State of the Art in Network-Related Extrusion Prevention Systems

Andreas Hackl, Barbara Hauer

4020 Linz, Austria
andreas_hackl@utanet.at, hauer@grz.at

08. January 2009

Abstract. In recent years organizations realized the value of intellectual property. Common security systems are able to protect the data against threats from outside, but approved researches showed that in most cases the danger arises from inside an organization. Under the term “Extrusion Prevention” new concepts were developed to meet the requirements against internal and external threats. This paper gives an overview about state of the art in network-related extrusion prevention systems.

1 Introduction

Extrusion Prevention, also referred to as Data Loss Prevention, Data Leakage Prevention or Content Monitoring and Filtering, helps companies to stop leakage of sensitive data through unauthorized disclosure by insiders [1].
Looking at current threats, unwanted discharge of information becomes more important because often companies have to pay for the consequences of loosing sensitive data, e.g. damage mitigation and victim remediation [3]. There are a lot of ways data can leak out of an organization like sensitive data sent via e-mail or instant messenger, transferred via insecure transmission or stored on lost or stolen devices without encryption [5].
Extrusion prevention systems take measures to identify, monitor, analyze and protect sensitive data in different states [2] by a content or context based approach.

Data in Motion. Describes data which is sent across specific communication channels over the network, e.g. via e-mail, web mail, instant messaging, a web based tool [2, 3] or a peer-to-peer application. An extrusion prevention system is monitoring the network traffic to identify sensitive data leaving the network [3].

Data at Rest. Protecting data at rest through extrusion prevention systems includes scanning the network and various sources like databases, document management systems and other content repositories for sensitive materials [2, 3].

Data in Use. Is a term for data a user interacts with [2]. An extrusion prevention system uses a software agent to monitor the data on and the data transport from an endpoint device/client via different output channels to peripherals [3]. E.g., a document with sensitive content is transferred to an USB device.

1.1 System Types

There are two types of extrusion prevention systems.

Network-Related Systems. See section 2.

Endpoint-Related Systems. These systems focus on data in use. Therefore, software agents are used which run directly on endpoint devices/clients. They monitor and control the access to data and the related server is responsible for administrative duties, policy distribution and generation of log events. Only a few solutions provide additional functionality depending on some conditions [4].

2 Network-Related Systems

These systems, also referred to as “Server-Based Architecture”, analyze the network traffic and search for unauthorized data transmissions. Therefore, they need one or more sensors in a network domain. Typically, network-related systems require proxies for every application protocol to monitor the specific traffic. Such dedicated proxies provide the ability for an efficient monitoring. Related data in motion will be analyzed by a store-and-forward principle [4], based on the defined policies. The analysis results in forwarding, throwing an event or blocking.
Current products provide protective and detective controls and include advanced capabilities to integrate existing network infrastructure.
Network-related systems consist of several system parts (cf. [2])
Network monitor. A passive network monitor is typically at or near the gateway on a SPAN port. The monitor is responsible for full packet capture, session reconstruction and analyzing the content of the network traffic in real time. A network monitor is often dedicated server hardware with an extrusion prevention software running on. The functionality for management, work flow and reporting is sometimes built in the network monitor, but more often offloaded to a separate server or appliance.
E-mail Integration. A lot of capabilities, like quarantine, encryption integration and filtering are gained by e-mail integration. Since most products embed a Mail Transport Agent (MTA), only a few of them support some major MTA security solutions.
Filtering/Blocking and Proxy Integration. Most of the traffic in a network is synchronous and runs in real time. Therefore, if an extrusion prevention system wants to filter or even block the traffic in the network it uses a bridge, proxy or TCP poisoning. A bridge acts like a man in the middle for data analysis by using two network cards. A proxy is queuing up traffic for analysis. Gateway proxies are

typically used for HTTP, FTP and instant messaging protocols. When TCP poisoning is applied a TCP reset packet is injected to destroy the connection.

An extrusion prevention system should be able to support multiple monitoring points like network monitors, proxies and e-mail servers. A central management server should then collect the events of all monitoring points for workflow, reporting, investigation and archiving [2].

