DO DATA LOSS PREVENTION SYSTEMS REALLY WORK?

Sara Ghorbanian, Glenn Fryklund, Stefan Axelsson

Chapter 20

DO DATA LOSS PREVENTION SYSTEMS REALLY WORK?

Sara Ghorbanian, Glenn Fryklund and Stefan Axelsson

Abstract The threat of insiders stealing valuable corporate data continues to escalate. The inadvertent exposure of internal data has also become a major problem. Data loss prevention systems are designed to monitor and block attempts at exposing sensitive data to the outside world. They have become very popular, to the point where forensic investigators have to take these systems into account. This chapter describes the first experimental analysis of data loss prevention systems that attempts to ascertain their effectiveness at stopping the unauthorized exposure of sensitive data and the ease with which the systems could be circumvented. Four systems are evaluated (three of them in detail). The results point to considerable weaknesses in terms of general effectiveness and the ease with which the systems could be disabled.

Keywords: Data leakage prevention systems, evaluation, forensic implications

1. Introduction

The theft of sensitive corporate information has always been, and continues to be, a serious problem. While the insider threat (malicious and accidental) should not be exaggerated, fully half of all security incidents reported by businesses are attributed to insiders [4]. The potential loss from insider crime is very high because malicious insiders are difficult to detect, they often have access to sensitive information and they have intimate knowledge of what to take.

Systems specifically designed to identify and protect sensitive data leakage were first introduced in 2006 [6]. They have come to be known as data leakage prevention systems or data loss prevention systems. The purpose of these systems is to detect and stop unauthorized attempts to

leak or export sensitive data. Several data loss prevention systems are available for various operating systems and mobile platforms.

Data loss prevention systems are commonly used in corporate environments and are, therefore, beginning to be encountered in forensic investigations of suspected data leaks. However, very little research has focused on how these systems operate and their ability to prevent data loss. This chapter examines the effectiveness of four well-known data loss prevention systems in a range of leakage scenarios. The results point to considerable weaknesses in terms of general effectiveness and the ease with which the systems could be disabled.

2. Related Work

Considerable enterprise-related research has been conducted in the data loss prevention area, but the academic research is relatively sparse. Most of the research has focused on identifying the best data loss prevention system based on user needs. One example is the report by Ouellet [10], which compares systems from Trustwave, McAfee, Symantec and eleven other vendors, and lists their strengths and weaknesses.

Blasco et al. [2] have examined methods for bypassing data loss prevention systems using trusted applications. A trusted application is a piece of software that has been approved to be used in an otherwise restricted environment. Blasco et al. demonstrate that, by encrypting secret data and using only trusted applications (in this case, an ordinary spreadsheet), a user is able to leak information. The data loss prevention system did not identify the data as sensitive and, because the application was classified as trusted, the data leakage was not detected.

Carvalho and Cohen [3] have proposed a technique for preventing email leakage in scenarios where emails with sensitive information are sent (intentionally or unintentionally) to unauthorized recipients. The technique, which relies on machine learning, was able to detect leakage in 82% of the test cases.

Kim and Kim [7] have proposed a data loss prevention architecture that takes user privacy into consideration. Their research examines the trade-off between information leakage prevention and privacy protection. A scoring module is suggested for computing the levels of security and privacy. The scores are used to discern the number of times private information has been reviewed by a data loss prevention system.

Luft [8] investigated if data loss prevention systems can actually stop data leakage. Evaluations of two data loss prevention systems indicated that the systems had problems preventing data from leaking. Luft also discovered that the systems did not properly secure communications be-

tween the data loss prevention server and agents, making it possible to intercept and eavesdrop on information such as incident reports about secret or sensitive data that had been being blocked. Since 2009, when the evaluation was conducted, data loss prevention systems have made rapid advancements; nevertheless, continued evaluations of the systems are required. Although Luft examined the security of data loss prevention systems, no research has specifically focused on how data loss prevention systems can be manipulated to cause data leakage. This is one of the important issues discussed in this work.

