www.iss.net

IBM Internet Security Systems Executive Brief

IBM Internet Security Systems X-Force® 2006 Trend Statistics

January 2007

IBM Internet Security Systems
Ahead of the threat.™

Table of Contents

Management Overview

2006 was a record year on many security fronts. IBM Internet Security Systems X-Force® research and development team closely observed and recorded new vulnerabilities and the status of varying threats throughout the year. This data has been compiled in this report.

2006 End-of-the-Year Highlights

Vulnerabilities

- There were a total of 7,247 vulnerabilities in 2006, which represents a 39.5 percent increase over 2005.
- June was the busiest month of the year with 696 vulnerabilities.
- Week 46 (the week before Thanksgiving) was the busiest week of 2006 for new vulnerabilities.
- The most popular day for vulnerability disclosures was Tuesday.
- Weekend disclosure of vulnerabilities in 2006 more than doubled that of 2005 to reach 17.6 percent of all disclosures.
- "High impact" vulnerabilities continue to decrease as a percentage of total vulnerabilities in 2006.
- 3 percent of vulnerabilities under the Common Vulnerability Scoring System (CVSS) were evaluated as being "critical impact" vulnerabilities with a score of 10.
- The top three vulnerable vendors in 2006 were Microsoft, Oracle and Apple.
- The top 10 vulnerable software vendors accounted for 14 percent of all 2006 vulnerabilities.
- 17 percent of the vulnerabilities identified within the top 10 vulnerable vendors' products were un-patched at the end of 2006. This contrasts with 65 percent un-patched for all other vulnerabilities recorded in the year.
- 88.4 percent of all 2006 vulnerabilities could be exploited remotely.
- Over half (50.6 percent) of 2006 vulnerabilities would allow an attacker to gain access to the host after successful exploitation.

Spam and Phishing

- The U.S., Spain and France are the three largest originators of spam worldwide.
- USA and China each host over 1/3rd of the world's destination websites sent in spam messages.
- More than 90 percent of spam messages now use HTML to present message content.
- More than 60 percent of spam messages are sent directly to the recipient's mail server – without passing through any intermediary relay agents.
- 92.99 percent of spam messages are written in English, with German being the next most popular language.
- South Korea accounts for the highest source of phishing e-mails – 16.33 percent.
- More than half (55.78 percent) of the world's phishing attacks have fake Web sites hosted in the U.S.
- U.S. based businesses are the most targeted organizations of phishing e-mails, accounting for 71.37 percent of all phishing e-mail.
- More than 95 percent of phishing e-mails rely upon HTML delivery.
- Image-based spam has increased linearly since 2005, and accounted for more than 40 percent of spam messages by the end of 2006.

Web Content

- 12.5 percent of Internet Web sites host “unwanted” content such as pornography, violence and crime, etc.
- Web sites that host pornographic or sex-related content account for 12.03 percent of the Internet.
- "Unwanted" content has risen by between 9-14 percent in 2006 (depending on web content)
- The U.S. is the top hosting country for “unwanted” content such as violence and crime, pornography and sex, computer crime, and illegal drugs.

Malcode

- The largest threat category of malware in 2006 were Downloaders (68,620 varieties in 2006) – accounting for 22 percent of all malware.
- The most frequently occurring malware on the Internet was Trojan-Downloader.Win32.Zlob.
- The most common worm in 2006 was Email-Worm.Win32.Luder, and the most successful family of network propagating worms was New-Worm.Win32.Mytob.

Web Browser Exploitation

- The most popular exploit used on the Internet to infect Web browsers with malware was Microsoft’s MS-ITS vulnerability (MS04-013).
- Approximately 50 percent of Web sites hosting exploit material designed to infect Web browsers now obfuscate their attack, with approximately 30 percent encrypting their payload.

Vulnerability Analysis

The IBM ISS X-Force has been cataloguing, analyzing and researching vulnerability disclosures since 1997. The X-Force database is the largest, most authoritative database in the world, with more than 30,000 security vulnerabilities catalogued. This unique database enables X-Force researchers to understand the dynamics that make up vulnerability discovery and disclosure.

In fact, X-Force researchers have analyzed many more “disclosures” than the 30,000+ recorded in the X-Force Vulnerability Database. On average each year, a sizable proportion of public disclosures are incorrect and are not recorded in the database. These “disclosures” are rejected because they are either re-discoveries of existing and older vulnerabilities, or (after careful research by X-Force) found to be merely software bugs, have no vulnerability context and are closer to audit-level notifications.

The exponential increase of vulnerabilities in 2006 over all previous years shattered many records. X-Force researchers catalogued, tracked and researched 7,247 vulnerabilities – an average of 20 vulnerabilities per day.

The next section covers the following areas of analysis:

- Per annum vulnerability count
- Vulnerabilities per month
- Vulnerabilities per week
- Vulnerabilities by day of week
- Weekday vs. weekend

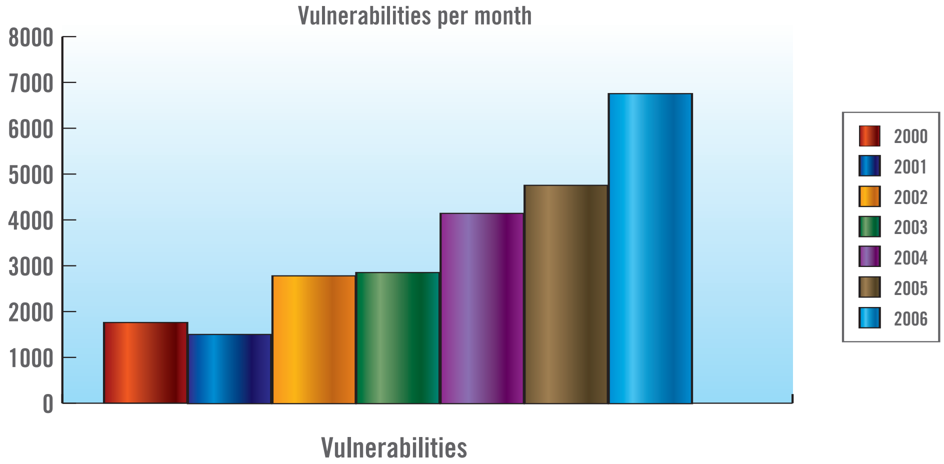
- Classic high/medium/low vulnerability impact breakdown
- Common Vulnerability Scoring System (CVSS) breakdown
- Top 10 vulnerable vendors
- Remote vs. local exploitation
- Consequences of exploitation

Per Annum Vulnerability Count

With 7,247 vulnerabilities disclosed in 2006, total vulnerability count increased nearly 40 percent over the previous year. Since the turn of the millennium, there has been a 261 percent increase in vulnerabilities, an average of 23 percent per annum. This trend is expected to continue throughout 2007.

While new operating systems such as Microsoft Vista provide more security functions and have undergone extensive security reviews and audits, their complexity has increased proportionally with the number of lines of new code. Although security has improved – and the likelihood of critical new security issues has decreased – the total number of new vulnerabilities likely to be uncovered with the introduction of new operating systems in 2007 is estimated to top 2006 totals. In addition, third-party software vendors typically release major updates to their products in conjunction with the introduction of a major new operating system. The parallel release of new applications will likely contribute to a record year for vulnerabilities in 2007.

The year-on-year increase in vulnerabilities can be observed in the following graph:

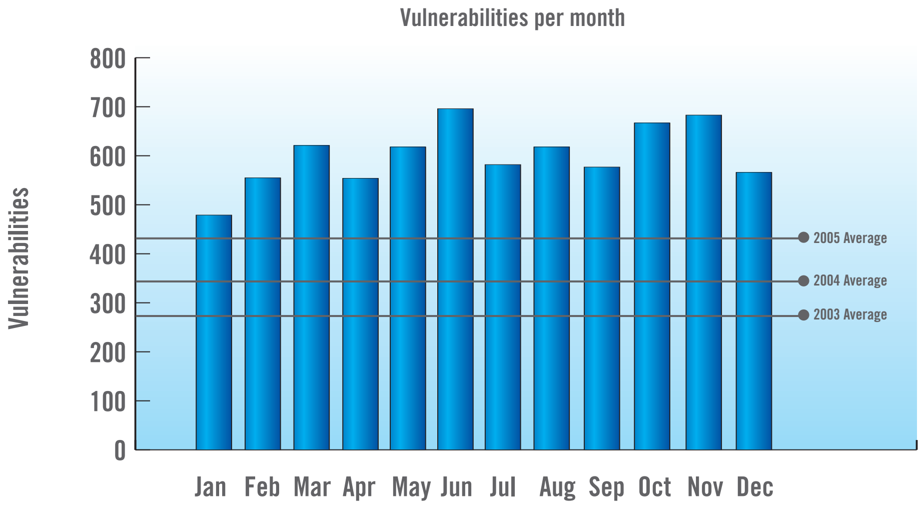


Year	Vulnerabilities	Avg per month	Avg per week	% increase year over year
2000	2007	167	39	
2001	1918	160	37	-4.4%
2002	3210	268	62	67.4%
2003	3156	263	61	-1.7%
2004	4606	384	89	45.9%
2005	5195	433	100	12.8%
2006	7247	604	139	39.5%

Vulnerabilities per Month

The average number of vulnerabilities per month has substantially increased since 2000, from an average of 167 in 2000 to 604 in 2006. June was the busiest month, reaching 696 vulnerabilities.

The following chart also shows the number of new vulnerabilities identified, catalogued and researched by X-Force for each month in 2006. The line plotted across the chart shows the average number of vulnerabilities, threats and security checks released each year, further illustrating the growth.

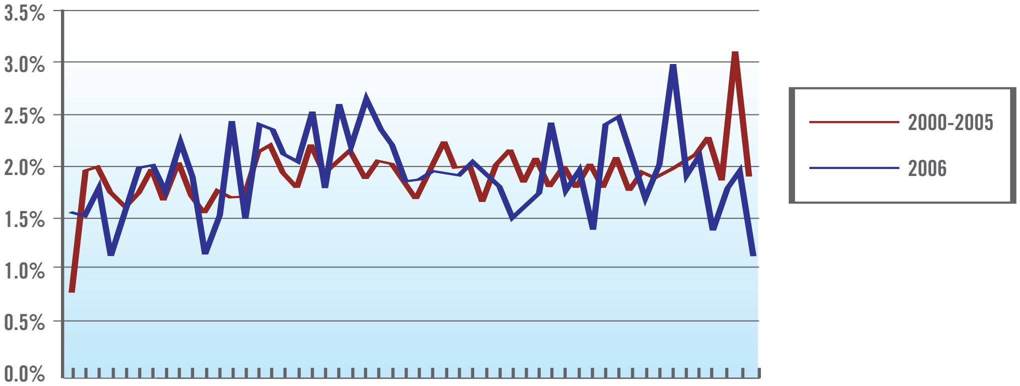


Month	Count
Jan-06	479
Feb-06	555
Mar-06	621
Apr-06	554
May-06	618
Jun-06	696
Jul-06	582
Aug-06	618
Sep-06	577
Oct-06	667
Nov-06	683
Dec-06	566

Vulnerabilities per Week

Vulnerability disclosure is subject to many external influences. Over the years, factors such as school holidays, religious festivals and significant world events have influenced when new vulnerabilities will be disclosed. The significance of these factors is best seen when studying the weekly breakdown of vulnerability disclosures.

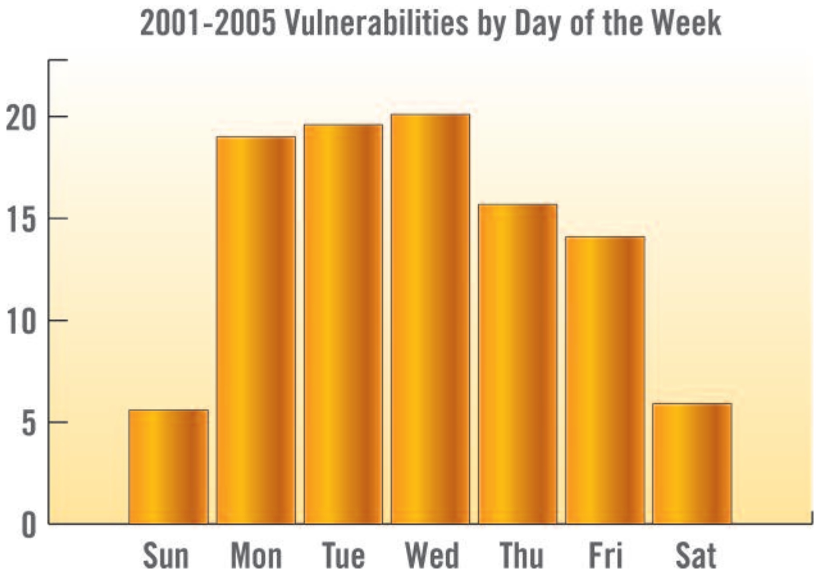
The following graph indicates the average weekly percentage of annual vulnerabilities between 2000 and 2005, with the weekly breakdown for 2006 superimposed on top. Traditionally, week 51 (the week prior to Christmas) is the busiest week of the year for vulnerability disclosures. However, X-Force found the week before Thanksgiving to be the busiest in 2006.



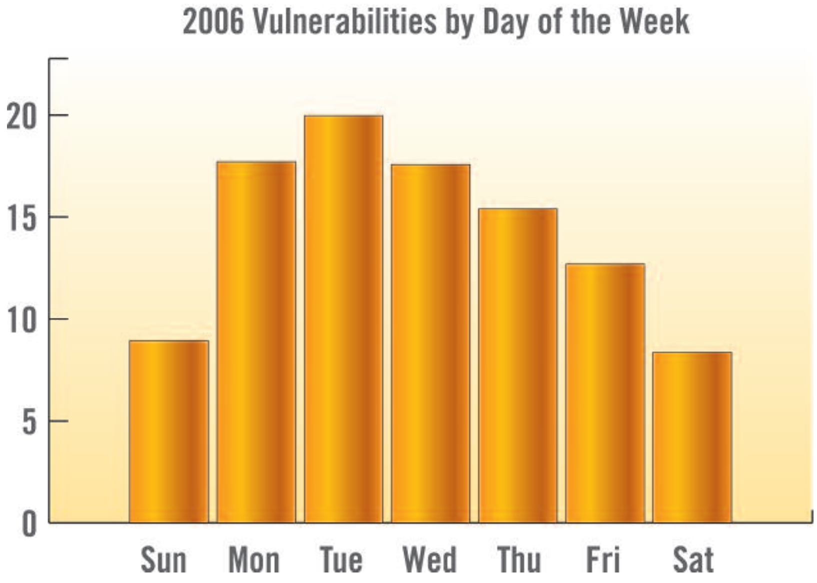
Graph: Percentage of annual vulnerabilities per week.

