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-