Balinsky et al. [1] and Wuchner and Pretschner [11] have studied agent-based implementations of data loss prevention systems. Balinsky et al. propose the interception of operating system calls to manage read and write access to data. Wuchner and Pretschner continued the work of Balinsky et al. by applying an adapted policy model. Both solutions require restrictions to be imposed on user permissions and access to critical files. The two groups of researchers also identify vulnerabilities in their data loss prevention solutions. One is the possibility of bypassing their solutions because all possible function calls that could be used to access data are not intercepted. Other weaknesses exist in policy management, which make it possible to execute man-in-the-middle attacks.

3. Evaluated Systems

Data loss prevention systems constantly monitor data to prevent the unauthorized movement of data from secured sites. In general, data exists in three states [2]:

- **Data in Motion:** This corresponds to data that is being transferred from one location to another. It could involve a file being moved from one hard drive or an email moving across the Internet.
- **Data in Use:** This corresponds to data that is actively being used by software. It could involve a Word document that has not recently been saved to the hard drive.
- **Data at Rest:** This corresponds to data that is stored on some media and is not actively being used or transferred.

A data loss prevention system operates under a set of policies that guide decisions and help achieve rational outcomes. The policies are set by administrators and incorporate rules governing the network and endpoints (i.e., what usage is allowed and what is not). The policies can be set on specific applications or web pages, and on anything that makes file transfer possible, such as email, instant messaging and transfers to

an external hard drive. Different policies may be set according to the levels of access granted to users. The more specific the policies, the lower the number of false positive and false negative alerts, which results in a more accurate data loss prevention system.

Data loss prevention systems can be categorized into three types: agent-based, agentless and hybrid systems. An agent-based system incorporates an agent at each endpoint, which communicates with a data loss prevention server that delegates policies [9]. An agentless system incorporates one or more servers that monitor and analyze network traffic; every endpoint is forced to have its traffic routed through a server and nothing is installed at the endpoints. A hybrid system is a combination of agent-based and agentless systems; it incorporates an agent at each endpoint as well as one or more monitoring servers.

To gain an understanding of how data loss prevention systems operate and their ability to prevent data loss, the following four systems were evaluated:

- **My Endpoint Protector:** `www.endpointprotector.com`
- **Trustwave:** `www.trustwave.com`
- **MyDL:** `www.mydlp.com`.
- **OpenDLP:** `code.google.com/p/opendlp`

My Endpoint Protector was chosen because it is a data loss prevention system that is deployed in the cloud; it engages the software as a service (SaaS) paradigm. A ten-day trial version of My Endpoint Protector with complete access to all the data loss prevention functionality was used in the study. Trustwave was chosen as a representative commercial data loss prevention system; access to its virtual test environment for the study was set up by Trustwave via a third party, Hatsize (`hatsize.com`). MyDLP and OpenDLP were chosen as representative open-source data loss prevention systems. Unfortunately, after closer examination, it turned out that OpenDLP was only able to operate on data at rest. Since the study focused on data in motion, extensive evaluations could not be performed for OpenDLP.

3.1 Agent-Based Solution

My Endpoint Protector is an agent-based, 100% cloud-managed data loss prevention solution. Everything is managed via a web user interface, from policy creation to downloading and deploying the agent at an endpoint. The trial version Endpoint Protector was used in the tests. It runs on a variety of platforms, including Windows 7 and 8, Mac OS

X, and Apple and Android tablets and phones. Tests were conducted on Windows 7 and Mac OS X Maverick to see if they differed in their protection services.