Vulnerabilities by Day of Week

From 2000 to 2005, the average busiest day of the week for vulnerability disclosures was a Wednesday, and the quietest day of the work week was Friday. This runs counter to traditional thought that Friday was the busiest day for vulnerabilities.

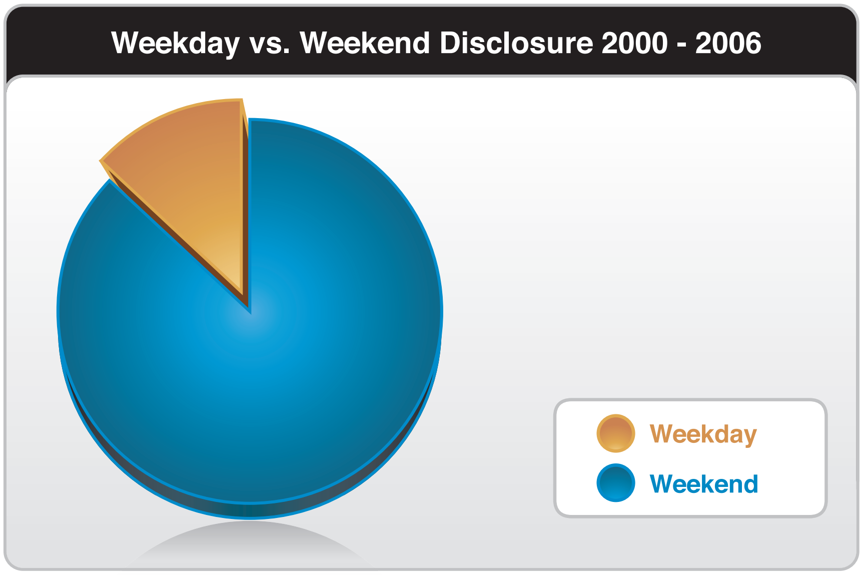


In 2006, X-Force observed that Tuesday became the most popular day for vulnerability disclosure. X-Force believes that the migration to Tuesday is likely due to the widespread adoption of Tuesday by commercial software vendors to coordinate vulnerability disclosure with patch availability – something that is most commonly attributed to Microsoft.



Weekday vs. Weekend

From 2000 to 2005, the trend in vulnerability disclosure leaned toward disclosing new vulnerabilities during the working week, with the number of vulnerabilities disclosed on the weekends decreasing. But in 2006, X-Force observed a change. Weekend disclosure has increased slightly, with the percentage of vulnerabilities disclosed on the weekend doubling that of 2005.



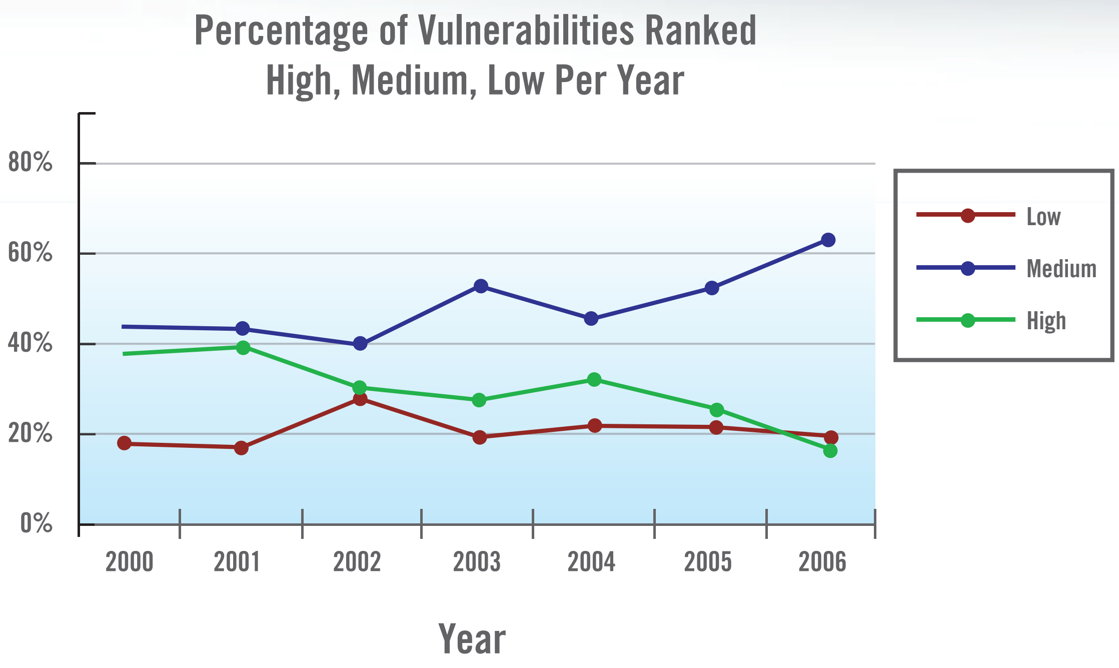
2000	2001	2002	2003	2004	2005	2006	Avg.
10.1%	10.7%	14.3%	15.9%	10.4%	8.1%	17.6%	13.1%
89.9%	89.3%	85.7%	84.1%	89.6%	91.9%	82.4%	86.9%

Classic High/Medium/Low Vulnerability Impact Breakdown

Each vulnerability catalogued by X-Force is researched and its exploitation impact is assessed. X-Force uses two impact categorization formats: the classical high/medium/low and CVSS scoring.

Looking at the high/medium/low impact breakdown for vulnerability disclosures since 2000, X-Force observed a marked decrease in the percentage of high impact vulnerabilities and a reciprocal increase in medium impact vulnerabilities. Low impact vulnerabilities have remained fairly consistent over the years. In 2006, high, medium and low impact vulnerabilities accounted for 18, 65 and 16 percent respectively.

The trend toward a lower percentage of high impact vulnerabilities is most likely due to the advances in vulnerability discovery using automated processes – such as fuzzing – which tend to yield high volumes of medium impact content-level vulnerabilities such as cross-site scripting, and desktop application file format vulnerabilities.



X-Force defines high, medium and low impact vulnerabilities according to the following criteria:

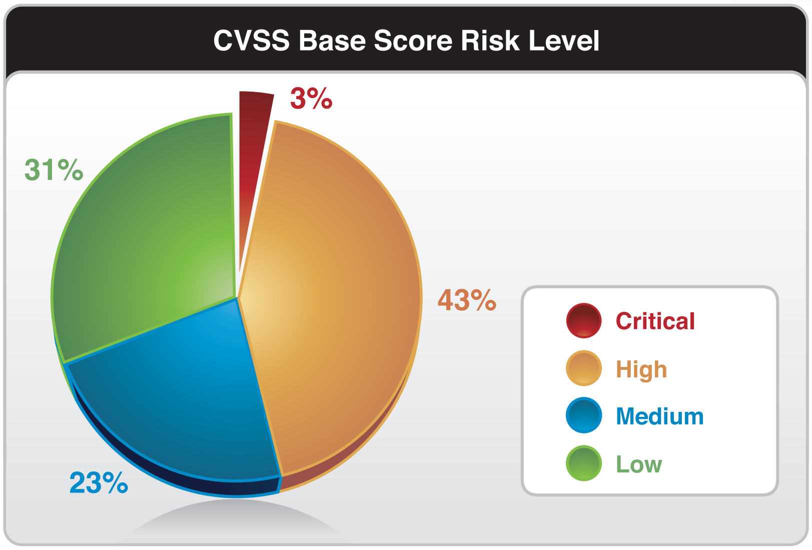
- High: Security issues that allow immediate remote or local access or immediate execution of code or commands with unauthorized privileges. Examples include most buffer overflows, backdoors, default or no password, and bypassing security on firewalls or other network components.
- Medium: Security issues that have the potential of granting access or allowing code execution via complex or lengthy exploit procedures, or low-risk issues applied to major Internet components. Examples are cross-site scripting, man-in-the-middle attacks, SQL injection, denial of service of major applications, and denial of service resulting in system information disclosure (such as core files).
- Low: Security issues that deny service or provide non-system information that could be used to formulate structured attacks on a target, but not directly gain unauthorized access. Examples include brute force attacks, non-system information disclosure (configurations, paths, etc.) and denial of service attacks.

Common Vulnerability Scoring System (CVSS) Breakdown

CVSS is an industry open standard for rating vulnerability severity and risk based on metrics and formulas. The base metrics are composed of characteristics that are intrinsic to a vulnerability and do not change over time, including the access vector, complexity and impact. The temporal metrics are composed of characteristics of a vulnerability that change over time, including public exploit availability, vendor confirmation and remediation.

In July 2006, X-Force began assigning CVSS base and temporal ratings to all new public vulnerabilities released. In addition, X-Force continued to assign the traditional high, medium or low rating.

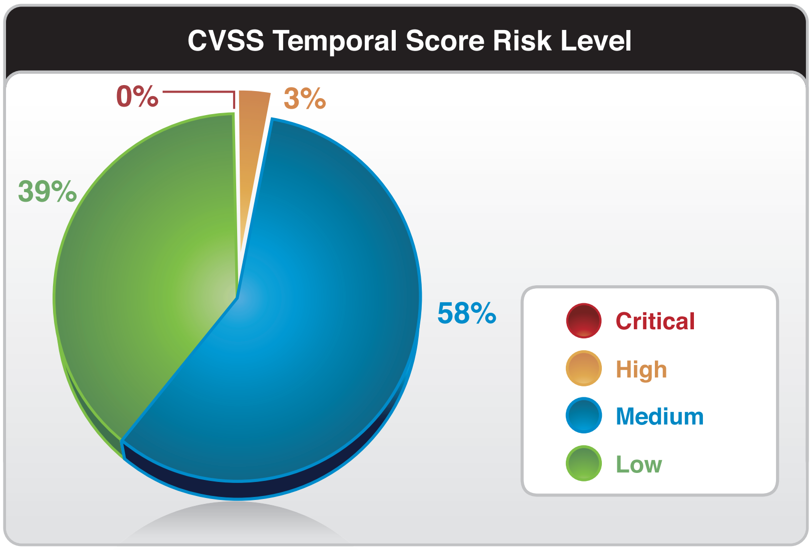
The following graphs show the percentages of critical, high, medium and low risk vulnerabilities based on the CVSS base and temporal score data. Since CVSS scores are based on a numerical value between 0 and 10.0, the graph below is formatted according to the same standard.



CVSS Base/Temporal Score	Risk Level
10.0	Critical
7.0 – 9.9	High
4.0 – 6.9	Medium
0.0 – 3.9	Low

Vulnerabilities identified as "critical" (i.e. have a CVSS base score of 10) are vulnerabilities that are typically installed by default, network-routable, do not require authentication to access, and when exploited will provide an attacker with system or root-level control of the host.

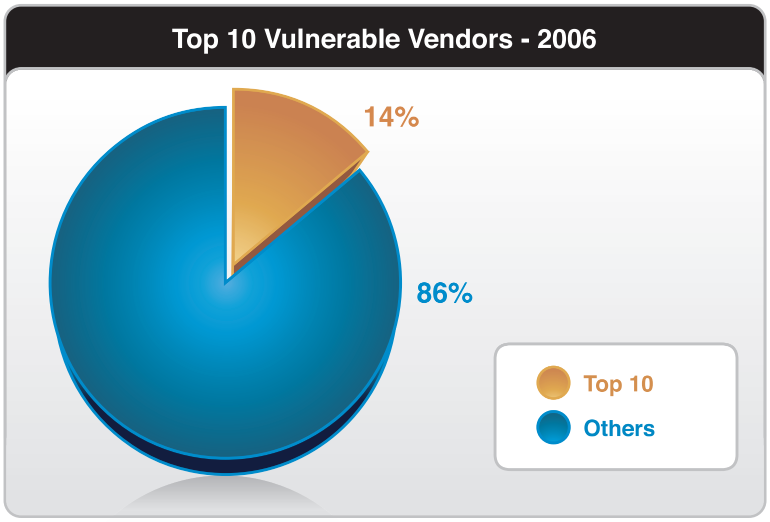
When considering patch, exploit and proof-of-concept code availability, the CVSS temporal score provides a more accurate risk analysis of a vulnerability. The following chart illustrates how the temporal data affects the CVSS scores of all vulnerabilities since X-Force began recording CVSS values in July 2006.



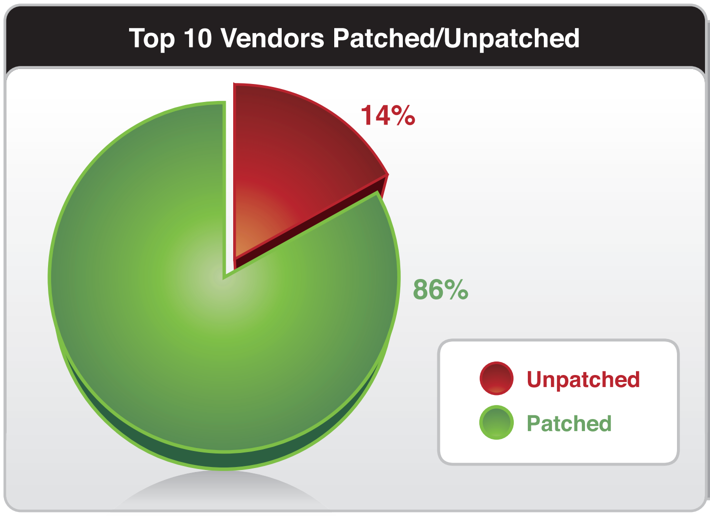
Top Ten Vulnerable Vendors

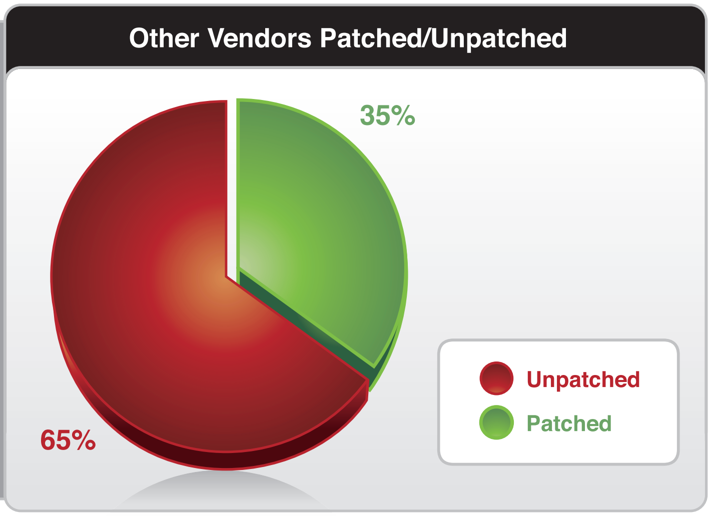
The top 10 vulnerable vendors in 2006 accounted for 964 of the total vulnerabilities disclosed. Although this seems like a large number of vulnerabilities that likely affected millions of systems worldwide, it only accounted for 14 percent of the total vulnerabilities disclosed during the year. If system administrators and end-users only implement workarounds or apply security patches and upgrades to vulnerabilities in top-vendor software and hardware, it is likely that several thousand vulnerable software packages go unnoticed and un-patched.