2.1 Concepts and Algorithms

There are two categories of analyses

- the context analysis and
- the content analysis.

Each type uses a variety of techniques. Not every technique is used in every product, because of that it is important to identify the requirements in order to find the appropriate solution.
Context analysis deals with source IP address and port, destination IP address and port, size, header information, metadata, time, format and in more advanced versions with context related to business processes and therefore also to content, which includes the environment and the use of the content at the time of analysis.
Since context analysis is based on many years of experience from previous network security technologies like firewalls, proxies, intrusion detection systems (IDS) and intrusion preventions systems (IPS), content analysis depends on enhanced and new approaches and turned out to be more complex.

As explained in [2], Securosis identified seven major types of content analysis techniques:

- Rule-Based/Regular Expressions
- Database Fingerprinting
- Exact File Matching
- Partial Document Matching
- Statistical Analysis
- Conceptual/Lexicon
- Categories

In order to define complex policies it is possible to combine and chain analysis techniques but this depends on the features of the extrusion prevention product.

2.2 Advantages/Disadvantages

The main advantage of network-related systems is that they do not need an installation on endpoints. This reduces the installation complexity and prevents conflicts with other software like virus scanners [4].

Main disadvantages in comparison to endpoint-based systems are the transfer of sensitive data from an endpoint device/client to a physical device which cannot be monitored [4] and encrypted data which cannot be analyzed.
Furthermore, used search patterns must have a high quality to avoid false positives.

2.3 Server-based Extrusion Prevention in comparison with established Security Solutions

Common security concepts are limited to the protection against industrial espionage and attacks from outside the organizations but delinquents are not only strangers, but also and more often dissatisfied or negligent employees as approved researches showed. In view of the fact that security solutions like firewalls, intrusion detection systems or intrusion prevention systems are not meant to provide protection against internal threats, it becomes clear that these systems do not prevent the discharge of sensitive data.
Besides, common security solutions concentrate on technical aspects, while "Extrusion Prevention" or "Data Leakage Prevention (DLP)" respectively is a concept with technical and organizational measures. This way of seeing DLP is not shared by all vendors which use the term to sale their products. Therefore, this market is split into DLP as a feature and DLP as a complete solution (cf. [2] p. 6). DLP features only provide some basic functions for detection and enforcement but are not enough for protection of data content, while DLP products try to solve business and technical problems through a comprehensive approach and include a centralized management and awareness. Because of that, network-based extrusion prevention must be seen as DLP feature because only a part of possible ways for input and output of data can be controlled. Besides, a DLP feature also can integrate existing security solutions like the following

2.3.1 Network Layer Firewall
This type of firewall is also called a “packet filter firewall”. The firewall decides on the basis of defined rules if a network packet is allowed to pass through or not. Filter criteria are packet attributes like the source IP address, the destination IP address, the source port or the destination service, or in modern firewalls also the protocol, the TTL value, the domain name and other attributes available in the packet.
Common network layer firewalls filter based on context information only.

2.3.2 Application Layer Firewall (ALF)
An ALF intercepts and inspects the packets to and from an application or Web Service. It can analyze the included data and drop unwanted traffic. This kind of firewall can be used for input and output validation.
An ALF is not able to filter all network traffic. Furthermore, it is limited to one or a small number of applications or Web Services. Mostly, an ALF is focused on a special application layer protocol but this protocol is filtered with content analysis techniques e.g. input validation based on regular expressions.

2.3.3 Proxy

Proxies can act as a firewall and additionally can hide the clients IP address. ALFs use “dedicated Proxies” to supervise application layer protocols, which means to have a dedicated proxy for every application layer protocol. Many vendors of extrusion prevention systems integrate customer’s existing gateway proxies into their solutions. If a reverse SSL proxy is integrated, SSL connections can be sniffed. This is applicable for all encrypted data and traffic since the proxy can intercept, decrypt, initiate analyzing of the content and encrypt for further routing, provided that the necessary private keys and certificates are available.