Upon using the Windows version of My Endpoint Protector, the following problems and suggestions for improvements were discerned:

Problems:

- An image was sent via Windows Live Mail, even if the action was to block images, because Live Mail uploads the image to Microsoft SkyDrive, which is not included in the list of checked applications for Windows.
- The data loss prevention agent disabled the print screen button, but was unable to disable the snipping tool for taking screenshots, which is pre-installed on Windows.
- The agent was unable to decompress files and analyze their contents, even if they contained sensitive data.
- Printing with a network printer was not detected by the agent.
- The agent could not detect if sensitive text was being written to a document or chat, only if the user was attempting to copy the text.
- When the connection between the agent and server was lost, and an incident occurred, the logs were not sent to the server after reconnection.

Suggested Improvements:

- Make it possible to add applications that should be monitored.
- Since the agent analyzes MIME-type files, the administrator should have the ability to add more file types as necessary.
- Save logs locally but encrypt them and send them to the server as soon as a connection is reestablished.
- Decompress compressed archived files and analyze their contents.

3.2 Agentless Solution

Trustwave is a commercial, agentless data loss prevention solution that performs content filtering and monitoring on dedicated servers in a network. In order for this to work, it is necessary to configure each

endpoint to route its traffic through the servers. To ensure that the traffic does not bypass the servers, it is recommended to prevent all outgoing SSL/HTTPS traffic (encrypted traffic) from leaving the network, with the exception of traffic from the data loss prevention servers. Ordinary non-encrypted traffic is covered by a monitoring server that receives copies of the outgoing traffic and logs instances of sensitive data leaving the network.

Access was provided to Trustwave's pre-installed virtual environment, which is used to demonstrate the product to new customers. The experimental setup comprised six servers and one Windows 7 desktop.

The virtual environment froze whenever the connection between data loss prevention components at the endpoint and at the server was cut by adding and activating rules in a Windows Firewall (as was done with the other data loss prevention systems). Unfortunately, this situation prevented further testing of Trustwave.

The following problems and suggestions for improvements were discerned for the Windows platform (because the pre-installed environment was used):

Problems:

- The uploading of `.zip` and `.rar` files was blocked, but not `.7z` files.
- Encrypted files containing sensitive data were not discovered.
- When attaching a document with sensitive data to an email, the only notification received was: "Something went wrong, could not attach document. Please try again."

Suggested Improvements:

- The Trustwave vendor was informed about the problem with `.7z` files; the problem will be fixed in a future version.
- Simplify the addition of file types from a scripting language to adding MIME types. The method for doing this should be clarified in the documentation.
- Present users with more information about the files that were blocked and why they were blocked.

3.3 Hybrid Solution

MyDLP is a combination of the two data loss prevention solutions discussed above. As in the case of Trustwave, content filtering is done

within the network using dedicated servers. However, MyDLP also enables data loss prevention agents to be installed at endpoints (these agents only run on Windows systems). Each endpoint must be configured to route its traffic through a data loss prevention server.

MyDLP is a open-source data loss prevention system with two licenses, Community Edition and Enterprise Edition. No support is available for the Community Edition.

MyDLP extracts compressed files before determining if they are sensitive or not; this is not done by My Endpoint Protector. MyDLP even extracts `.docx` files to check if anything is hidden in `.docx` folders. As in the case of Trustwave, it was not possible to force Skype and Teamviewer to go through the proxy.

Unlike the situation with My Endpoint Protector, when the agent-server is disabled and a file that is blocked is attempted to be transmitted, a log entry is sent to the server as soon as the connection is reestablished. MyDLP has built-in filesystem scan functionality called Discovery, but this has a very high impact on performance.

One advantage of MyDLP is that, even if the data loss prevention agent is disabled on a machine, the rules enforced via the proxy are still operational; however, the use of USB storage and local printers is no longer prevented. If the connection to the data loss prevention server is blocked, the last active operational policies are still enforced.