Vendor	Percentage of 2006
Microsoft Corporation	3.1%
Oracle Corporation	2.1%
Apple Computer, Inc.	1.9%
Mozilla Corporation	1.4%
IBM	1.2%
Linux Kernel Organization, Inc.	1.2%
Sun Microsystems, Inc.	1.0%
Cisco Systems, Inc.	0.9%
Hewlett-Packard	0.6%
Adobe Systems Incorporated	0.4%



Even more troubling, the other software and hardware vendors that account for 86 percent of the total vulnerabilities released make up a much larger percentage of the vulnerabilities that remain un-patched. Out of the top 10 vulnerable vendors, only 14 percent of the publicly-disclosed vulnerabilities remain un-patched, while 65 percent of all other publicly-disclosed vulnerabilities remain un-patched.



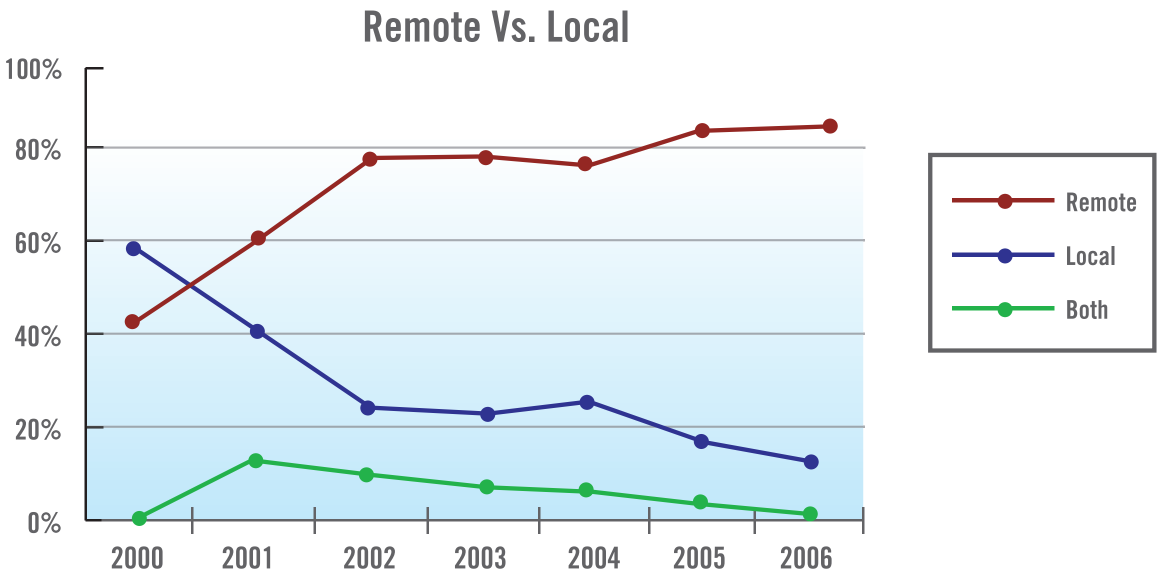


These calculations take into account vendors that have publicly acknowledged a vulnerability and released a corresponding fix or patch. They do not take into account cases where a vendor silently fixes a vulnerability without an announcement, or a patch is released by a third-party vendor.

Remote vs. Local Exploitation

In a world of vulnerability exploitation, the most highly prized vulnerabilities are those that can be exploited remotely – thereby providing an attacker with the greatest opportunity for host compromise. In 2006, the percentage of remotely exploitable vulnerabilities reached an all-time high of 88.4 percent.

In the following graph, “Remote” represents vulnerabilities that can be exploited over a network. “Local” represents vulnerabilities that can only be exploited after logging in to the local host or from the desktop. And “Both” represents vulnerabilities that can be exploited both remotely and locally. For the purposes of the graph below, 'Both' is an assumed subcategory of the 'Remote' total.

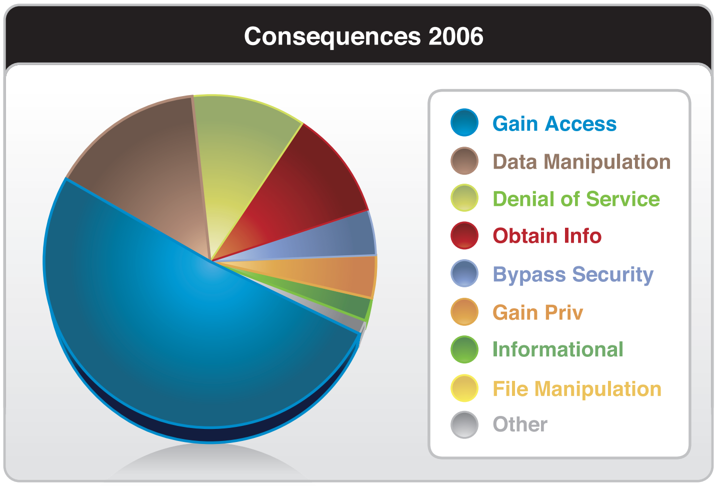


When comparing the number of actual vulnerabilities that can be exploited by a remote attacker vs. those that can be exploited by a local attacker, X-Force observed an alarming trend that has been developing since the year 2000. In 2000, only 43.6 percent of all vulnerabilities disclosed were remotely exploitable. The number of vulnerabilities that can be exploited remotely has continued to grow each year since, reaching its highest point in 2006.

Year	% of Remote Vulnerabilities	% of Local Vulnerabilities
2000	43.6%	56.4%
2001	57.4%	42.6%
2002	75.7%	24.3%
2003	76.6%	23.4%
2004	73.3%	26.7%
2005	84.8%	15.2%
2006	88.4%	11.6%

Consequences of Exploitation

As part of the research into each vulnerability disclosed in 2006, X-Force records the primary consequence of exploitation. The consequences are divided into nine categories. In 2006, the most common consequence of exploitation was “Gain Access,” which accounted for 50.6 percent of all vulnerabilities.



Gain Access	Data Manipulation	Denial of Service	Obtain Info	Bypass Security	Gain Privileges	Informational	File Manipulation	Other
50.6%	14.6%	11.0%	10.4%	4.5%	4.1%	2.2%	1.5%	1.1%

The nine categories described above include:

- Bypass Security – An attacker can bypass security restrictions such as a firewall or proxy, an IDS system or a virus scanner.
- Data Manipulation – An attacker is able to manipulate data stored or used by the host associated with the service or application.
- Denial of Service – An attacker can crash or disrupt a service or system or take down a network.
- File Manipulation – An attacker can create, delete, read, modify or overwrite files.
- Gain Access – An attacker can obtain local and remote access. This also includes vulnerabilities by which an attacker can execute code or commands, because this usually allows the attacker to gain access to the system.
- Gain Privileges – Privileges can be gained on the local system only.
- Obtain Information – An attacker can obtain information such as file and path names, source code, passwords or server configuration details.
- Informational – Service name disclosure.
- Other

Spam and Phishing Analysis

The IBM ISS premier content filtering services provide a world-encompassing view of spam and phishing attacks. With millions of e-mail addresses actively monitored, X-Force has identified numerous advances in the spam and phishing technologies attackers use.

Within the last 12 months, the volume of spam has increased by 100 percent. Consequently, even though spam detection algorithms and technologies have improved throughout the year, it often appears that more spam e-mails make their way to the user's inbox and therefore subjectively the users "feel" that spam detection has not improved.

On an average day, IBM ISS analyzes more than 150,000 unique spam messages – whereby, using a fuzzy-fingerprint technology, a "unique" spam message is one that is at least 10 percent different to any other spam message ever received.

This section includes the following analysis:

- From which countries does spam originate?
- Where are the Web pages contained in spam messages hosted?
- What is the average byte size of spam?
- How much spam uses HTML?
- How many e-mail servers did spam pass through before reaching its destination?
- What are the most popular subject lines of spam?
- What amount of spam exhibited a Reply-To: different from the From: message data?
- What amount of spam had a Return-Path: different from the From: message data?
- What is the language distribution of spam?
- Where do phishing e-mails come from?
- Where are the Web pages contained in phishing e-mails hosted?
- Where are the phishing targets located?
- How much phishing uses HTML?
- How many e-mail servers was phishing passed through?
- What is the effect of geographical distribution of spam?
- What is the history and future prospect of image-based spam?

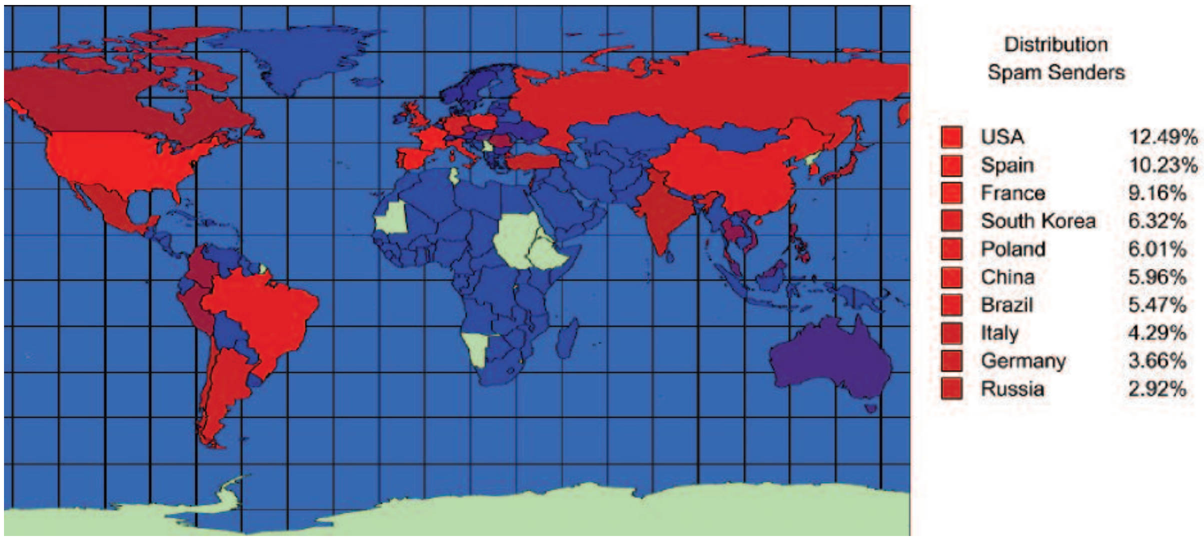
Basics about the determination of geographical distributions

The following statistics use the IP-to-Country Database provided by WebHosting.Info (http://www.webhosting.info), available from http://ip-to-country.webhosting.info.

The geographical distribution was determined by requesting the IP addresses of the hosts (in the case of the content distribution) or of the sending mail server (in the case of spam and phishing) to the IP-to-Country Database.

From which countries does spam originate?

The following map shows the origination point for spam globally. Here we see that the U.S. accounts for one-eighth of worldwide spam.

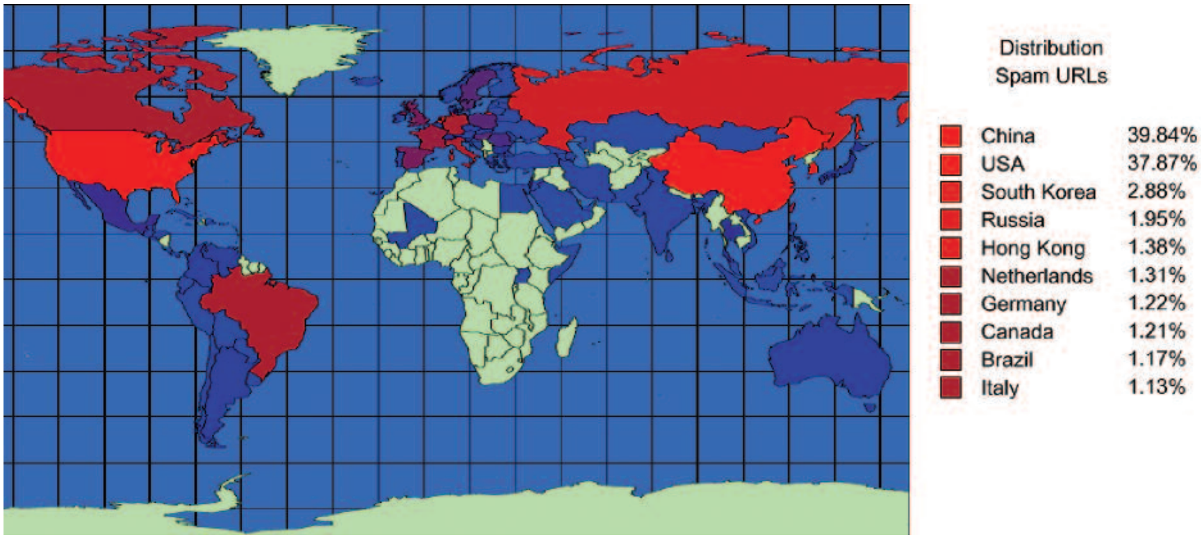


🟥	USA	12.49%
🟥	Spain	10.23%
🟥	France	9.16%
🟥	South Korea	6.32%
🟥	Poland	6.01%
🟥	China	5.96%
🟥	Brazil	5.47%
🟥	Italy	4.29%
🟥	Germany	3.66%
🟥	Russia	2.92%

Figure 1- Geographical distribution of spam senders

Where are the Web pages contained in spam messages hosted?