2.3.4 Intrusion Detection System (IDS) and Intrusion Prevention System (IPS)

Often integrated into firewalls, they are used to monitor and analyze occurring events in order to identify an incident, e.g. possible attack, by comparison with predefined traffic patterns. IPS differs from IDS as it can not only identify and alert suspicious traffic but also block packets [15].
This kind of technology is restricted to identification of suspicious activity, mostly based on monitoring network traffic. Content analysis is neither applied nor provided.

3 Weakness analysis

This section focuses on data in motion only and presents two probable weaknesses or attack approaches. Further scenarios would go beyond the scope of this paper.

If a company allows encryption or encrypted connections without the possibility to decrypt the data and have a look at the content, it has no chance to detect leakage of sensitive data. Besides this fact, a lot has to be considered like the topics listed below.

3.1 Transport of Application Protocols across unusual Channels

In case the employees of a company are not allowed to use the internet, the used firewall or proxy usually blocks TCP/IP connections. But most of the companies allow ICMP requests from their network to go outside and even worse, they do not analyze the content of the ICMP packets. Because of that it is possible to transfer HTTP or SSH hidden [14]. This works because ICMP is part of the Internet Protocol Suite and the data is sent in the payload of the ICMP packets. In the majority of cases, the client inside the company network will send ICMP request packets, while the used proxy outside the network will answer with ICMP reply packets. In early attacks, ICMP tunneling was used to execute remote commands and get information from a compromised machine. These days this kind of tunneling still serves the purpose of smuggling data through a firewall, with the difference that it starts very often from the inside of the organizations.
However, security systems can be configured to prevent these data leakage (see [13]).

3.2 Data Privacy Laws

In the last few months a lot of articles were written dealing with the subject “Data Leakage Prevention” or synonyms like extrusion prevention. They all concentrate more or less detailed on technology issues. But only a fistful address the issue of data privacy laws and even if they do, they only state out, that private data of possible victims has to be protected. But what about the legal protection of the private data of a conventional employee who works within a company secured by an extrusion prevention product or what about the private data of a “bad person”? They can also claim the protection of law and at this point it starts to get interesting but complicated too. In fact some vendors deleted extrusion prevention for e-mails from the feature list of their products because of the variable and often very strict regulations within different countries.
In Austria the legal situation is substantiated by the “Datenschutzgesetz 2000 (DSG 2000 [10])”, the European Convention on Human Rights (ECHR [11]) and the “Arbeitsverfassungsgesetz (ArbVG [10])”. According to Art. 1 § 1 DSG (Basic right on data protection) and Art. 8 ECHR (Right to respect for private and family life) it is not allowed to read private communication data like private e-mails of a person, even if the private use of e-mail and Internet is not allowed due to company restrictions. The company is allowed to reprimand an employee without having a look into the content of his communication (cf. [12]). The problem with e-mails is that an operational system cannot decide if it is a private or a business e-mail. Furthermore, the content cannot be analyzed without saving (hard disk or memory) and opening the e-mail, and at the latest when the e-mail is opened, it is relatable to a person. Because of that, there is no legal way for an analysis of data transferred through an e-mail.
Beyond e-mail, referring to § 96 ArbVG even the insertion of arrangements for inspection of employees is not allowed if the action is concerning human dignity, irrespective of the works council consent, and this concerns all sensors.
Network-related extrusion prevention systems make sense anyway because a company is obligated to protect sensitive data of their customers and employees. Furthermore it is allowed to monitor and analyze the network traffic if it cannot be associated to a person.

4 State of the Art Products

The market research company Forrester Research, Inc. evaluated data leakage prevention products of leading vendors in 2008 (see [6]). As result of the evaluation, the vendors Websense, Reconnex and Verdasys were ranked on top, achieving best on the 74 evaluation criteria.

Websense. Websense offers a product called “Websense Data Security Suite”. The Data Security Suite protects data in all states. The analysis engine is language agnostic and includes templates for many regions. To identify structured and unstructured data a special fingerprinting technology is used. Multiple data fields can be correlated to define the data to protect [6]. The Websense Data Security Suite

consists of four system components, the Websense Data Discover, Websense Data Monitor, Websense Data Protect and Websense Data Endpoint [8].
Sensitive data at rest is discovered and classified by an agent-less network solution, called Data Discover. The Data Monitor provides the ability for passive monitoring of business communication within a company. The Data Protect extends the functionality of the Data Monitor and provides built-in, automated policy controls and real-time reporting. The fourth system component is the Data Endpoint for discovering, monitoring and protecting sensitive data at the endpoint.