Problems:

- Attachments in web mail clients were blocked, but not attachments in local mail clients. Blocking was not implemented even after the proxy settings were set up.
- An email containing a `.docx` file with an image added in the extracted folder was detected, but not when the file was sent via Facebook.
- Only printing to a locally-connected printer can be blocked. Printing to a wirelessly-networked printer is not blocked.
- Notifications are not presented to users when something has been blocked. This could lead to frustration if the user is not aware of the information or the file type that was blocked.
- When applying a rule for blocking applications/octet-streams and MIME types for *inter alia* encrypted files, the entire web pages contained octet-streams, which made it impossible to test encrypted files as attachments.

- The screenshot rule was able to block the snipping tool in Windows, but not the print screen button. Other screenshot tools were not blocked.

Suggested Improvements:

- Extract `.docx` files for all types of file transfers.
- Present information about the reasons for blocking to users.
- Differentiate between file transfers and web browsing when it comes to octet-streams.
- Block all types of screenshot attempts.

3.4 File System Scanning Tool

OpenDLP is marketed as "a free and open source, agent- and agentless-based, centrally-managed, massively distributable data loss prevention tool" [5]. However, while this data loss prevention tool analyzes locally-stored data, it does not monitor network traffic or prevent data from being leaked. As such, OpenDLP is an example of what happens when the notion of a "proper data loss prevention solution" is not well-defined. By the time the scanner reaches a file containing sensitive information in order to quarantine it or block its dissemination, the file could be attached to an email or printed without being detected. The scanning tool does not affect performance as much as in the case of MyDLP, but because every scanning log entry is saved in a database, the performance could become noticeable if the database is overwhelmed with log entries.

This research focuses on data loss prevention systems that monitor data in motion. Since OpenDLP only analyzes data at rest, no tests were performed on the data loss prevention tool.

4. Experimental Setup

The virtual test environments incorporated the following operating systems:

- Windows 7 running Service Pack 1
- Linux running Ubuntu 13.10
- Mac OS X Maverick

The advantage of using a virtual environment is that it is fairly easy to create clones of entire experimental setups, which significantly reduces the time required for experimentation.

The four data loss prevention systems were installed and thoroughly tested one at a time on the different platforms. The experiments involved leaking sensitive data using webmail clients (Gmail), local mail clients (Windows Live Mail), cloud syncing software (Google Drive), social media (Facebook), instant messaging software (Skype), remote control software (Teamviewer), printers (network and local) and USB devices. The experiments also considered a situation where a user has different privileges on a local machine. The data that was attempted to be leaked included:

- A text file containing the secret words: "Hemlig," "Stopp," "Credit card information" and "Confidential." These were added to a custom content dictionary to see if the system could detect files containing the words. The credit card number: "4485630591171087" was also added to the text file to test the built-in credit card number finders.
- Compressed versions of the text file in the `.zip`, `.7z` and `.rar` formats.
- Encrypted version of the text file (stored in a Truecrypt file container).
- A 5,370 KB `.mp3` audio file.
- A 15 KB `.png` image file.
- Compressed version of the encrypted text file in the `.zip` format.
- A Microsoft Office Word `.docx` file; this file type is an archive that can be decompressed and files added to it without anyone being the wiser.

In every experiment, one file type at a time was set to be blocked with the remaining files being allowed; this was done to see if there was any way to fool the system. The experiments with the compressed files were done twice: the first experiment was set to block compressed file types while the second experiment allowed them, but blocked their contents (i.e., the text file). A similar set of experiments was done with the encrypted files. Experiments with the Microsoft Office file were also performed twice: the first was set to block the `.docx` file type and the second was set to allow it while blocking image files. The `.docx` file was also extracted and an image was added in the archive.

The experiments were performed in three phases for each data loss prevention system. During the first phase, a data loss prevention system was set up as recommended to prevent data leakage. Attempts

to move/leak the data were made using the applications and software mentioned above and the results were graded as Passed, Passed with Comments or Failed, depending on whether or not data leakage was prevented. Experiments that could not be conducted were graded as Comments while experiments for which a system did not cover the functionality were graded as N/A. The first phase also involved a study of how well each data loss prevention solution reported a blocked attempt to administrators and users.