The map shows where the spam URLs are hosted.



🟥	China	39.84%
🟥	USA	37.87%
🟥	South Korea	2.88%
🟥	Russia	1.95%
🟥	Hong Kong	1.38%
🟥	Netherlands	1.31%
🟥	Germany	1.22%
🟥	Canada	1.21%
🟥	Brazil	1.17%
🟥	Italy	1.13%

Fig. 2: Geographical distribution of Spam URLs

What is the average byte size of spam messages?

Spam messages have grown in size over the last two years, increasing from an average of 6 kilobytes to 9.5 kilobytes. This is largely due to the increased inclusion of random data designed to help the spam bypass first-generation anti-spam technologies, and the use of images to convey message content.

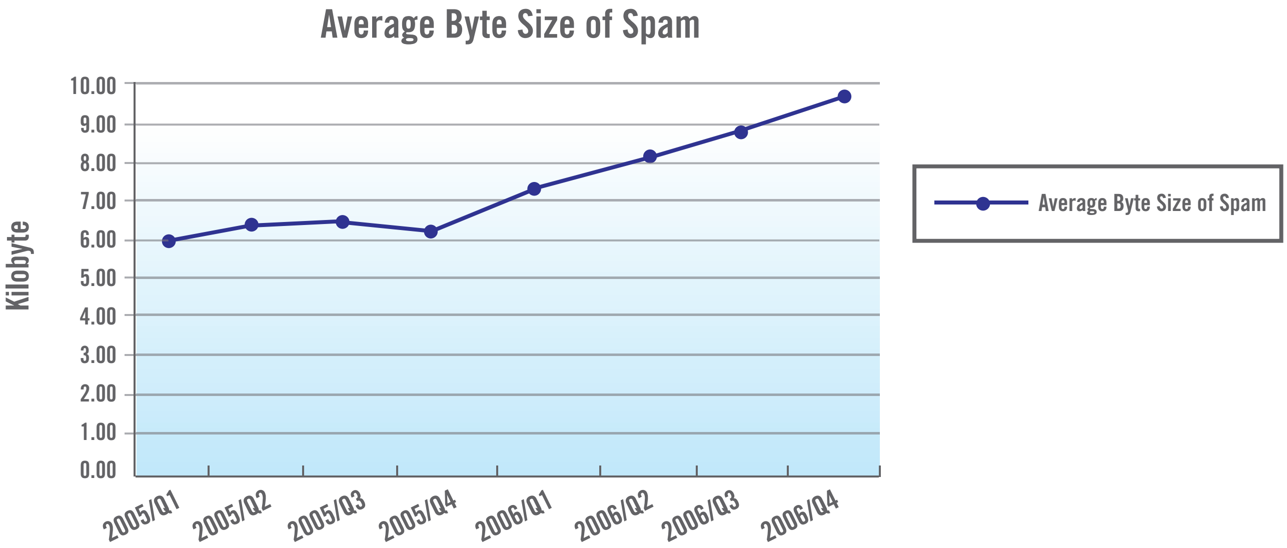


Fig. 3: Average Byte Size of Spam 2005 and 2006

This trend correlates with the increase of image-based spam (see below).

How much spam uses HTML?

In 2006 there was a significant tendency to use HTML in spam:

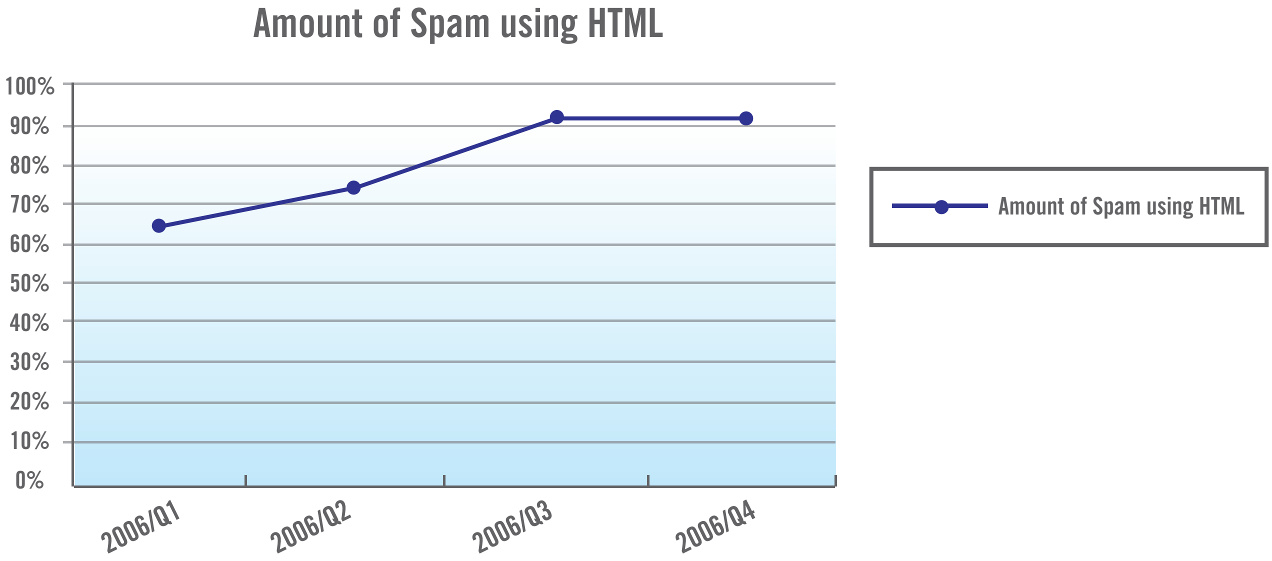


Fig. 4: Amount of Spam using HTML in 2006

How many e-mail servers did spam pass through before reaching its destination?

Since IBM ISS began monitoring e-mail server relay hops in June 2006, most spam was sent directly to the recipient's email server:

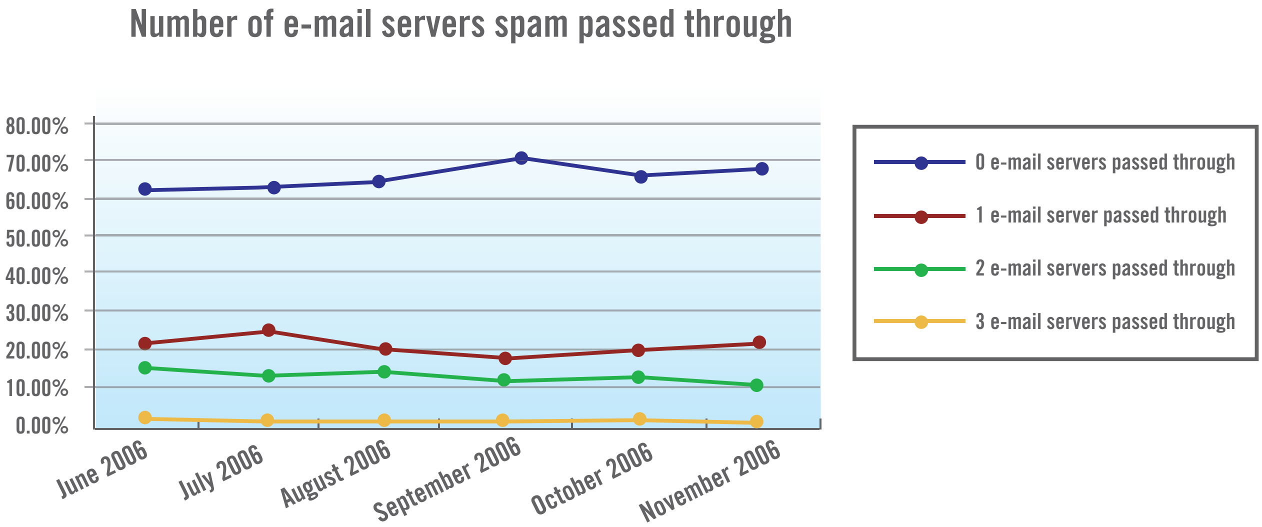


Fig. 5: Number of e-mail servers spam passed through H2/2006

Since most spam and phishing messages are sent via botnets, the botnet agents mostly send spam messages directly to the recipient – which accounts for the "0 e-mail servers passed through" category in the above chart.

What are the most popular subject lines of spam?

The most popular subject lines of spam in 2006 appear below:

Subject line	Quota
Re: hi	1.47%
Canadian online drugstore	0.77%
<empty subject line>	0.68%
Re: VIhAGRA	0.57%
Re:	0.46%
Re: VkAGRA	0.46%
Re: new	0.38%
Re: 482	0.31%
hello	0.30%
Re: VIpAGRA	0.28%

What amount of spam exhibited a Reply-To: different from the From: message data?

The usage of Reply-To: data differing from From: data remains at a low stage, but can be seen ascending slightly.

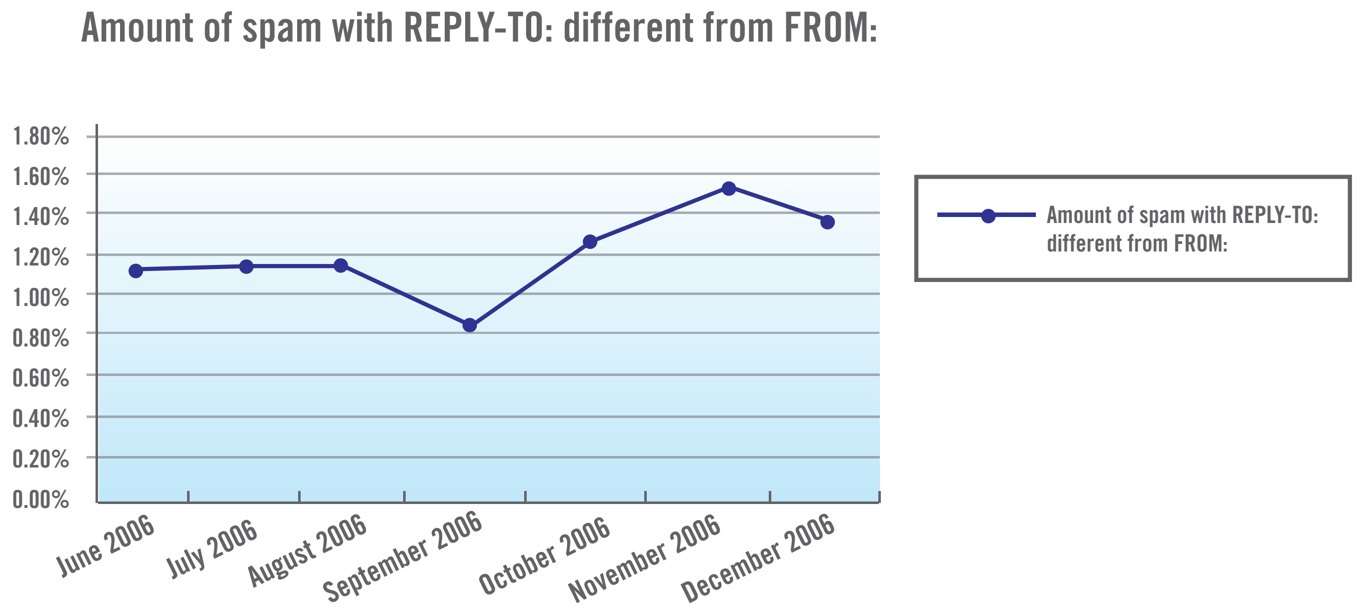


Fig. 6: Amount of spam with Reply-To: different from From:

What amount of spam had a Return-Path: different from the From: message data?

The usage of Return-Path: data differing from From: data is declining.

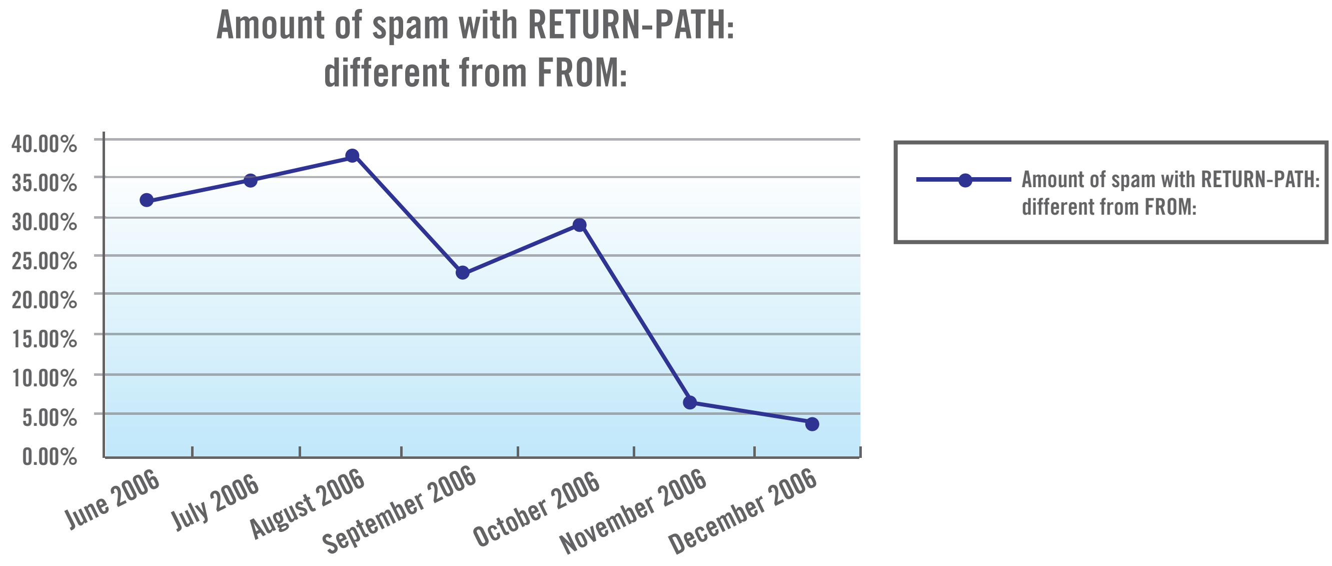


Fig. 7: Amount of spam with Return-Path: different from From:

What is the language distribution of spam?