Reconnex. This company offers an extrusion prevention system which protects sensitive data in motion, at rest and in use. The Reconnex solution consists of two system components, the “iGuard Appliance” and the “inSight Console” [7].
The iGuard Appliance is responsible for information monitoring and protection. The inSight Console handles management of device configurations, reporting, incidents, policies, rules and case management [7]. Reconnex provides port agnostic application classification for data at rest, offers appliance based network capabilities to protect data in motion and an endpoint agent to monitor data in use [6].

Verdasys. Verdasys “Digital Guardian” discovers, monitors, classifies and controls confidential data by context and content. The framework consists of several system components for file, e-mail and full disk encryption, monitoring data at the point of use. Centralized management for policies, data, classification, settings and reporting supports in customization and use through flexible configuration [6, 9].

5 Conclusion

The main challenge in extrusion prevention is the definition and the adaptation of plenty of technical and organizational measures exacerbated by a lack of standards and norms. Furthermore, the systems have to protect the data without disrupting the business process.
Network-related extrusion prevention systems are not able to deal with all potential ways of data leakage, but they are the first step in protection of sensitive data and intellectual property because they address one of the most significant vectors for data loss: electronic communication.
Since extrusion prevention is in its early stages, the technology faces many challenges to meet the requirements of the organizations. At the moment available extrusion prevention systems can avoid accidental leakage of data but they are not able to provide full protection of intellectual property against a targeted attack.

References

1. Abbadi, I.M., Alawneh, M.: Preventing Insider Information Leakage for Enterprises. In: 2008 International Conference on Emerging Security Information, Systems and Technologies, pp. 99--106 (2008)
2. Mogull, R.: Understanding and Selecting a Data Loss Prevention Solution. Securosis, The SANS Institute, Whitepaper (2008)
3. Lawton, L.: New Technology Prevents Data Leakage. Computer, vol. 41, no. 9, pp. 14--17 (2008)
4. Schmidt, M. A.: Ausgangskontrolle - Funktionsweise und Grenzen von Data Leakage Prevention. <kes> - Die Zeitschrift für Informations-Sicherheit. vol. 4, pp. 12 - 16, SecuMedia Verlags-GmbH, Ingelheim (2008)
5. Filkins, B., Radcliff D.: Data Leakage Landscape: Where Data Leaks and How Next Generation Tools Apply. The SANS Institute, Whitepaper (2008)
6. Raschke, T.: The Forrester Wave™: Data Leak Prevention, Q2 2008. Forrester Research, Inc.. Research Document (2008)
7. Reconnex, Inc.: Product overview, http://www.reconnex.net/products/index.php
8. Websense, Inc.: Data Security, http://www.websense.com/content/DataSecurity.aspx
9. Verdasys, Inc.: Digital Guardian Overview, http://www.verdasys.com/data_loss_prevention.php
10. Rechtsinformationssystem des Bundes (RIS), http://www.ris2.bka.gv.at/default.aspx
11. European Convention on Human Rights and Additional Protocols, European Court of Human Rights, Strasbourg-Cedex, http://www.echr.coe.int/echr/
12. Kammer für Arbeiter und Angestellte für Wien als Büro der Bundesarbeitskammer: Private E-Mail-Nutzung am Arbeitsplatz. http://www.arbeiterkammer.at/online/private-emails-im-buero-25261.html
13. Singh, A., Lu, C., Nordstrom, O., Santos, A.: ICMP Tunneling: Defense Against the Vulnerability. Georgia Tech. Information Security Center (GTISC), Center for Experimental Research in Computer Systems(CERCS), College of Computing, Georgia Institute of Technology, Paper, http://www.2factor.us/icmp.pdf
14. Stödle, D.: Ping Tunnel - For those times when everything else is blocked. Tutorial (2008)
15. Scarfone, K., Mell P.: Guide to Intrusion Detection and Prevention Systems (IDPS). National Institute of Standards and Technology, Special Publication (2007)