The second phase involved attempts to disable or bypass a data loss prevention solution and then investigate how well it performed in a "crippled" state. The experiments were run again and graded using the same scores as before. The following attacks were implemented on the data loss prevention solutions:

- Booting the computer from another media (i.e., CD) and accessing the hard drive (i.e., a live CD attack).
- Corrupting or manipulating important binaries or configuration files used by a data loss prevention solution.
- Stopping or killing a data loss prevention solution to prevent it from executing.

The third phase involved disabling communications between the central server and the endpoint. To accomplish this, rules were configured in the local firewall at the endpoint to prevent communications between the central server and the endpoint. The same experiments as in the two previous phases were then conducted to evaluate the effects with respect to data loss prevention.

Table 1 presents the test cases and configurations used in the experiments.

5. Experimental Results

This section summarizes the experimental results obtained for the three phases. The data loss prevention systems tested were: (i) My Endpoint Protector (MEP); (ii) Trustwave (TW); and (iii) MyDLP (MD). As mentioned above, the following grades were used to assess data loss prevention performance:

- **Passed (P):** The system successfully blocked the leakage attempt.
- **Passed with Comments (PC):** The system was partially successful at blocking or identifying the leakage attempt, but this did not qualify as a pass and comments are provided.

Table 1. Test cases.

Test Case	Configuration
Transfer a text file	Policy to block text files is active
Transfer a text file with sensitive data	Policy to block text files is inactive; policy to block sensitive data is active
Transfer a compressed file	Policy to block compressed files is active
Transfer a compressed file with sensitive data	Policy to block compressed files is inactive; policy to block sensitive data is active
Transfer an encrypted file	Policy to block encrypted files is active
Transfer an encrypted file with sensitive data	Policy to block encrypted files is inactive; policy to block sensitive data is active
Transfer an encrypted file	Policy to block encrypted files is active
Transfer an encrypted file with sensitive data	Policy to block encrypted files is inactive; policy to block sensitive data is active
Transfer a media file	Policy to block media files is active
Transfer an image	Policy to block image files is active
Write sensitive data	Type blocked words from content dictionary
Transfer a compressed file containing an encrypted file	Compress an encrypted file while the policy to block encrypted files is active; policy to block compressed files is inactive
Transfer a Microsoft Office file	Policy to block Microsoft Office files is inactive; policy to block image files is active

- **Failed (F):** The system did not block the leakage attempt.

- **Comments (C):** The test case was not conducted for some reason.

- **Not Applicable (N/A):** The system functionality did not cover the test case.

Phase 1: Clean Install. The experiments involved installing the data loss prevention software at the endpoints, setting up policies according to the documents and manuals, and transferring the test files. The data loss prevention software was not tampered with during the first phase. Figure 1 shows the detailed results of the Phase 1 experiments. Note that the systems tested were: My Endpoint Protector (MEP); (ii) Trustwave (TW); and (iii) MyDLP (MD). Table 2 summarizes the main results.