The top five languages used in spam messages appear below:

Language	Quota
English	92.99%
German	1.76%
Korean	1.00%
Portuguese	0.58%
Russian	0.42%

Where do phishing e-mails come from?

The following map highlights countries of origin for phishing e-mails.

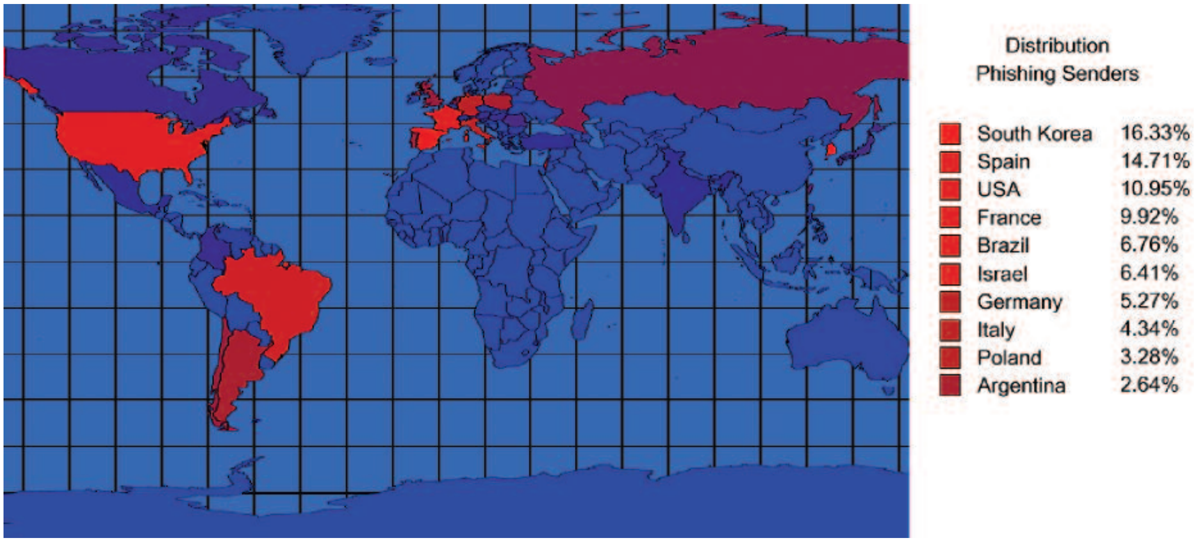


Fig. 8: Geographical distribution of phishing senders

Where are Web pages contained in phishing e-mails hosted?

The map shows where the Phishing URLs are hosted.

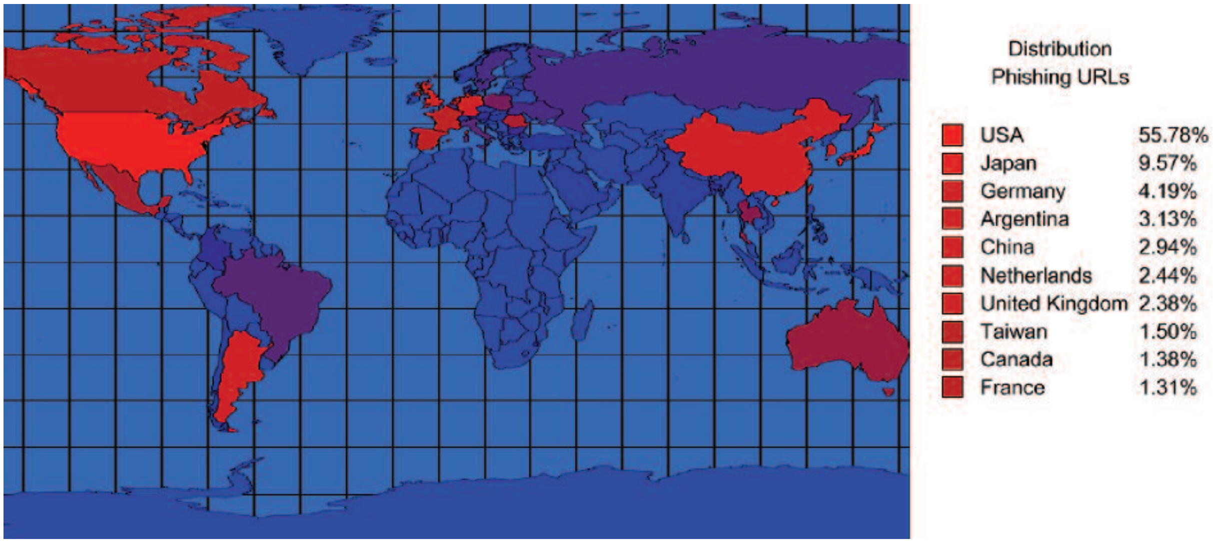


Fig. 9: Geographical distribution of Phishing URLs

Where are phishing targets located?

The map below indicates where phishing targets are located.

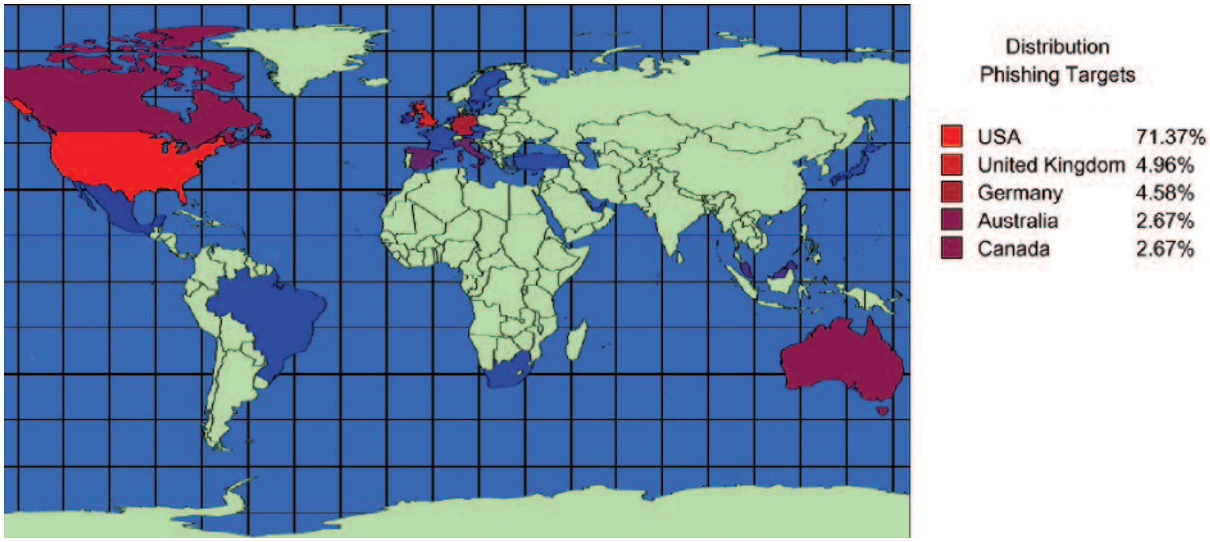


Fig. 10: Geographical distribution of the phishing targets

How much phishing uses HTML?

In 2006, nearly all phishing e-mails have used HTML.

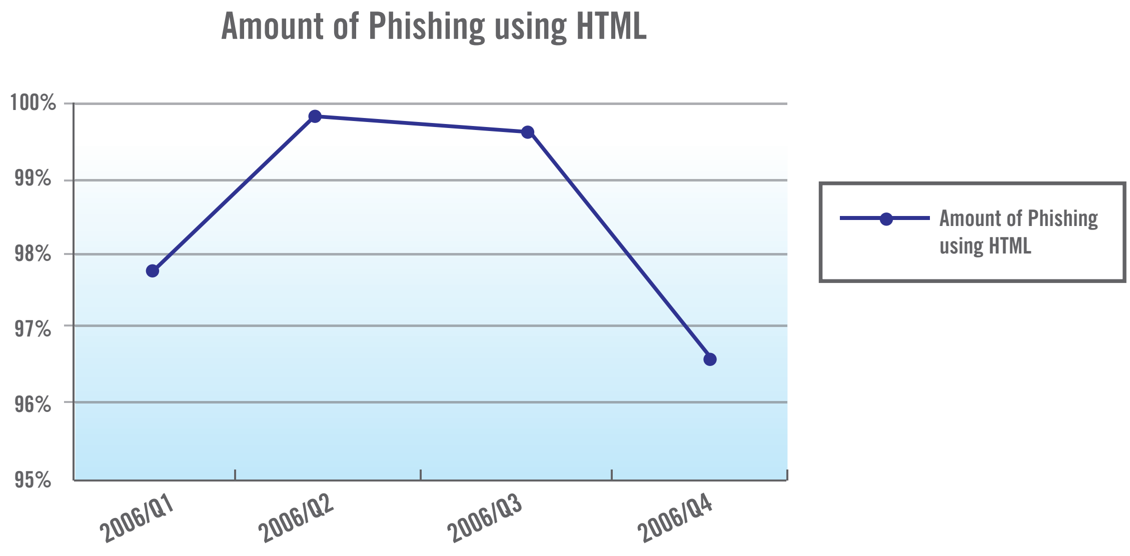


Fig. 11a: Amount of Phishing using HTML in 2006

How many e-mail servers was phishing passed through?

Phishing e-mails are either sent directly to the recipient or pass through one e-mail server.

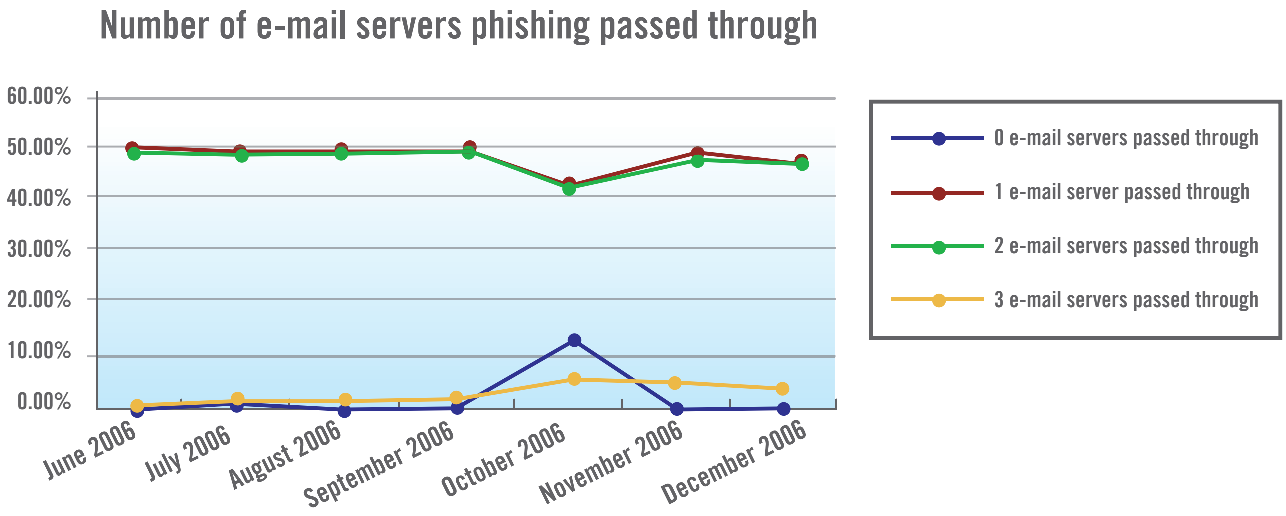


Fig. 11b: Number of e-mail servers phishing passed through H2/2006

What is the effect of geographical distribution?

In the case of spam and phishing, the geographical distribution of the senders and the geographical distribution of the spam/phishing URLs diverge. The reasons for this are two-fold:

- In many cases (e.g. in the case of image-based spam), the spam messages do not contain any URLs.
- The majority of spam and phishing e-mails are sent via botnets – thus, there is no relationship between the location of origin for the spam and the location where the spam/phishing URLs are hosted.

Furthermore, observers will note the high correlation between the geographic distribution of spam and phishing and the economic strength of the country. This also manifests in the geographical distribution of shopping and banking Web sites.

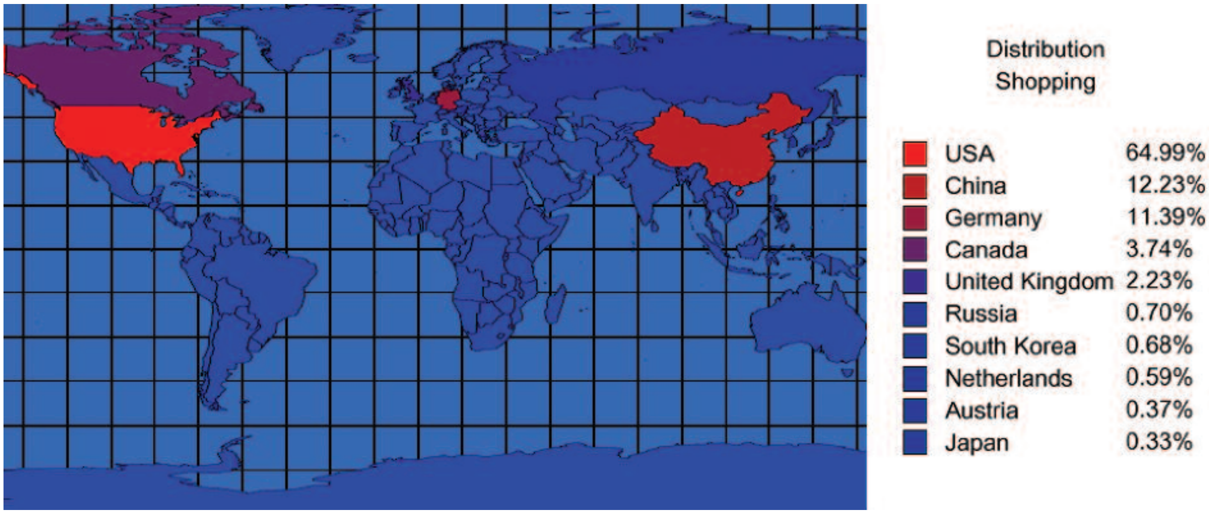


Fig. 12: Geographical distribution of shopping Web sites

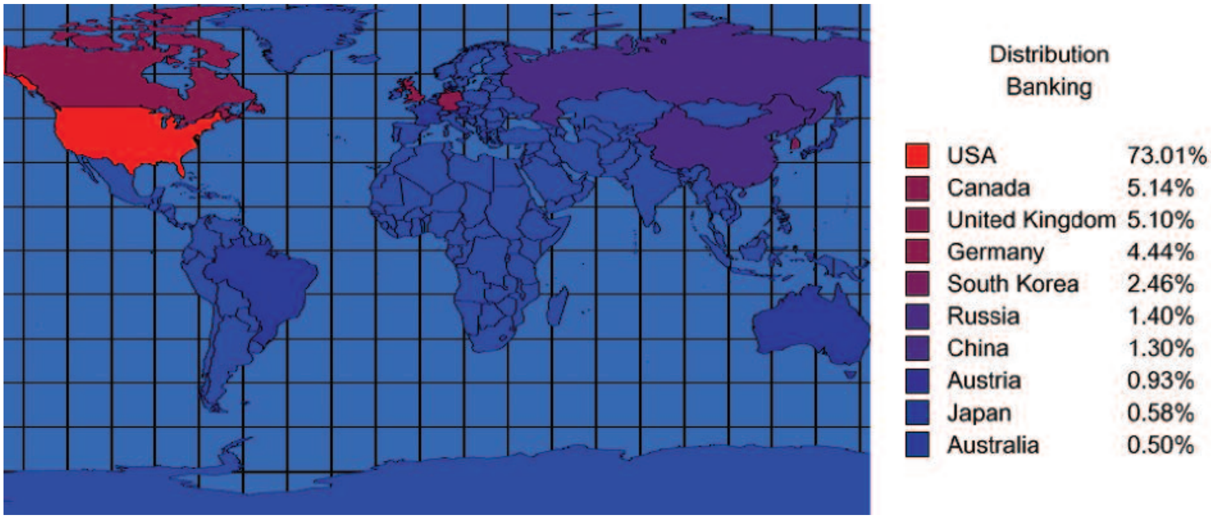


Fig. 13: Geographical distribution of banking Web sites

As expected, four of the top five countries targeted by phishing scams are also part of the top 10 countries hosting shopping and banking Web sites.

