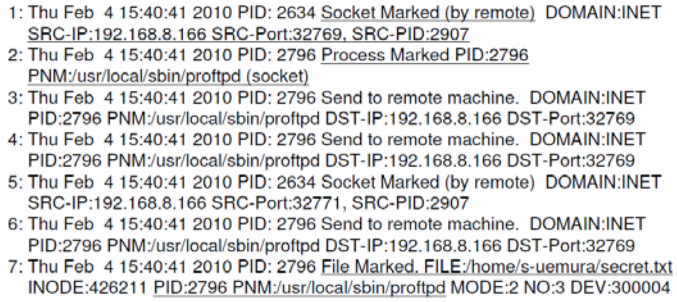
In order to evaluate overhead of the diffusion tracing function for classified information among multiple computers, we measured the time required to transfer managed files to other computers. We measured the time it takes to upload a managed file (1 MB to 10 MB) from an FTP client to an FTP server.

We used the following configuration as our measurement environment: the computer of FTP client has a Celeron D 2.8 GHz CPU and 768 MB memory. The computer of FTP server has a Pentium 4 3.0 GHz CPU and 512 MB memory. The two computers are connected using Ethernet 100Base-TX. OS used on both computers is Linux-2.6.0.

The measurement results are shown in Table1. We found that the transfer time using the managed kernel function is a maximum of about 2.8 times of the transfer time using the original kernel function. This is because the communication monitor communicates with the management server when the send system call is called. Therefore, in order to reduce overhead, it is necessary to reduce the exchange of information between the communication monitor and the management server.

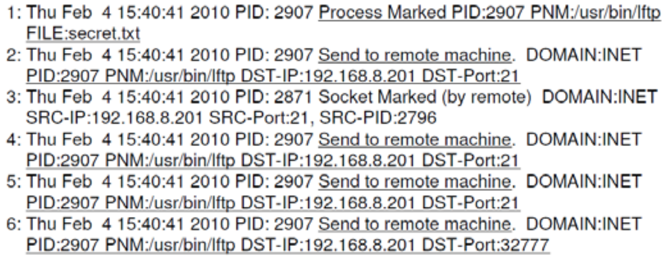
4.2 Evaluation of diffusion tracing of classified information from logs

The diffusion of classified information among multiple computers can be analyzed by tracing the flow of classified information, using data in collected logs during FTP file transfer of the classified information. Logs from two the computer of FTP client and the computer of FTP server were evaluated after uploading files with classified information from the client to the server. ProFTP was used for the FTP server; LFTP was used for the FTP client.



```
1: Thu Feb  4 15:40:41 2010 PID: 2634 Socket Marked (by remote)  DOMAIN:INET
   SRC-IP:192.168.8.166 SRC-Port:32769, SRC-PID:2907
2: Thu Feb  4 15:40:41 2010 PID: 2796 Process Marked PID:2796
   PNM:/usr/local/sbin/proftpd (socket)
3: Thu Feb  4 15:40:41 2010 PID: 2796 Send to remote machine.  DOMAIN:INET
   PID:2796 PNM:/usr/local/sbin/proftpd DST-IP:192.168.8.166 DST-Port:32769
4: Thu Feb  4 15:40:41 2010 PID: 2796 Send to remote machine.  DOMAIN:INET
   PID:2796 PNM:/usr/local/sbin/proftpd DST-IP:192.168.8.166 DST-Port:32769
5: Thu Feb  4 15:40:41 2010 PID: 2634 Socket Marked (by remote)  DOMAIN:INET
   SRC-IP:192.168.8.166 SRC-Port:32771, SRC-PID:2907
6: Thu Feb  4 15:40:41 2010 PID: 2796 Send to remote machine.  DOMAIN:INET
   PID:2796 PNM:/usr/local/sbin/proftpd DST-IP:192.168.8.166 DST-Port:32769
7: Thu Feb  4 15:40:41 2010 PID: 2796 File Marked. FILE:/home/s-uemura/secret.txt
   INODE:426211 PID:2796 PNM:/usr/local/sbin/proftpd MODE:2 NO:3 DEV:300004
```