Description	Mail			Printer			Google Drive			Facebook			Skype			Teamviewer			USB		
	MEP	TW	MD	MEP	TW	MD	MEP	TW	MD	MEP	TW	MD	MEP	TW	MD	MEP	TW	MD	MEP	TW	MD
Attach/send text file	P	N/A	PC	P	N/A	F	PC	N/A	F	P	N/A	P	P	N/A	C	P	N/A	C	P	N/A	P
Attach/send text file with sensitive data	P	P	PC	P	N/A	PC	P	C	F	P	C	P	P	C	C	P	C	C	P	N/A	P
Attach/send compressed file	PC	N/A	PC	N/A	N/A	N/A	PC	N/A	F	P	N/A	P	P	N/A	C	P	N/A	C	P	N/A	P
Attach/send compressed file with sensitive data	F	PC	PC	N/A	N/A	N/A	F	C	F	F	C	P	F	C	C	F	C	C	F	N/A	P
Attach/send encrypted file	P	N/A	F	N/A	N/A	N/A	PC	N/A	F	P	N/A	F	P	N/A	C	P	N/A	C	P	N/A	F
Attach/send encrypted file with sensitive data	F	F	F	N/A	N/A	N/A	F	C	F	F	C	F	F	C	C	F	C	C	F	N/A	F
Attach/send media file	P	N/A	PC	N/A	N/A	N/A	PC	N/A	F	P	N/A	P	P	N/A	C	P	N/A	C	P	N/A	P
Attach/send image	PC	N/A	PC	PC	N/A	F	PC	N/A	F	P	N/A	P	P	N/A	C	P	N/A	C	P	N/A	P
Write sensitive data	F	P	PC	N/A	N/A	N/A	F	C	F	F	C	P	F	C	C	N/A	N/A	N/A	N/A	N/A	N/A
Attach/send compressed file with encrypted file within.	F	F	F	N/A	N/A	N/A	F	C	F	F	C	F	F	C	C	F	C	C	F	N/A	F
Attach/send Microsoft Office files	P	N/A	PC	P	N/A	F	PC	N/A	F	P	N/A	PC	P	N/A	C	P	N/A	C	P	N/A	P

Figure 1. Test results for a clean install.

Description	Mail			Printer			Google Drive			Facebook			Skype			Teamviewer			USB		
	MEP	TW	MD	MEP	TW	MD	MEP	TW	MD	MEP	TW	MD	MEP	TW	MD	MEP	TW	MD	MEP	TW	MD
Attach/send text file	F	N/A	PC	F	N/A	F	F	N/A	F	F	N/A	P	F	N/A	C	F	N/A	C	F	N/A	F
Attach/send text file with sensitive data	F	P	PC	F	N/A	F	F	C	F	F	C	P	F	C	C	F	C	C	F	N/A	F
Attach/send compressed file	F	N/A	PC	N/A	N/A	N/A	F	N/A	F	F	N/A	P	F	N/A	C	F	N/A	C	F	N/A	F
Attach/send compressed file with sensitive data	F	PC	PC	N/A	N/A	N/A	F	C	F	F	C	P	F	C	C	F	C	C	F	N/A	F
Attach/send encrypted file	F	N/A	F	N/A	N/A	N/A	F	N/A	F	F	N/A	F	F	N/A	C	F	N/A	C	F	N/A	F
Attach/send encrypted file with sensitive data	F	F	F	N/A	N/A	N/A	F	C	F	F	C	F	F	C	C	F	C	C	F	N/A	F
Attach/send media file	F	N/A	PC	N/A	N/A	N/A	F	N/A	F	F	N/A	P	F	N/A	C	F	N/A	C	F	N/A	F
Attach/send image	F	N/A	PC	F	N/A	F	F	N/A	F	F	N/A	P	F	N/A	C	F	N/A	C	F	N/A	F
Write sensitive data	F	P	PC	N/A	N/A	N/A	F	C	F	F	C	P	F	C	C	N/A	N/A	N/A	N/A	N/A	N/A
Attach/send compressed file with encrypted file within.	F	F	F	N/A	N/A	N/A	F	C	F	F	C	F	F	C	C	F	C	C	F	N/A	F
Attach/send Microsoft Office files	F	N/A	PC	F	N/A	F	F	N/A	F	F	N/A	PC	F	N/A	C	F	N/A	C	F	N/A	F

Figure 2. Test results for a disabled data loss prevention agent.

Table 2. Summary of results from Figure 1.

	My Endpoint Protector	**Trustwave**	**MyDLP**
Passed	37	2	14
Passed with Comments	8	1	10
Failed	22	2	23
Comments	0	19	21
N/A	9	53	9
Total	77	77	77

Table 3. Summary of results from Figure 2.