What is the history and future prospect of image-based spam?

Image-based spam became one of the anti-spam challenges of the year 2006. Over the last two years, the rate of image-based spam has grown rapidly as viewed in the graph below:

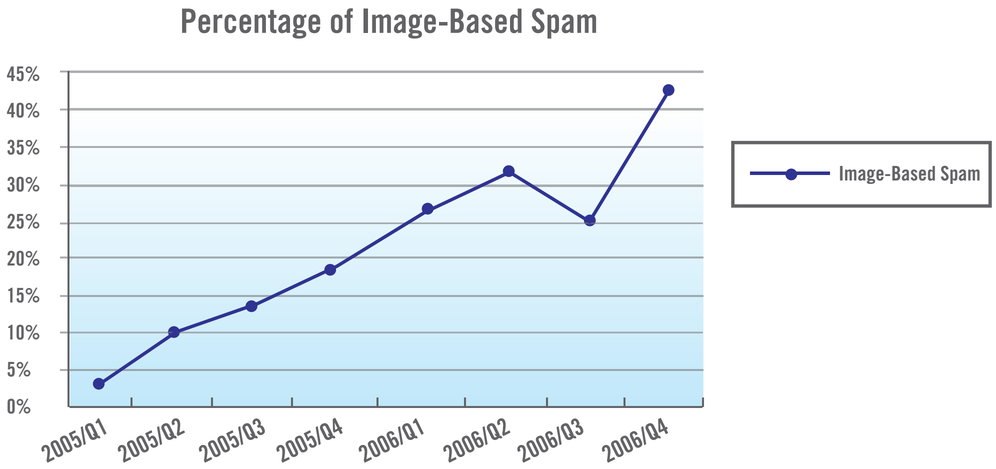


Fig. 14: Development of Image-based Spam 2005 and 2006

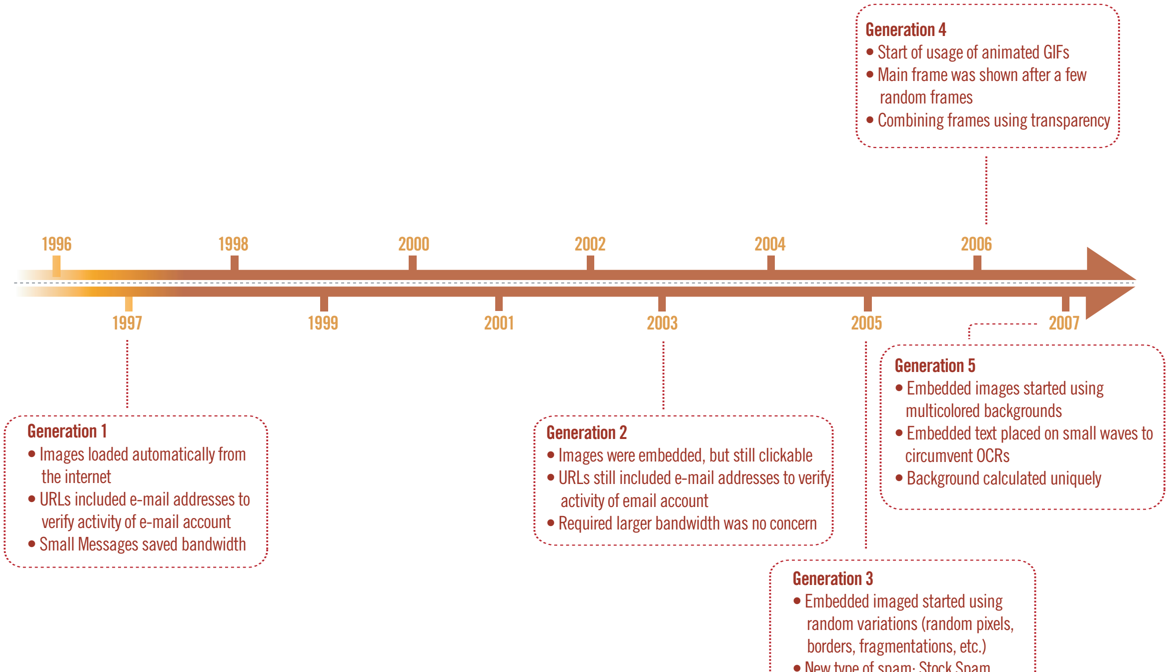
What is image-based spam?

Spam containing embedded images is called image-based spam. The actual message of the spam is not contained in the text of the e-mail, but solely within the image that looks like regular text (embedded text).

The challenge of today's anti-spam products is to detect image-based spam even though there is no text available that indicates a spam-related topic.

How quickly did image-based spam develop?

After a slow start in the early years of spam, image-based spam has developed more rapidly within the last two years.



Generation 1
- Images loaded automatically from the internet
- URLs included e-mail addresses to verify activity of e-mail account
- Small Messages saved bandwidth

Generation 2
- Images were embedded, but still clickable
- URLs still included e-mail addresses to verify activity of email account
- Required larger bandwidth was no concern

Generation 3
- Embedded imaged started using random variations (random pixels, borders, fragmentations, etc.)
- New type of spam: Stock Spam

Generation 4
- Start of usage of animated GIFs
- Main frame was shown after a few random frames
- Combining frames using transparency

Generation 5
- Embedded images started using multicolored backgrounds
- Embedded text placed on small waves to circumvent OCRs
- Background calculated uniquely

Fig. 15: Timeline of Image-based Spam

Next Generation Spam

Spammers dedicated to image-based spam have many new ways to outsmart detection methods. For example, animated GIF images and multi-color, embedded text in images cannot be easily separated from the background by an algorithm.

When dealing with animated GIF images, the protection technology must decide which frame should be analyzed. When spammers make use of all the features available within the animated GIF image standard – like transparency of pixels, overlaying frames and showing the last frame several minutes after the frame before – antispam technologies that rely on image analysis techniques have a hard time deciding which frame to use.

In addition, multi-color images require expensive image analysis methods, which many technologies try to avoid because, when receiving 80-90 percent spam, one-third of current e-mails are image-based spam.

Therefore, more general methods have to be developed that do not only detect one or two of the image-based spam types mentioned above, but also detect types of image-based spam that are not yet in the wild – but expected to arrive during 2007.

Web Content Trends

The growth of the Internet is still unchecked. With respect to content filtering, the development of the "bad" Web filter categories around "pornography," "computer crime," etc. is a matter of particular interest.

This section covers the following analysis topics:

- Growth of bad content within the last 12 months
- Current distribution of violence and crime-related Web sites
- Current distribution of porn and sex-related Web sites
- Current distribution of computer crime-related Web sites
- Current distribution of illegal drug-related Web sites

Analysis

The content distribution of the Internet and its growth were determined by counting the hosts classified in the corresponding Web filter categories of the IBM ISS Web Filter Database.

Counting hosts is the most common method to determine content distribution of the Internet and provides the most realistic overview. When using another methodology (like counting Web pages/sub pages), other results may arise.

The IBM ISS data center is constantly reviewing and analyzing new Web content data. Consider the following statistics related to the IBM ISS data center:

- It analyzes 150 million new Web pages and images each month.
- Since 1999, it has analyzed 6.2 billion Web pages and images.

The IBM ISS Web Filter Database maintains the following characteristics:

- 62 filter categories
- 74 million entries
- 100,000 new or updated entries added each day

Current Status of Unwanted Internet Content

Currently, about 12.5 percent of the Internet deals with unwanted content such as pornography, violence and crime, and illegal drugs.

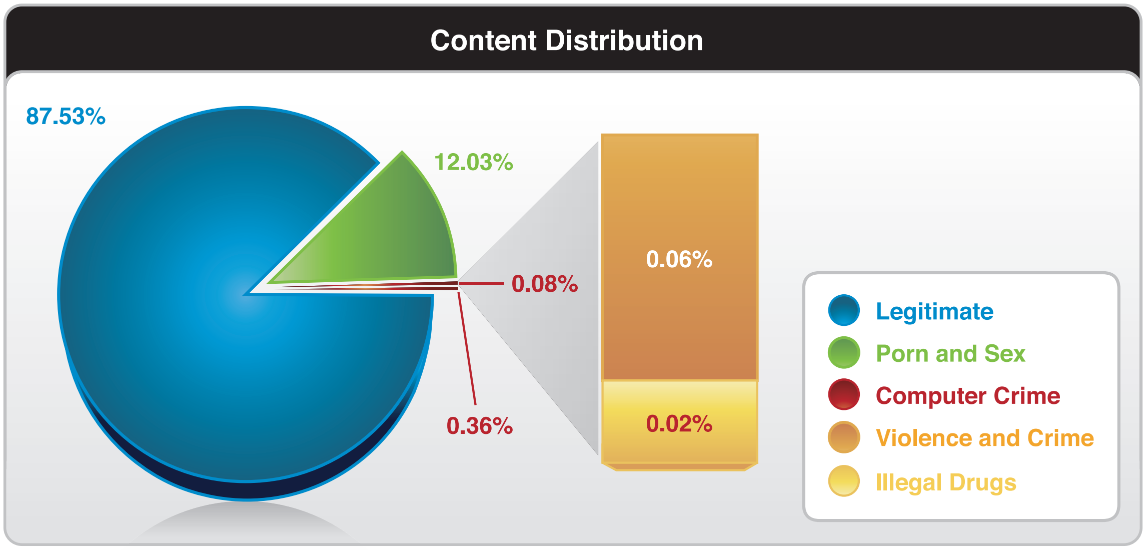


Fig. 1: Content distribution of the Internet

Growth of Bad Content within the Last 12 Months

The amount of Web sites containing unwanted content has grown by 8-15 percent in 2006. The table below provides further detail.

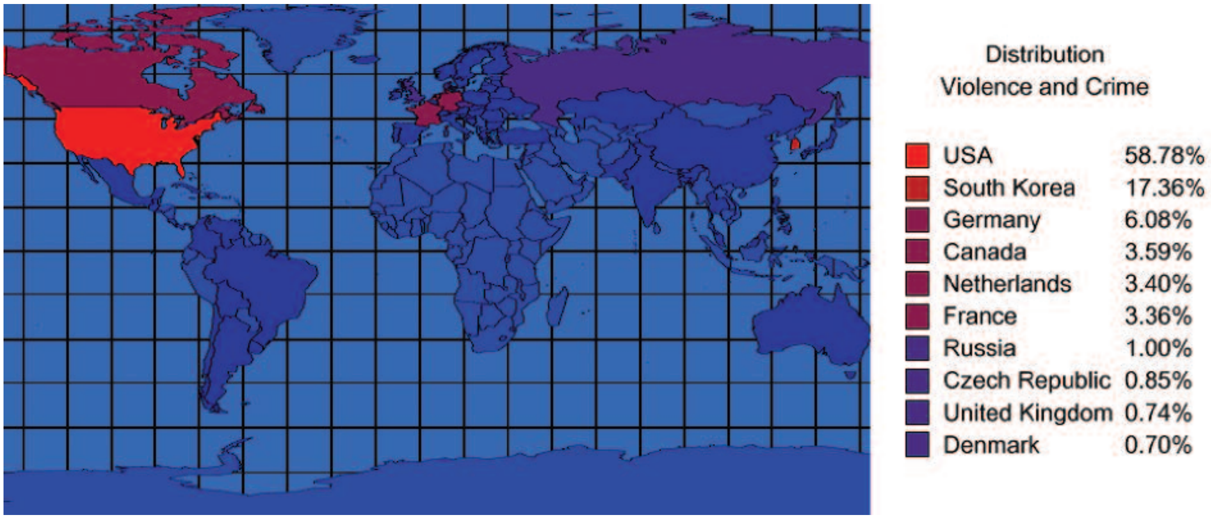
The explosive growth of the Internet has been driven by several aspects:

- Huge growth of blogs and small business Web sites (http://news.netcraft.com/archives/2006/11/01/november_2006_web_server_survey.html)
- A strong tendency to park domains, rather than to use them as active Web sites (http://news.netcraft.com/archives/2006/03/06/march_2006_web_server_survey.html)

The growth of the normally-legitimate topics above outstrips even the growth of the unwanted content categories mentioned.