Fig. 3. Log output from the computer of FTP client



```
1: Thu Feb  4 15:40:41 2010 PID: 2907 Process Marked PID:2907 PNM:/usr/bin/lftp
   FILE:secret.txt
2: Thu Feb  4 15:40:41 2010 PID: 2907 Send to remote machine.  DOMAIN:INET
   PID:2907 PNM:/usr/bin/lftp DST-IP:192.168.8.201 DST-Port:21
3: Thu Feb  4 15:40:41 2010 PID: 2871 Socket Marked (by remote)  DOMAIN:INET
   SRC-IP:192.168.8.201 SRC-Port:21, SRC-PID:2796
4: Thu Feb  4 15:40:41 2010 PID: 2907 Send to remote machine.  DOMAIN:INET
   PID:2907 PNM:/usr/bin/lftp DST-IP:192.168.8.201 DST-Port:21
5: Thu Feb  4 15:40:41 2010 PID: 2907 Send to remote machine.  DOMAIN:INET
   PID:2907 PNM:/usr/bin/lftp DST-IP:192.168.8.201 DST-Port:21
6: Thu Feb  4 15:40:41 2010 PID: 2907 Send to remote machine.  DOMAIN:INET
   PID:2907 PNM:/usr/bin/lftp DST-IP:192.168.8.201 DST-Port:32777
```

Fig. 4. Log output from the computer of FTP server

Figure 3 shows the log output from the computer of FTP client. From the first line of the log, we find that the process PID 2907 of the FTP client “lftp” is a managed process, because it read the classified information file secret.txt. From lines 2 and 4-6 of the log, we see that the process PID 2907 has sent the data to the computer with the IP address 192.168.8.201.

Figure 4 shows the log output from the computer of FTP server. From the first line of the log, we find that a socket receiving classified information from the computer with the IP address 192.168.8.166 is marked as managed socket. From the second line of the log, the process PID 2796 of the FTP server “proftp” is marked to be managed. Lines 2-4 and 6 indicate that the managed process PID 2796 will exchange data with a computer with the IP address 192.168.8.166. Line 7 of the log shows that the process PID 2796 writes data to a managed file called secret.txt; this file is marked as a managed file.

From these logs, the proposed methods can trace the flow of classified information from the computer with the IP address 192.168.8.166 to the computer with the IP address 192.168.8.201.

5 Related Work

By using another connection to send address range of the data when sending tainted data, received data is to be tainted, to trace the diffusion of classified information among multiple computers [3]. However, there is a problem that unable to trace the diffusion of classified information by UDP. On the other hand, by storing the address range of the data to improve the header of the packet, taint is imparted to the header, to trace the diffusion of classified information among multiple computers [4]. However, there is a problem that when send the packet to a computer that does not eliminate the header of the packet, improved header is treated as data. Multiple virtual machines can be processed on a trusted virtual machine monitor, providing isolated virtual environments on a per-machine basis to serve as mechanisms that prevent other users from accessing files [5]. To limit access to files by isolating environmental in the program unit, the mechanism used is to assign the domain name of the group to the files, and same domain user only access the files [6].

6 Conclusion

We proposed a diffusion tracing function of classified information among multiple computers to prevent information leakage outside internal networks. The classified information exchanged among the processes in multiple computers should be traced. Therefore, before computers exchange classified information, the proposed function sends IP address and the port number of the receiver socket. Then, the diffusion tracing function for classified information in the receiving computer traces the receiving classified information. An evaluation of logs shows that the proposed function traces a diffusion of classified information among multiple computers.

In future work, we will reduce the overhead of the proposed function.

References

1. Japan Network Security Association, 2008 Information Security Incident Survey Report, http://www.jnsa.org/result/incident/data/2008incident_survey_e_v1.0.pdf
2. Tabata, T., Hakomori, S., Ohashi, K., Uemura, S., Yokoyama, K., Taniguchi, H.: Tracing Classified Information Diffusion for Protecting Information Leakage. IPSJ Journal. Vol.50, No.9, pp. 2088-2012 (2009) (in Japanese)
3. Kim, C. H., Keromytis, D. A., Covington, M., Sahita, R.: Capturing Information Flow with Concatenated Dynamic Taint Analysis. 2009 International Conference on Availability, Reliability and Security (ARES 2009). pp. 355-362 (2009)
4. Zavou, A., Portokalidis, G., Keromytis, D. A.: Taint-Exchange: a Generic System for Cross-process and Cross-host Taint Tracking. The 6th International Workshop on Security (IWSEC 2011). LNCS. Vol.7038. pp. 113-128 (2011)
5. Garnkel, T., Pfaff, B., Chow, J., Rosenblum, M., Boneh, D.: Terra: a virtual machine-based platform for trusted computing. Proceedings of 19th ACM SIGOPS Symposium on Operating System Principles (SOSP 2003). pp. 193-206 (2003)
6. Katsuno, Y., Watanabe, Y., Furuichi, S., Kudo, M.: Chinese-Wall Process Confinement for Practical Distributed Coalitions. Proceedings of 12th ACM Symposium on Access Control Models and Technologies (SACMAT2007). pp. 225-234 (2007)