	My Endpoint Protector	**Trustwave**	**MyDLP**
Passed	0	2	14
Passed with Comments	0	1	10
Failed	68	2	23
Comments	0	19	21
N/A	9	53	9
Total	77	77	77

Comments:

- **Trustwave:** SSL/HTTPS could not be tested because of configuration problems. Also, Trustwave does not cover endpoints, so the printer and USB received N/A scores.
- **MyDLP:** Skype and Teamviewer could not be forced to go through the proxy.

Phase 2: Disabled Agent. The experiments involved attempts to disable data loss prevention software at the endpoints (described previously). Figure 2 shows the detailed results of the Phase 2 experiments. Table 3 summarizes the main results.

Comments:

- **Trustwave:** Trustwave is an agentless data loss prevention system, so there is no agent to disable. Also, the comments are the same as in Phase 1.
- **MyDLP:** The comments are the same as in Phase 1.

Table 4. Summary of results from Figure 3.

	My Endpoint Protector	Trustwave	MyDLP
Passed	34	0	7
Pass with Comments	8	0	1
Failed	26	0	39
Comments	0	24	21
N/A	9	53	9
Total	77	77	77

Phase 3: Disconnected Server. The experiments involved disabling communications between the central server and the endpoint. Figure 3 shows the detailed results of the Phase 3 experiments. Table 4 summarizes the main results.

Comments:

- **Trustwave:** The virtual environment halted when a firewall rule was added and activated. Also, the comments are the same as in Phase 1.
- **MyDLP:** The comments are the same as in Phase 1.

6. Discussion

The experimental results demonstrate that the data loss prevention systems have severe "blind spots" in some test cases. This is understandable given that they implement different strategies based on host operation or network operation. Even so, some of the systems draw complete blanks for certain test cases, which means that they can hardly be relied on in operational environments. Also, it is not unreasonable to assume that some users could inadvertently identify these flaws and leverage them to transfer and store sensitive files. This, of course, makes the situation trickier for a forensic investigator because any evidence of leakage could be argued to be the result of an accident instead of malicious behavior.

The evaluation results with regard to disabling the data loss prevention systems are even more disheartening. The systems were relatively easy to disable or fool (even by non-expert users), which renders their effectiveness more questionable, especially as standalone solutions. It also appears that attempts at disabling the systems could go unnoticed. For example, stopping a data loss prevention agent from contacting a server by activating firewall rules would leave few traces on the server.

Description	Mail			Printer			Google Drive			Facebook			Skype			Teamviewer			USB		
	MEP	TW	MD	MEP	TW	MD	MEP	TW	MD	MEP	TW	MD	MEP	TW	MD	MEP	TW	MD	MEP	TW	MD
Attach/send text file	P	N/A	F	F	N/A	F	PC	N/A	F	P	N/A	F	P	N/A	C	P	N/A	C	P	N/A	P
Attach/send text file with sensitive data	P	C	F	F	N/A	PC	P	C	F	P	C	F	P	C	C	P	C	C	P	N/A	P
Attach/send compressed file	PC	N/A	F	N/A	N/A	N/A	PC	N/A	F	P	N/A	F	P	N/A	C	P	N/A	C	P	N/A	P
Attach/send compressed file with sensitive data	F	C	F	N/A	N/A	N/A	F	C	F	F	C	F	F	C	C	F	C	C	F	N/A	P
Attach/send encrypted file	P	N/A	F	N/A	N/A	N/A	PC	N/A	F	P	N/A	F	P	N/A	C	P	N/A	C	P	N/A	F
Attach/send encrypted file with sensitive data	F	C	F	N/A	N/A	N/A	F	C	F	F	C	F	F	C	C	F	C	C	F	N/A	F
Attach/send media file	P	N/A	F	N/A	N/A	N/A	PC	N/A	F	P	N/A	F	P	N/A	C	P	N/A	C	P	N/A	P
Attach/send image	PC	N/A	F	F	N/A	F	PC	N/A	F	P	N/A	F	P	N/A	C	P	N/A	C	P	N/A	P
Write sensitive data	F	C	F	N/A	N/A	N/A	F	C	F	F	C	F	F	C	C	N/A	N/A	N/A	N/A	N/A	N/A
Attach/send compressed file with encrypted file within.	F	C	F	N/A	N/A	N/A	F	C	F	F	C	F	F	C	C	F	C	C	F	N/A	F
Attach/send Microsoft Office files	P	N/A	F	F	N/A	F	PC	N/A	F	P	N/A	F	P	N/A	C	P	N/A	C	P	N/A	P