Content	Growth
Total Growth of the Internet	43.6%
Violence and Crime	14.4%
Porn and Sex	12.8%
Computer Crime	10.0%
Illegal Drugs	8.7%

Current Distribution of Violence and Crime-related Web Sites

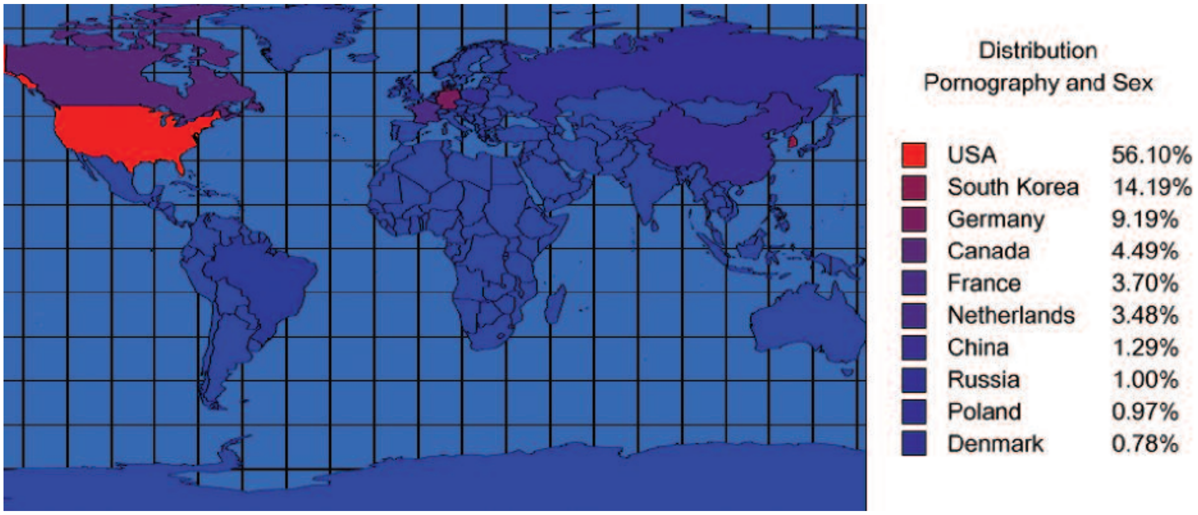


Distribution
Violence and Crime

USA	58.78%
South Korea	17.36%
Germany	6.08%
Canada	3.59%
Netherlands	3.40%
France	3.36%
Russia	1.00%
Czech Republic	0.85%
United Kingdom	0.74%
Denmark	0.70%

Fig. 2: Geographical distribution of violence and crime-related Web sites

Current Distribution of Porn and Sex-related Web Sites

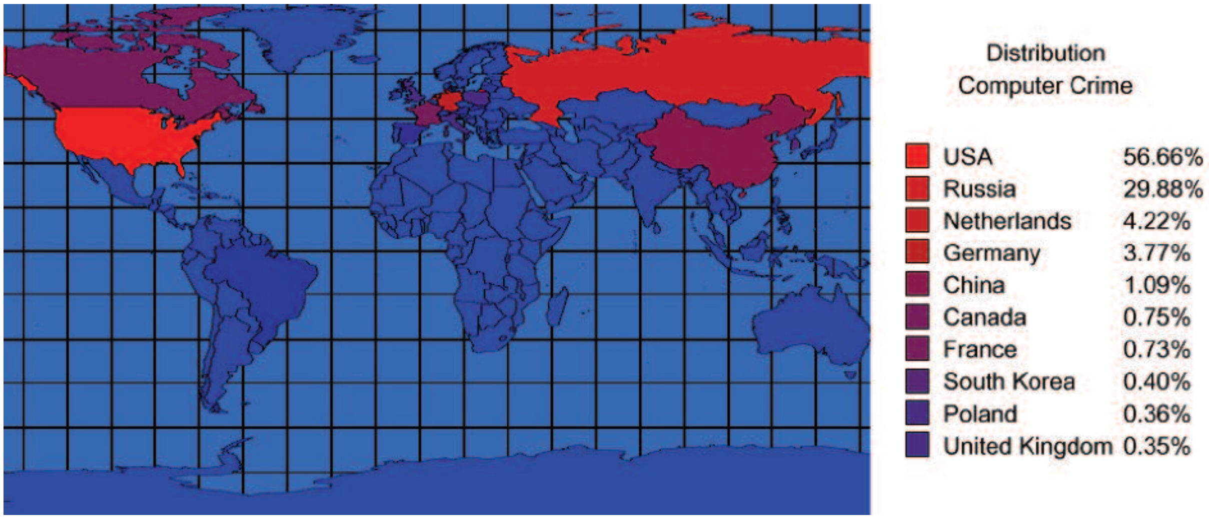


Distribution
Pornography and Sex

USA	56.10%
South Korea	14.19%
Germany	9.19%
Canada	4.49%
France	3.70%
Netherlands	3.48%
China	1.29%
Russia	1.00%
Poland	0.97%
Denmark	0.78%

Fig. 3: Geographical distribution of porn and sex-related Web sites

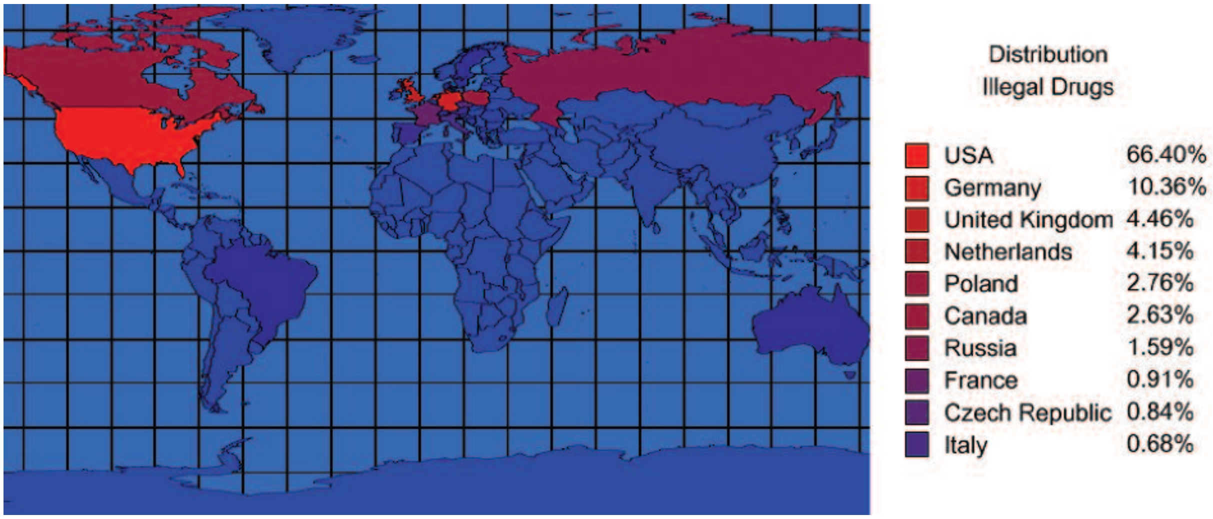
Current Distribution of Computer Crime-related Web Sites



🟥	USA	56.66%
🟥	Russia	29.88%
🟥	Netherlands	4.22%
🟥	Germany	3.77%
🟪	China	1.09%
🟪	Canada	0.75%
🟪	France	0.73%
🟪	South Korea	0.40%
🟦	Poland	0.36%
🟦	United Kingdom	0.35%

Fig. 4: Geographical distribution of computer crime-related Web sites

Current Distribution of Illegal Drug-related Web Sites



🟥	USA	66.40%
🟥	Germany	10.36%
🟥	United Kingdom	4.46%
🟥	Netherlands	4.15%
🟥	Poland	2.76%
🟥	Canada	2.63%
🟪	Russia	1.59%
🟪	France	0.91%
🟪	Czech Republic	0.84%
🟦	Italy	0.68%

Fig. 5: Geographical distribution of illegal drug, alcohol and tobacco-related Web sites

Malcode Analysis

2006 was a big year for malware, with new records in volume and sophistication occurring on a monthly basis. X-Force identified, studied and analyzed more than 200,000 new malware samples throughout the year.

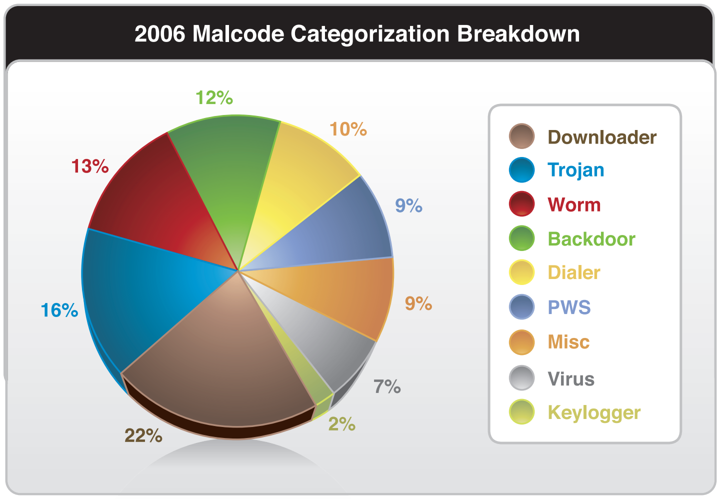
An important trend observed during 2006 was the way malcode continued to become less distinct in its categorization. Malcode continued to absorb or borrow new technologies being used by other successful malware. As such, the classical "buckets" of virus, worm, spyware, backdoor, etc. are largely irrelevant as we move in to 2007. Modern malware is now the digital equivalent of the Swiss Army knife.

Instead, classification (or "bucketing") must be done with regard to the most dominant or primary feature-set of the malware. In this section we review 2006 malware and, for analysis purposes, divide them into the following buckets:

- Worm – malware that can self-propagate over a network.
- Backdoor – malware that provides functionality for an attacker to connect back to the victim's system without supplying authorized login credentials.
- Virus – malware that infects a host and does some form of damage to the host, but cannot self-propagate.
- Password Stealer (PWS) – malware that is designed to steal the login credentials for specific online applications and is a key component in identity theft attacks.
- Downloader – low-profile malware that exists to install itself so that it can then download and install a more sophisticated or updated malware agent.
- Keylogger – malware that captures all key presses and stores the information away for later retrieval by the attacker.
- Dialer – malware that uses modem connections to either dial back to the attacker, or causes the victim to use primary-rate billing numbers when making connections.
- Trojan – malware that appears to be a legitimate file before installing itself – often with rootkit functionality.
- Miscellaneous – all other malware not falling into one of the above primary categories.

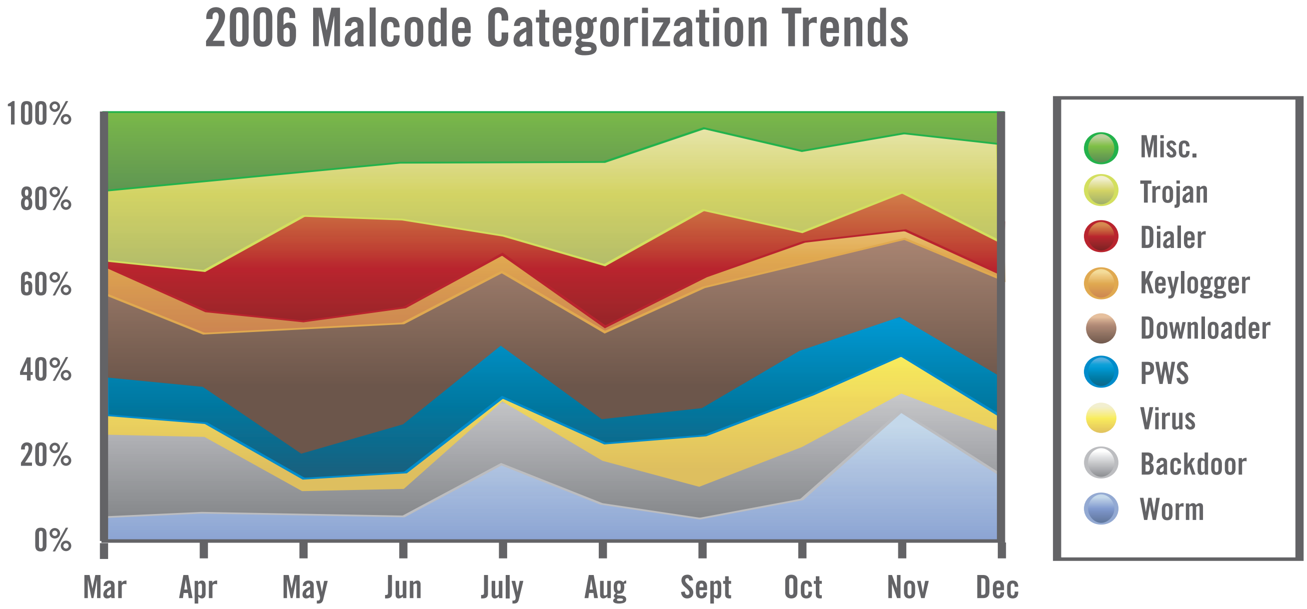
Malcode Categorization

The malware samples collected by X-Force during 2006 can be broken down into a number of key categories. The following graph indicates that the biggest class of malware was downloaders.



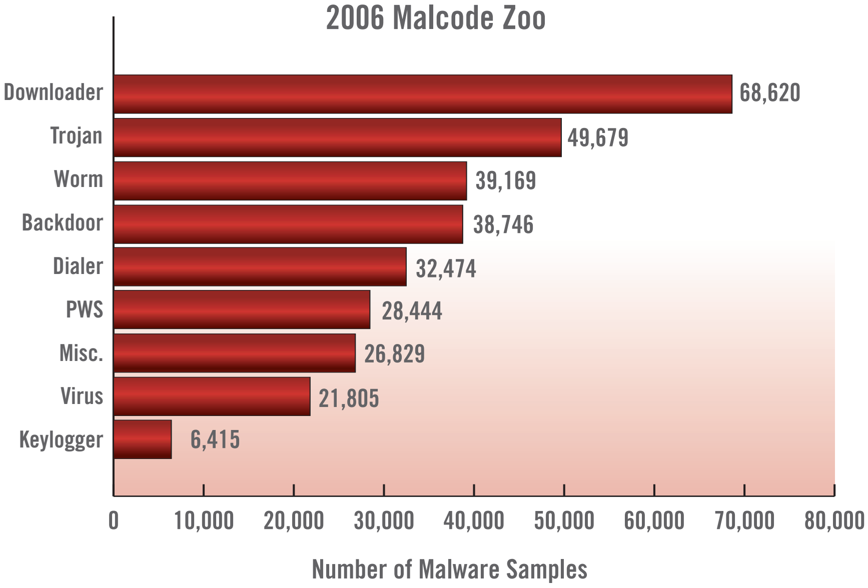
Malcode Categorization Trends

In 2006, the categorization distribution changed on a monthly basis to reflect large outbreaks of specific malware families. This is most clearly seen in November through December with the serial variant attacks of various worms.



The X-Force Malware

The following chart depicts the absolute volume of new malware identified by X-Force throughout 2006.



Top 10 Malware Families Most Frequently Encountered

The top 10 most popular exploits for each category researched are listed below.

Top 10 2006 Malcode
Trojan-Downloader.Win32.Zlob
Trojan-Downloader.Win32.Small
Email-Worm.Win32.Luder
Trojan-Downloader.Win32.Agent
Trojan-Spy.Win32.Banker
Trojan-Downloader.Win32.Obfuscated
Trojan-Downloader.Win32.Tibs
Trojan-Downloader.Win32.Banload
Backdoor.Win32.Hupigon
Backdoor.Win32.Rbot

Top 10 2006 Trojan
Trojan-Spy.Win32.Banker
Trojan.Win32.Agent
Trojan-Spy.Win32.Perfloger
Trojan-Spy.Win32.SCKeyLog
Trojan-Spy.Win32.BZub
Trojan-Dropper.Win32.Agent
Trojan.Win32.Small
Trojan-Dropper.Win32.Delf
Trojan-Spy.Win32.Bancos
Trojan-Spy.Win32.Ardamax

Top 10 2006 PSW
Trojan-PSW.Win32.QQPass
Trojan-PSW.Win32.Delf
Trojan-PSW.Win32.Lineage
Trojan-PSW.Win32.Sinowal
Trojan-PSW.Win32.Nilage
Trojan-PSW.Win32.LdPinch
Trojan-PSW.Win32.Agent
Trojan-PSW.Win32.QQRob
Trojan-PSW.Win32.Lmir
Trojan-PSW.Win32.WOW

Top 10 2006 Backdoor
Backdoor.Win32.Hupigon
Backdoor.Win32.Rbot
Backdoor.Win32.Bifrose
Backdoor.Win32.Delf
Backdoor.Win32.Agent
Backdoor.Win32.SdBot
Backdoor.Win32.Prorat
Backdoor.Win32.VB
Backdoor.Win32.Ciadoor
Backdoor.Win32.IRCBot

Top 10 2006 Worm
Email-Worm.Win32.Luder
Email-Worm.Win32.Warezov
Email-Worm.Win32.Scano
Email-Worm.Win32.Bagle
Email-Worm.Win32.Banwarum
Net-Worm.Win32.Mytob
Worm.Win32.Viking
Email-Worm.Win32.Glowa
Email-Worm.Win32.NetSky
Worm.Win32.Feebs

Top 10 2006 Downloader
Trojan-Downloader.Win32.Zlob
Trojan-Downloader.Win32.Small
Trojan-Downloader.Win32.Agent
Trojan-Downloader.Win32.Obfuscated
Trojan-Downloader.Win32.Tibs
Trojan-Downloader.Win32.Banload
Trojan-Downloader.Win32.Adload
Trojan-Downloader.Win32.Delf
Trojan-Downloader.Win32.VB
Trojan-Downloader.Win32.Ani

Top 10 2006 Rootkit
Rootkit.Win32.Vanti
Rootkit.Win32.Agent
Rootkit.Win32.SMA
Rootkit.Linux.Agent
Rootkit.Win32.Delf
Rootkit.Win32.HideProc
Rootkit.Win32.Fuzen
Rootkit.Win32.PePatch
Rootkit.SunOS.Agent
Rootkit.Win32.Woshi

Top 10 2006 Viruses
Virus.Win32.Parite
Virus.Win32.Virut
Virus.Win32.Hidrag
Virus.DOS.PS-MPC-based
Virus.DOS.Jerusalem
Virus.Win32.Sality
Virus.Win32.Xorala
Virus.DOS.Trivial
Virus.DOS.Leprosy
Virus.DOS.Pixel

2006 Total
W32.Mydoom.M@mm
W32.Netsky.P@mm
W32.Blackmal.E@mm!enc
W32.Erkez.D@mm
W32.Blackmal.E@mm
W32.Erkez.B@mm
W32.Beagle@mm!zip
W32.Mytob.EA@mm
W32.Netsky.Z@mm
W32.Lovgate.R@mm

Web Browser Exploitation Trends

This year, X-Force observed considerable Web browser exploitation through its various Whiro crawlers and analysis of IBM ISS Managed Security Services operational alerting data.

Processing this data and extracting trend information is difficult due to the relationship model the delivery mechanism uses. For example, if there is one site with a particular exploit, but a thousand URLs have been hacked to link to that particular URL, a straight count of one-to-one sites does not work very well.

Recently, the X-Force observed a compromised Web hosting provider that redirected invalid page requests to a dynamically-created exploit JavaScript file which was obfuscated – not encrypted – and utilized two known exploits. In this example, even a dynamically-generated payload file does not lend itself towards a true one-to-one relationship. In the X-Force analysis that follows, items such as most popular exploit vs. most notorious exploit for 2006 will be discussed, as well as encoded exploits, delivery mechanism relationships and browser exploit statistics for Internet Explorer and FireFox on the Windows platform.

Most Popular Exploit

1. **MS04-013 MS-ITS**
2. **MS06-014, RDS.Dataspace ActiveX**
3. **MSFA2005-50, InstallVersion.compareTo Exploit**

Most Notorious Exploit

1. **CVE-2006-1359, IE createTextRange()**
2. **MS06-055, MS VML Vulnerability**
3. **MS06-071, MS XMLHTTP 4.0 ActiveX**

It is interesting that the two most popular Web browser vulnerabilities that were exploited during 2006 are not from 2006 at all. None of the most notorious new vulnerabilities made the top three most popular vulnerabilities to exploit during the year. In fact, the most popular vulnerability in 2006 is actually from 2004. The people behind the malicious Web sites discovered in 2006 must have cause to believe that these patched vulnerabilities are still useful. In one explanation, unless attackers have a true zero-day exploit, only users that regularly patch will apply newly-available protection. Thus, developing exploits for patched vulnerabilities may be a waste of time and therefore many attackers focus on un-patched, older vulnerabilities.

Obfuscation and Encryption

During 2006, X-Force observed a strong increase in Web exploit obfuscation and encryption utilization. Encrypted exploits are contained in streams of encrypted data present in a script such as JavaScript that is decoded on the client's machine and then executed. Obfuscation may be used by an encrypted exploit, but in general it is not. Obfuscated exploits simply are rearranged in a way that makes it difficult for IDS and IPS to match a signature.

Prior to 2006, the use of obfuscated Web-browser exploits were statistically insignificant, and were almost exclusively used in targeted attacks designed to breach known failings in the organization's perimeter security defenses.

The changes in underground exploit sales and the emergence of a market for Web-browser exploits resulted in an emphasis on customizing the exploits for particular customers and bypassing signature-based protection engines.

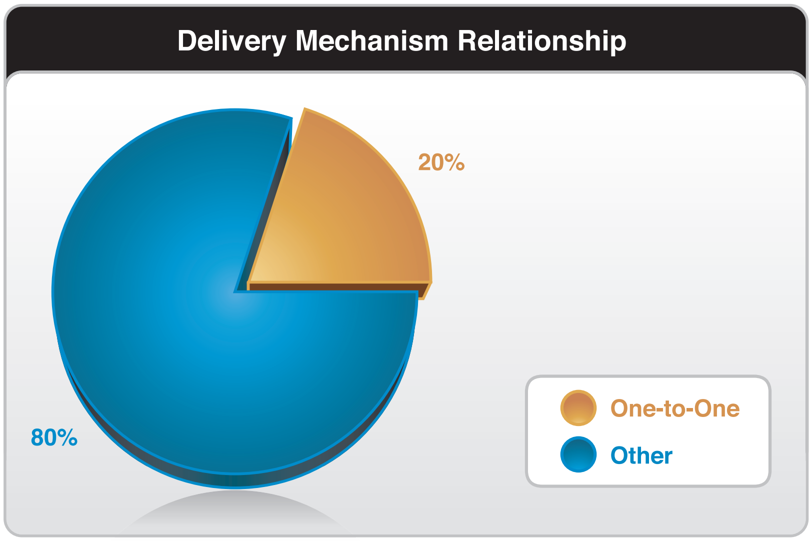
Based upon X-Force analysis, by the end of 2006 there was almost a 50/50 split between the use of obfuscated and non-obfuscated exploits. More importantly, by the end of the year encrypted exploit delivery accounted for 70 percent of in-the-wild exploits.

Delivery Mechanism Types

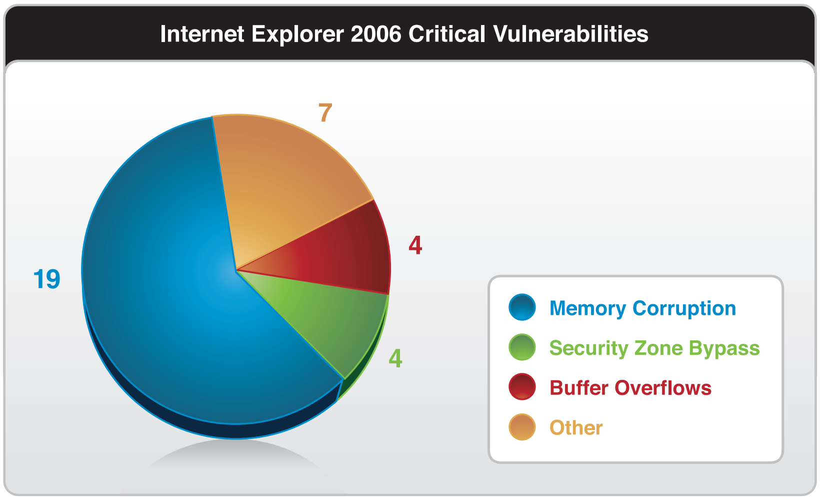
There are various mechanisms by which an attack may be carried out:

1. **One-to-one**
2. **One-to-many**
3. **Many-to-one**

The following graph shows the relative breakdown between one-to-one exploit delivery mechanisms vs. the more complex delivery methods.

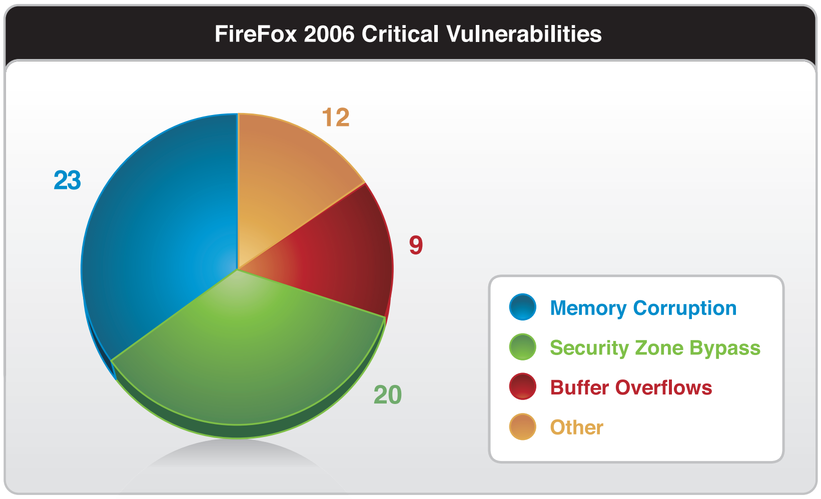


A one-to-one relationship is the traditional exploitation technique where, for each malicious URL, there is one exploit hosted there. A one-to-many relationship is similar in that the origin is a single site but then there are multiple exploits that are hosted. In a many-to-one relationship, there are many sites that point to a single malicious Web site. There is a special case where there are many sites that link to unique URLs with either a one-to-one or one-to-many relationship. Since there is so much overlap, the graph above only presents an estimated view of one-to-one vs. other relationships.



Windows-based Web Browser Wrap-up

Internet Explorer has had 34 critical vulnerabilities reported in 2006. Of these, the majority were memory corruption issues that could lead to remote code execution. With the "heap spray" technique, as previously discussed in the October 2006 X-Force Newsletter, these memory corruption issues are more dangerous than initially thought. Security zone bypass techniques and traditional buffer overflows each account for about 12 percent of the total. The remaining 20 percent is attributed to various other vulnerabilities that can be significant, such as information leakage.



In 2007, X-Force expects to see a continued higher ratio of memory corruption to other vulnerabilities. Even with IE7's enhanced protection, we are still likely to see these affecting plug-in components both offered by Microsoft and third parties. We also expect to see an increase in the "other" category.

FireFox has had 64 critical vulnerabilities reported in 2006. Unlike IE, both memory corruption issues and security zone bypass techniques have been reported in virtually the same amount. Thus while memory corruption issues are still significant problems for FireFox, unlike IE, there is a greater potential for security zone bypass attacks that can often lead to malicious code execution.

GLOBAL HEADQUARTERS

6303 Barfield Road
Atlanta, GA 30328
United States
Phone: (404) 236-2600
e-mail: sales@iss.net

REGIONAL HEADQUARTERS

Australia and New Zealand
Internet Security Systems Pty Ltd.
Level 6, 15 Astor Terrace
Spring Hill Queensland 4000
Australia
Phone: +61 (0)7 3838-1555
Fax: +61 (0)7 3832-4756
e-mail: aus-info@iss.net

Asia Pacific
Internet Security Systems K. K.
JR Tokyu Meguro Bldg. 3-1-1
Kami-Osaki, Shinagawa-ku
Tokyo 141-0021
Japan
Phone: +81 (3) 5740-4050
Fax: +81 (3) 5487-0711
e-mail: jp-sales@iss.net

Europe, Middle East and Africa
Ringlaan 39 bus 5
1853 Strombeek-Bever
Belgium
Phone: +32 (2) 479 67 97
Fax: +32 (2) 479 75 18
e-mail: isseur@iss.net

Latin America
6303 Barfield Road
Atlanta, GA 30328
United States
Phone: (404) 236-2709
Fax: (509) 756-5406
e-mail: isslatam@iss.net

IBM Internet Security Systems
Ahead of the threat.™