Figure 3. Test results for a disconnected data loss prevention server.

As a result, checking for information that is not in the logs becomes as important as checking what is actually in the logs.

In summary, the data loss prevention field is currently immature. Consequently, it is imprudent to place too much trust on data loss prevention solutions when attempting to prevent data leakage as well as when conducting forensic investigations.

7. Conclusions

This chapter describes the evaluation of four data loss prevention systems in order to ascertain their effectiveness at stopping the unauthorized exposure of sensitive data and the ease with which the systems could be circumvented. Although the focus was on agent-based systems that do not have connections to data loss prevention servers, agentless systems were considered as well. The experimental results reveal that none of the data loss prevention systems is 100% secure and all have potential weaknesses and lack complete coverage when identifying and protecting data. In general, the evaluated systems perform rather poorly and several improvements have been suggested to address the problems.

The research has identified various threats to the data loss prevention agents themselves; the measures to protect agents include restricting user privileges, using full disk encryption and monitoring process status. Another key issue is to create policies that cover the proper identification and categorization of data, especially when the data is hidden or obfuscated.

Future research will investigate the forensic aspects of data loss prevention systems, especially the traces of data leakage that remain and those that are not retained. Also, additional applications will be used to reveal more information about data loss prevention systems and their weaknesses. Finally, improved agent-based and agentless solutions will be developed to prevent data leakage from devices such as network printers.

References

[1] H. Balinsky, D. Perez and S. Simske, System call interception framework for data leak prevention, *Proceedings of the Fifteenth IEEE International Conference on Enterprise Distributed Object Computing*, pp. 139–148, 2011.

[2] J. Blasco, J. Hernandez-Castro, J. Tapiador and A. Ribagorda, Bypassing information leakage protection with trusted applications, *Computers and Security*, vol. 31(4), pp. 557–568, 2012.

[3] V. Carvalho and W. Cohen, Preventing information leaks in email, *Proceedings of the SIAM International Conference on Data Mining*, pp. 68–77, 2007.

[4] Computer Security Institute, 2010/2011 Computer Crime and Security Survey, New York, 2011.

[5] A. Gavin, OpenDLP – Data Loss Prevention Suite, version 0.5.1 (`code.google.com/p/opendlp`), 2012.

[6] P. Kanagasingham, Data Loss Prevention, InfoSec Reading Room, SANS Institute, Bethesda, Maryland, 2008.

[7] J. Kim and H. Kim, Design of internal information leakage detection system considering the privacy violation, *Proceedings of the International Conference on Information and Communication Technology Convergence*, pp. 480–481, 2010.

[8] M. Luft, Can Data Leakage Prevention Prevent Data Leakage? Bachelor's Thesis, Laboratory for Dependable Distributed Systems, University of Mannheim, Mannheim, Germany, 2009.

[9] R. Mogull, Understanding and Selecting a Data Loss Prevention Solution, Technical Report, Securosis, Phoenix, Arizona, 2010.

[10] E. Ouellet, Magic Quadrant for Content-Aware Data Loss Prevention, Technical Report G00224160, Gartner, Stamford, Connecticut, 2013.

[11] T. Wuchner and A. Pretschner, Data loss prevention based on data-driven usage control, *Proceedings of the Twenty-Third IEEE International Symposium on Software Reliability Engineering*, pp. 151–160, 2012.